

Electronic Journal for Research in **Science & Mathematics Education**

*Flagship Journal of the International Consortium for
Research in Science & Mathematics Education (ICRSME)*



ICRSME

EJRSME

Electronic Journal for Research in Science & Mathematics Education

EDITORS

Kelly Feille, Science Education, *University of Oklahoma*
Jacob Pleasants, Science Education, *University of Oklahoma*
Richard Velasco, Mathematics Education, *University of Florida*

MANAGING EDITOR

Samantha Hittson, *University of Oklahoma*

COPY EDITORS

Sofia Alvarez-Briglie, *University of Oklahoma*
Torrle Epperson, *University of Oklahoma*
Mackenzie Flynn, *University of Oklahoma*
Harriet Lumula, *University of Oklahoma*
Kayla Sutcliffe, *University of Florida*

ASSOCIATE EDITORS

Daniel Alston, *University of North Carolina at Charlotte*
Stacey Britton, *University of West Georgia*
Stephen R. Burgin, *University of Arkansas*
Ryan Fox, *Belmont University*
Heather Gallivan, *University of Northern Iowa*
Gabriela Jonas-Ahrend, *Paderborn University*
Yin-Yin Ko, *Indiana State University*
Christopher S. Long, *University of North Texas*
Stefanie Livers, *Missouri State University*
Eryn Maher, *Georgia Southern University*
Cherie McCollough, *Texas A&M University – Corpus Christi*
Matthew Perkins Coppola, *Purdue University Fort Wayne*
Erin Peters-Burton, *George Mason University*
Laurie Rubel, *University of Haifa*
Sara Salloum, *Ohio University*
Edgar Sintema, *Weizmann Institute of Science*
Karthigeyan Subramania, *University of North Texas*
Robert Wieman, *Ronan University*
Dawn Woods, *Oakland University*
Xiangquan Yao, *Pennsylvania State University*
Ismail Ozgur Zembat, *University of Glasgow*
Issa Ndungo, *Mountains of the Moon University*

EDITORIAL REVIEW BOARD

Gabriel Adeoye, Kwara State University of Education, Ilorin
Franklin Allaire, University of Houston- Downtown
Piata Allen, University of Auckland
Reuben Asempapa, Penn State Harrisburg
Joseph Baidoo, University of Iceland
Eleonor Basilic, Bulacan State University
Francesco Beccuti, Cagliari State University
Pavneet Kaur Bharaj, California State University Bakersfield
Kristen Brown, Texas Christian University
Mila Rosa Carden, University of North Texas
Praveen Chhikara, University of Illinois at Urbana-Champaign
Matthew Clay, Fort Hays State University
Amelia Cook, University of Oklahoma

Richard Cox Jr., Winthrop University
 Yenealem Degu, Kotebe University of Education
 Tamara Diaz Chang, Universidad Austral de Chile
 Nur Banu Duran, Middle East Technical University
 Elizabeth Forde, State University of New York (SUNY) New Paltz
 Ryan Fox, Belmont University
 Rachel Gisewhite, University of Southern Mississippi
 Jenna Gist, Purdue University
 David Goodman, Western Oklahoma State College
 Christopher Irwin, Florida International University
 Benjamin Janney, University of Utah
 Tonya Jeffrey, University of Houston- Downtown
 Austin Jenkins, Purdue University
 Cheryl Johnson, Asbury University
 Delayne Johnson, Delaware State University
 Rachelle Johnson, University of Oklahoma
 Bona Kang, Ohio Wesleyan University
 Firdevs Iclcl Karatas Aydin, Giresun University
 Young Rae Kim, Texas A&M University- San Antonio
 Midhat Noor Kiyani, McGill University
 Lindsay Lightner, Washington State University Tri-cities
 Balagopal Madhu, Regional Institute of Education (NCERT), Mysuru
 Mariam Makramalla, New Giza University
 Kim Megyesi-Brem, Claremont Graduate University
 Duncan Mhakure, University of Cape Town
 Dana Morris, University of Texas-Tyler
 Corey Nagle, CT River Academy
 Tegan Nusser, Bradley University
 Michael Odell, University of Texas at Tyler
 Stephen Ofori, Louisiana State University
 Albolfazl Rafiepour, Shahid Bahonar University of Kerman, Kerman, Iran & Nord university, Norway
 Peter Rillero, Arizona State
 Renata Rodrigues de Matos Oliveira, Seduc (Secretaria de Educaçao de Contagem)/ Universidade Federal de
 Minas Gerais (UFMG)
 Matos Oliverira, Gerais (UFMG)
 Laurie Rubel, University of Haifa
 Rachel Rupnow, Northern Illinois University
 Anchula S J Achari, University of Hyderabad
 Akash Saini, University of Illinois Urbana-Champaign
 Lex Salazar, Texas Christian University
 Wesam Salem, University of Memphis
 Bima Sapkota, The University of Texas Rio Grande Valley
 Laura Schisler, Missouri Southern State University
 Mario Schmiedebach, Bielefeld University, Engelbert-Kaempfer Gymnasium Lemgo
 Rafikh Shaikh, Tata Institute of Social Sciences
 Kristen Shelton, University of Oklahoma
 Edgar Sintema, Weizman Institute of Science
 Karthikeyan SS, Azim Premji Foundation
 Mariam Takkouch, Brock University
 Ozdemir Tiflis, Minister of National Education
 Khahn Tran, Purdue University
 Anaa Wernberg, Malmö university
 Christopher Yarkwah, University of Cape Coast
 Sandra Zuniga Ruiz, San José State University

PUBLISHER

ICRSME, International Consortium for Research in Science & Mathematics Education

CONTENTS

Volume 30 No. 2 | Summer 2026

RESEARCH / EMPIRICAL

Quantitative and Covariational Reasoning in Calculus: Exploring Students' Thinking About Functions and Derivatives	1
Thembinkosi P. Mkhathshwa	
Examining Middle School Mathematics Teachers' Burnout and Self-Efficacy During the Pandemic	25
Sinem Elif Elma, Ekin Balci, Mine İşiksal-Bostan	
High School Students Explore Science Identity Through Participatory Science in Informal Education Program	46
Charlie K. Blake, Andreia F. Dexheimer, Carol E. Colaninno, Candice L. Johnson, Ben K. Greenfield, Lillian C. Ward, Ozaveshe Paul Amune, Adriana E. Martinez, Sharon M. Locke, Georgia Bracey	
A Review of Teaching and Learning Sequences for the Scientific Practice of Modeling	86
Aaron J. Sickel	
Understanding How an Environmental Educator Identifies and Thinks about Environmental Issues in Their Education Programming	116
Hamza Malik and Stephen B. Witzig	
PCK Grain Size in the 21st Century	137
William Veal and Jan Sermeus	
Beyond Problem Solving: Designing Concept-Building Worked Examples for Chemistry Education	165
Markus Emden, Elke Sumfleth, Katrin Schuessler	

ALL RIGHTS RESERVED

© International Consortium for Research in Science & Mathematics Education (ICRSME) Electronic Journal for Research in Science & Mathematics Education (EJRSME)

Quantitative and Covariational Reasoning in Calculus: Exploring Students' Thinking About Functions and Derivatives

Theminkosi P. Mkhathshwa 
Miami University

ABSTRACT

This qualitative study investigated calculus students' quantitative reasoning and covariational reasoning about functions that are situated in real-world non-kinematics contexts. Specifically, the study examined students' thinking about the relationship between a function and its first derivative, second derivative, or the combination of the two derivatives. Furthermore, the study examined students' thinking about open intervals where a function is positive or negative. Analysis of student work and verbal responses given by students while working on two tasks revealed that a majority of the 10 students who participated in the study were able to make sense of the relationship between a function and its first derivative. However, most of the students demonstrated a limited understanding of the relationship between a function and its second derivative. Additionally, most of the students exhibited a limited understanding of what it means for a function to be positive or negative on an open interval. Taken together, the students exhibited weak quantitative and covariational reasoning abilities. This is particularly concerning because calculus often serves as a gateway course for many STEM (Science, Technology, Engineering, and Mathematics) majors. Specifically, weak quantitative and covariational reasoning abilities such as those exhibited by the students in the current study could not only lead to high failure rates in calculus, but also to attrition from STEM majors. Implications for calculus instruction are included.

Keywords: Quantitative reasoning, covariational reasoning, functions, derivatives, calculus

Introduction

Oehrtman et al. (2008) argued that “the concept of function is central to undergraduate mathematics, foundational to modern mathematics, and essential in related areas of the sciences” (p. 27). On a related note, the concept of function is essential to the study of first-semester calculus (Calculus I) in the United States. Much research has reported on students' thinking about the concept of function over the last few decades (Beichner, 1994; Carlson, 1998; Dubinsky & Wilson, 2013; Habre & Abboud, 2006; Haciomeroglu et al., 2010; Malambo et al., 2018; Martínez-Plunell & Gaisman, 2012; Melhuish et al., 2020; Monk, 1992; Monk & Nemirovsky, 1994; Trujillo et al., 2023; Yerushalmy & Swidan, 2012). In fact, some of this research has examined students' thinking about the relationship between a function and its first derivative or the relationship between a function and its second derivative (Burns-Childers & Vidakovic, 2018; Carlson et al., 2002; Çetin, 2009; García-García & Dolores-Flores, 2021; Tsamir & Ovodenko, 2013). There is, however, a dearth of research that has investigated students' thinking about the relationship between a function's first derivative and the function's second derivative, as well as research that has examined students' understanding of open intervals where a function is positive or negative, which is the primary motivation for the present

study. Understanding open intervals where a function is positive or negative is a fundamental skill, as it is foundational for understanding calculus ideas such as open intervals where a function is increasing or decreasing. Additionally, this study examined students' ability to sketch the graph of the second derivative when presented with the graph of a function, an area of research that has not been well explored.

Most of the existing literature on students' thinking about functions or their derivatives is situated in motion contexts (e.g., Altindis, 2025; Carlson et al., 2002; Kertil, 2020; Mkhathswa, 2020; Nagle et al., 2017; Nguyen et al., 2025; Swidan et al., 2019). A growing number of researchers have called for the examination of students' thinking about derivatives in non-kinematics contexts (e.g., Biza et al., 2022; Jones, 2017; Kertil & Dede, 2022). In fact, Jones (2017) argued that "given the wide range of applications in science and engineering that are not based on kinematics ... it is important to know how students understand applied derivatives in non-kinematics contexts." (p. 95). In an effort to contribute toward addressing the gap in knowledge identified by Jones (2017), the present study examined students' thinking about derivatives in two non-kinematics contexts, namely volume of water in a tank and the spread of a disease. This examination takes place through the lens of quantitative reasoning and covariational reasoning, two modes of reasoning that have long been identified by influential mathematics educators as essential for students hoping to understand quantities and relationships among quantities that can be represented using functions (Carlson et al., 2002; Thompson, 2011).

The primary goal of the study was to gain insight into students' quantitative reasoning and covariational reasoning in the context of working with functions situated in real-world, non-kinematics contexts. The secondary goal of the study was to gain insight into students' understanding of open interval where a function is positive, negative, or zero. Thus, taken together, the present study aims to answer the following research questions:

1. What does calculus students' thinking about the relationship between a function and its first and second derivative reveal about their quantitative reasoning?
2. What does calculus students' thinking about the relationship between a function and its first and second derivative reveal about their covariational reasoning?
3. What does calculus students' thinking about the relationship between a function and its first and second derivative reveal about their understanding of open intervals where a function is positive, negative, or zero?

Theoretical Underpinnings: Quantitative and Covariational Reasoning

This study uses the theoretical lenses of quantitative reasoning and covariational reasoning.

Quantitative reasoning

"Quantity" refers to a measurable attribute of an object (Thompson, 1993; 1994b). A water tank is one example of an object used in the present study, while the volume of the tank is an example of a quantity used in the same study. A key element of quantitative reasoning involves making sense of quantities or relationships between quantities (Smith & Thompson, 2007; Thompson, 1993). Making sense of quantities entails assigning numerical values to quantities, a process Thompson (1994d) refers to as quantification. Additionally, making sense of quantities entails reasoning about their units of measure. Furthermore, interpreting quantities is an important aspect of quantitative reasoning.

Covariational reasoning

“Covariational reasoning” refers to how two variables are changing in tandem in a dynamic situation or event. Carlson et al. (2002) described five levels of covariational reasoning (coordination, direction, quantitative coordination, average rate, and instantaneous rate) that are based on the mental actions reproduced in Table 1.

Table 1. Mental Actions of Covariational Reasoning (Reproduced from Carlson et al., 2002, p. 357)

Mental action	Description of mental action	Behaviors
Mental Action 1 (MA1)	Coordinating the value of one variable with changes in the other [variable].	Labeling the axes with verbal indications of coordinating the two variables (e.g., y changes with changes in x).
Mental Action 2 (MA2)	Coordinating the direction of change of one variable with changes in the other variable.	- Constructing an increasing straight line. - Verbalizing an awareness of the direction of change of the output while considering changes in the input.
Mental Action 3 (MA3)	Coordinating the amount of change of one variable with changes in the other variable.	- Plotting points/constructing secant lines. - Verbalizing an awareness of the amount of change of the output while considering changes in the input.
Mental Action 4 (MA4)	Coordinating the average-rate - of change of the function with uniform increments of change in the input variable.	- Constructing contiguous secant lines for the domain. - Verbalizing an awareness of the rate of change of the output (with respect to the input) while considering uniform increments of the input.
Mental Action 5 (MA5)	Coordinating the instantaneous rate of change of the function with continuous changes in the independent variable for the entire domain of the function.	- Constructing a smooth curve with clear indications of concavity changes. - Verbalizing an awareness of the instantaneous changes in the rate of change for the entire domain of the function (direction of concavities and inflection points are correct).

The aforementioned levels of covariational reasoning are increasingly sophisticated in the sense that a student engages in the lowest level of covariational reasoning (coordination) if they exhibit behaviors associated with the first mental action (MA1), while engaging at the second level of covariational reasoning (direction) requires exhibiting behaviors associated with the first two mental actions (MA1 and MA2), and so forth.

Related Literature

This section provides an overview of the of the existing literature that is related to the present study, namely students' understanding of functions, students' understanding of the first derivative, and students' understanding of the second derivative.

Students' understanding of functions

The importance of functions at the K-16 level has been underscored by several influential bodies of education and reputable mathematics education researchers (Dubinsky & Wilson, 2013; Monk, 1994; National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010; National Council of Teachers of Mathematics, 2000; Oehrtman et al., 2008; Thompson, 1994c). One reoccurring theme from the literature that has examined students' thinking about functions is that in as much as functions can be represented in multiple ways such as using graphs, numerical tables of values, or equations/formulas, students have a proclivity to work with functions in algebraic form (Habre & Abboud, 2006; Herman, 2007; Ibrahim & Rebello, 2012; Thompson, 1994b). On a related note, Thompson (1994c) posited that while "tables, graphs and expressions might be multiple representations of functions" (p. 23) from the perspective of mathematics instructors or researchers, they are not "multiple representations of anything to students" (p. 23). Another common finding from research that has investigated students' thinking about functions is that making sense of graphs of functions or creating graphs of functions is often a challenge for students (Beichner, 1994; Dominguez et al., 2017; Carlson et al., 2002; Monk, 1994). In a similar vein, other research indicates that students tend to envision graphs of functions as pictures of physical events, an idea that Monk (1992) refers to as iconic translation. Lastly, findings from research on students' reasoning about functions indicate that translating information from one representation of a function to another, such as selecting the correct kinematics graph that matches a given textual description about a functional situation, is sometimes problematic for students (Beichner, 1994; Martínez-Planell & Gaisman, 2012; Van den Eynde et al., 2019).

Students' understanding of the first derivative

A number of studies have reported on students' thinking about the first derivative. One major theme that emerges from this research is that students are prone to confusing a function's instantaneous rate of a change with either the function's output value or the accumulation of the function's output (Jones, 2017; Mkhathswa, 2020, 2023; Mkhathswa & Doerr, 2018; Prince et al., 2012; Rasmussen & Marrongelle, 2006). For instance, Mkhathswa and Doerr (2018) reported on students who confused marginal cost (the rate of change of cost) with total cost (the accumulation of cost), while Rasmussen and Marrongelle (2006) reported on students who did not make a conceptual distinction between the rate of change in the amount of salt and the amount of salt. On a related note, other findings indicate that students have a penchant for confusing units of derivative quantities such as speed with units of amount quantities such as distance (Lobato et al., 2012; Mkhathswa & Doerr, 2018; Mkhathswa, 2020). Other research has reported on students' difficulties related to sketching the

graph of the first-derivative function when given the graph of a function (García-García & Dolores-Flores, 2021; Orhun, 2012; Ubuz, 2007).

Students' understanding of the second derivative

Research on students' thinking about the second derivative is limited. A common finding of the few available studies is that making sense of inflection points is not only challenging for calculus students, but also for students who have completed a course in ordinary differential equations (Carlson et al., 2002; Jones, 2019; Ovodenko & Kouropatov, 2019; Tsamir & Ovodenko, 2013). Carlson et al. (2002) reported on students who had “difficulty forming images of continuously changing rate and [students who] could not accurately represent or interpret inflection points or increasing and decreasing rate for dynamic situations” (p. 352). Some of the students in Tsamir and Ovodenko's (2013) study confused inflection points with local extrema, an observation also made by Jones (2019). Additionally, none of the 52 students who participated in Tsamir and Ovodenko's (2013) study “correctly identified all inflection points on the five graphs” (p. 416) of the functions they were presented with. Jones (2019) reported that some of the students who participated in his study “struggled to use appropriate aspects of covariational reasoning to make sense of concavity and inflection points” (p. 541) in various real-world contexts, including economics. On a positive note, Ovodenko and Kouropatov (2019) proposed a teaching unit that has the potential to address common student misconceptions or errors as it relates to the concept of an inflection point. The unit utilizes several technological and educational tools, including GeoGebra, a dynamic mathematics software.

In summary, the aforementioned bodies of literature have established that the concept of function is generally not well-understood by students. Specific student challenges include making sense of functions in multiple representations and a proclivity for working with functions in algebraic form. There is, however, a dearth of research that has explored students' thinking about functions in connection to open intervals where functions are positive, negative, or zero, a knowledge gap that the present study seeks to address in the literature. The concept of open intervals where a function is positive, negative, or zero is worthy of greater attention because it is foundational for understanding fundamental concepts or topics in the study of first-semester calculus, including optimization problems, curve sketching, critical points of a function, the first derivative test, and the intermediate value theorem. Furthermore, the literature has established that students have difficulties making sense of relationships between functions and their derivatives, especially in motion contexts. The present study examines students' thinking about functions and their derivatives in real-world and non-motion contexts to understand how students' reasoning in these contexts might inform the field of calculus education. Finally, there is a paucity of research that has examined how students make sense of the relationship between a function's first derivative and a function's second derivative, and how students sketch graphs of second derivative functions when given the graph of a function. The present study uses the lens of quantitative reasoning and covariational reasoning to address the previously noted knowledge gaps in the literature.

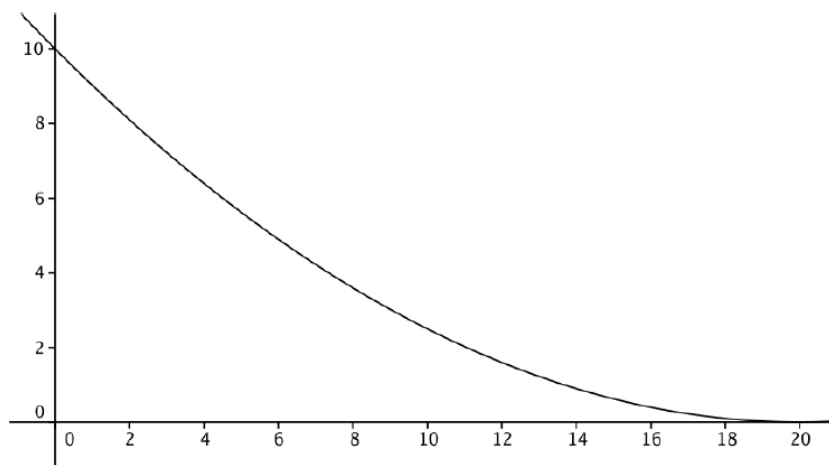
Methodology

Research design

This exploratory study utilized a qualitative research design (Creswell, 2014). Specifically, to answer the research questions listed in the Introduction section, I used structured task-based interviews (Goldin, 2000). According to Goldin, “the design of structured task-based interviews needs to take into account their research purposes” (p. 519). He added, the research purposes “may include (for example) exploratory investigation; refinement of observation, description, inference, or analysis

techniques; development of constructs and conjectures; investigation or testing of advance hypotheses; and/or inquiry into the applicability of a model of teaching, learning, or problem solving.” (p. 519). In the present study, I used structured task-based interviews with the sole purpose of exploring students’ quantitative or covariational reasoning in the context of working with functions and their derivatives in real-world non-motion contexts. In designing or conducting the task-based interviews, I followed key principles and techniques outlined by Goldin (2000), including choosing “tasks that are accessible to the participants” (p. 540), designing “task-based interviews to address advance research question” (p.540), and encouraging “free problem solving” (p. 542). The interviews covered two tasks:

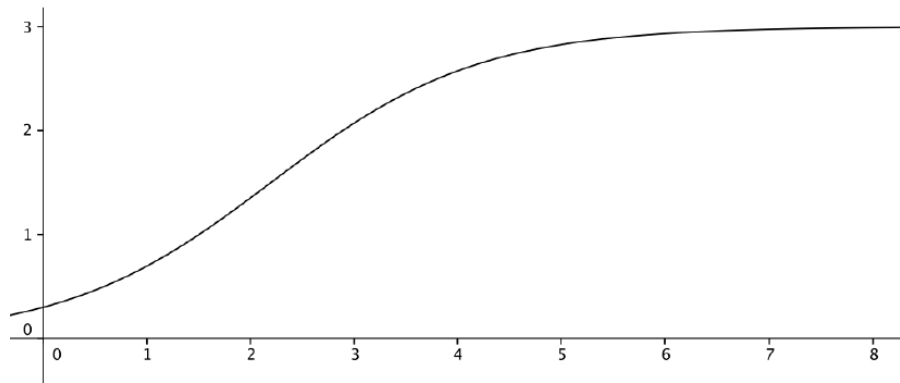
Task 1 (Volume of water in a tank): The graph shows the volume $f(t)$ (in gallons) of water in a tank t seconds after a drain hole is unplugged.



- a) What parts of the domain of f correspond to f being positive, negative, or zero? Explain.
- b) Based on your observations in part a), how can you interpret f in the context of volume of water in the tank and time? Explain.
- c) What parts of the domain of f correspond to f' being positive, negative, or zero? Explain.
- d) Based on your observations in part c), what does f' tell you about f in the context of volume of water in the tank and time? Explain.
- e) What parts of the domain of f correspond to f'' being positive, negative, or zero? Explain.
- f) Based on your observations in part e), what does f'' tell you about f in the context of volume of water in the tank and time?
- g) Considering the domain of f , what does f'' tell you about f' in the context of volume of water in the tank and time?

- h) What are the units of the quantities represented by f' and f'' ?
- i) Construct a rough sketch of the graph of f' .
- j) Construct a rough sketch of the graph of f'' .

Task 2 (Spread of a disease): The graph shows the number $f(t)$ (in hundreds) of laboratory mice that have become sick with the flu by the t^{th} day after the start of an experiment.



- a) What parts of the domain of f correspond to f being positive, negative, or zero? Explain.
- b) Based on your observations in part a), how can you interpret f in the context of the spread of the disease and time? Explain.
- c) What parts of the domain of f correspond to f' being positive, negative, or zero? Explain.
- d) Based on your observations in part c), what does f' tell you about f in the context of the spread of the disease and time? Explain.
- e) What parts of the domain of f correspond to f'' being positive, negative, or zero? Explain.
- f) Based on your observations in part e), what does f'' tell you about f in the context of the spread of the disease and time?
- g) Considering the domain of f , what does f'' tell you about f' in the context of the spread of the disease and time?
- h) What are the units of the quantities represented by f' and f'' ?
- i) Construct a rough sketch of the graph of f' .
- j) Construct a rough sketch of the graph of f'' .

Each participant completed Task 1 prior to working on Task 2. At the conclusion of each task, the interviewer asked the following questions, in addition to clarifying questions (For example, what do you mean when you say f' is decreasing with a negative slope?) that were asked by the interviewer as students worked through Tasks 1 and 2: What was the easiest part when solving this problem? Explain. What was the difficult part when solving this problem? Explain.

The purpose of each task was four-fold. First, to examine students' reasoning about the relationship between a function (f) and its first derivative (f'). Second, to examine students' reasoning about the relationship between a function (f) and its second derivative (f''). Third, to examine students' reasoning about the relationship between the first derivative of a function (f') and the second derivative of a function (f''). Fourth, to examine students' reasoning about open intervals where a function is positive or negative. Average completion task time for each interview was 36 minutes (19 minutes on Task 1 and 17 minutes on Task 2). The shortest interview lasted for eight minutes while the longest interview lasted for 34 minutes.

Setting and participants

This study took place at a public research university in the United States. The subjects were 10 students (Pseudonyms Dean, Mary, Claire, Camila, Cooper, Nate, Kiara, Ken, Adrian, and Ada) who were enrolled in three sections of a traditional Calculus I course during the fall semester of 2022. The three sections were taught by two seasoned professors. It is worth noting that the researcher is not one of the professors. The reader may find additional information about the course and study participants from an earlier manuscript (Mkhatshwa, 2025a).

Data collection and analysis

There are two sources of data for the current study, namely transcriptions of interviews and work written by students during each interview session. The data were analyzed using a thematic analysis approach (Braun & Clark, 2006). According to Braun and Clark, thematic analysis is a method “for identifying, analysing and reporting patterns (themes) within data” (p. 79). In the context of the present study, a theme is a similar response to the prompts included in Tasks 1 and 2 that was given by at least two students. Specifically, the data were analyzed in three stages. In the first stage of the analysis, I examined students' quantitative reasoning. Specifically, I examined how students interpreted the quantities represented by the functions f , f' , and f'' in the contexts of the two tasks. Furthermore, I examined how students reasoned about the units of the quantities represented by the functions f , f' , and f'' . Additionally, I examined how students reasoned about the relationship between the quantities represented by the functions f' and f'' . In the second stage of the analysis, I examined students' covariational reasoning. In particular, I examined how students engaged (or did not engage) in the five levels of covariational reasoning described in the theoretical underpinnings section, namely coordination, direction, quantitative coordination, average rate, and instantaneous rate. In the third stage of the analysis, I examined how students reasoned about open intervals where the functions f , f' , and f'' are positive, negative, or zero (i.e., neither positive nor negative). Table 2 shows examples of how the codes were applied.

Table 2

Sample codes from the data

Code	Description of Code	Examples
Interpreting (an aspect of quantitative reasoning) f' or f'' as a rate of change.	Using the word “per” when prompted to assign units of measure to the quantities represented by the functions f' or f'' in Task 1 or Task 2.	Researcher: What are the units of the quantities represented by f' [Task 1] and f'' [Task 1]? Dean: The units of f' is gallons per second and the units of f'' is gallons per second per second.
Attention to direction of change (the second level of covariational reasoning).	Describing how two quantities are changing simultaneously in Task 1 or Task 2.	Researcher: Based on your observations in part a) [Task 2], how can you interpret f in the context of the spread of the disease and time? Explain. Ken: As time goes on, more mice get sick.
Reasoning about positive functions.	Describing intervals where at least one of the functions f , f' , or f'' in Task 1 or Task 2 is positive.	Researcher: What parts of the domain of f [Task 1] correspond to f'' being positive, negative, or zero? Explain. Mary: Between $t = 0$ and $t = 20$, the second derivative [f''] will be positive because the graph of $f'(t)$ is increasing as the slope of $f(t)$ decreases.

To ensure rigor and trustworthiness of the data analysis, I used peer debriefing (Spall, 1998). According to Spall, “peer debriefing supports the credibility of the data in qualitative research and provides a means toward the establishment of the overall trustworthiness of the findings” (p. 280). In the context of the present study, peer debriefing entailed discussing the research design, the theoretical underpinnings for the study, the tasks, and the emergent themes (findings) from the study with one mathematics education researcher. Discrepancies in emergent themes were resolved as part of the peer debriefing process. Permission (Project Reference Number: 04117e) to conduct this study was granted by Miami University's Research Ethics & Integrity Program.

Results

Students' quantitative reasoning

Overall, students' interpretations of the quantities represented by the functions f , f' , and f'' across the two tasks suggests that while students are able to correctly interpret quantities represented by first derivatives in real-world non-kinematics contexts, interpreting quantities represented by second derivatives in the same contexts is particularly challenging for students. Additionally, analysis of students' work across the two tasks suggests that interpreting the relationship between the quantities represented by f' and f'' in each of the two tasks was problematic for a majority of the students who participated in this study. Table 3 summarizes students' interpretations of the quantities represented by the functions f , f' , and f'' in the two tasks. The symbol ✓ indicates that a student correctly interpreted the indicated quantity (or relationship between quantities) and the symbol X means that a

student incorrectly interpreted the indicated quantity (or relationship between quantities). Additionally, the symbol IDK indicates that a student stated that they “did not know” how to interpret the indicated quantity (or relationship between quantities). Furthermore, the symbol T1 indicates that the interpretation of the indicated quantity (or relationship between quantities) pertains to Task 1 whereas the symbol T2 indicates that the interpretation of the indicated quantity (or relationship between quantities) pertains to Task 2.

Table 3

A Summary of Students' Interpretations of Quantities Represented by the Functions f , f' , and f''

Student	f		f'		f''		Relationship between f' and f''	
	T1	T2	T1	T2	T1	T2	T1	T2
Dean	✓	✓	✓	✓	✓	X	✓	✓
Mary	✓	✓	X	X	X	X	X	X
Claire	✓	X	✓	X	X	X	X	X
Camila	✓	✓	IDK	X	IDK	X	IDK	IDK
Cooper	✓	✓	✓	X	X	✓	X	X
Nate	X	✓	X	X	X	X	X	X
Kiara	✓	X	✓	X	X	X	X	X
Ken	✓	✓	✓	✓	X	X	X	✓
Adrian	✓	X	✓	✓	X	X	X	X
Ada	✓	X	✓	✓	IDK	X	IDK	X

As can be seen in Table 3, nine of the study participants correctly interpreted the quantity represented by the function f in Task 1 compared to six students who correctly interpreted the quantity represented by the function f in Task 2. To some extent, it can be argued that students' correct interpretations of the function f in both tasks is due to the fact that an explanation of the function f was provided as part of the description of each task.

Six students correctly interpreted the quantity represented by the function f' in Task 1 as the rate at which the volume of water is decreasing in the tank after the drain hole is unplugged, compared to four students who correctly interpreted the quantity represented by the function f' in Task 2 as the rate at which the number of sick mice is increasing over time. As an example of one of the incorrect interpretations given by students when prompted to explain the meaning of f' in the context of Task 1 (i.e., water flowing out through a drain hole), Nate remarked that “a first derivative [referring to the function f' in Task 1] would tell you the critical values where a function is increasing or decreasing, and the local maximum and minimum values of the graph [that appears in Task 1].” As an example of

one of the incorrect interpretations given by students when prompted to explain the meaning of f' in the context of Task 2 (i.e., the spread of a disease), Kiara commented that the first derivative tells us that “as time increases, the disease also increases until it remains somewhat constant.” It would seem that in her interpretation of the function f' in Task 2, Kiara paid attention to the physical features of the graph given in the task.

Two students correctly interpreted the quantities represented by the function f'' in both tasks. In particular, one student (Dean) correctly interpreted the quantity represented by the function f'' in Task 1 as how quickly the rate of water loss in the tank is changing over time and another student (Cooper) correctly interpreted the quantity represented by the function f'' in Task 2 as how quickly the rate of sick mice is changing over time. As an example of one of the incorrect interpretations given by students when prompted to explain the meaning of f'' in the context of Task 1, Adrian explained that the function f'' tells us that “the volume of water in the tank decreases as time increases.” Like Kiara, it would seem that Adrian paid attention to the physical features of the graph presented in Task 2 when interpreting the quantity represented by the function f'' in the task. As an example of one of the incorrect interpretations given by students when prompted to explain the meaning of f'' in the context of Task 2, Claire noted that the function f'' tells us that “as time goes on, more mice will get sick.”

Two students correctly interpreted the relationship between the quantities represented by the functions f' and f'' in at least one of the two tasks. One of these students (Dean) correctly interpreted the relationship between the quantities represented by the functions f' and f'' in both tasks. When prompted to explain what f'' tells us about f' in the context of Task 1, Dean explained that “the speed of the decrease of water is slowing down, which means that the volume of water in the tank decreases slower and slower in a certain time period.” Similarly, when prompted to explain what f'' tells us about f' in the context of Task 2, he reasoned that “the speed of the spread of the disease increases at the beginning and decreases from the second day [of the outbreak].” Most of the students who incorrectly interpreted the relationship between the quantities represented by the functions f'' and f' either remarked about the physical features of the graphs shown in each task, or they provided responses that I coded as not making sense. One such response that I coded as not making sense was given by Nate who reasoned that “the f'' in context with f' would give you a better representation of the relationship between volume of water in the task [Task 1] and time.” In response to the same prompt regarding Task 2, Nate’s response suggest that he paid attention to the physical features of the graph given in the task in an effort to make sense of the relationship between the quantities represented by the functions f'' and f' . Specifically, he explained that “the f'' tells you about f' that as the number of days increase, the number of sick mice increases reaching a peak on the eighth day.”

Only four students correctly determined the units of measure for the quantity represented by the function f' in Task 1 as “gallons per second”, compared to only two students who correctly determined the units of measure for the quantity represented by the function f' in Task 2 as “sick mice per day.” Common incorrect units of measure assigned by students to the quantity represented by the function f' in Task 1 include “gallons” or “gallons per second squared”, whereas common incorrect units of measure assigned by students to the quantity represented by the function f' in Task 2 include “number of sick mice” or simply “number of mice.” Likewise, four students correctly determined the units of measure for the quantity represented by the function f'' in Task 1 as “gallons per second squared”, compared to two students who correctly determined the units of measure for the quantity represented by the function f'' in Task 2 as “sick mice per day squared.” Typical incorrect units of measure assigned by students to the quantity represented by the function f'' in Task 1 include “gallons” or “time in seconds” whereas typical incorrect units of measure assigned by students to the

quantity represented by the function f'' in Task 2 include “hundreds of mice per day” or “number of mice.”

Students' covariational reasoning

In assessing students' covariational reasoning, I focused on how students constructed first-derivative graphs and second-derivative graphs corresponding to the graphs of the functions that are presented in each of the two tasks. Because students' covariational reasoning is similar in each task, regardless of which derivative graph (first or second) they constructed, I limit my reporting in this section to how students engaged (or did not engage) in covariational reasoning while creating first-derivative graphs corresponding to the functions given in the two tasks.

An examination of students' covariational reasoning in the two tasks revealed that five students engaged in the highest level of covariational reasoning (i.e., instantaneous rate) while working on Task 1, compared to only three students who engaged in the same level of covariational reasoning while working on Task 2. Ada's graph (shown in Figure 1) is representative of students' graphs that I coded as showing evidence of engaging in the highest level of covariational reasoning in Task 1, while Dean's graph (shown in Figure 2) is representative of students' graphs that I coded as showing evidence of engaging in the highest level of covariational reasoning in Task 2.

Figure 1

Ada's First-Derivative Graph in Task 1

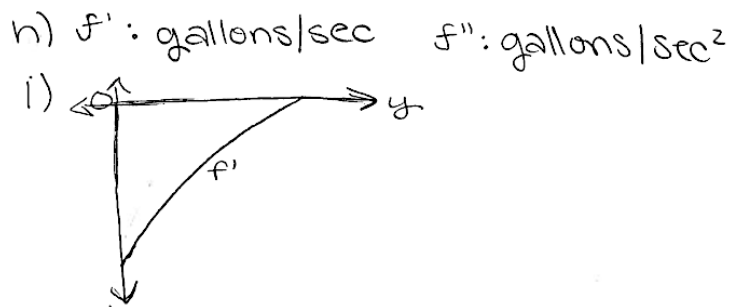
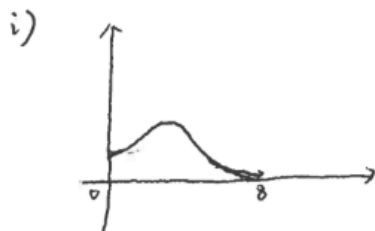


Figure 2

Dean's First-Derivative Graph in Task 2

h) the unit of f' is number/days
the unit of f'' is number/(days)²



Ada and Dean created vertical and horizontal axes for their graphs, which is evidence of engaging in the first level of covariational reasoning (i.e., coordination). I note that even though Dean's axes are not labeled, the student verbally and correctly stated that the quantity [time] on the horizontal axis would be measured in seconds while the quantity [hundreds of sick mice per day] on the vertical axis would be measured in number [of sick mice] per day. Ada correctly stated that the units of the quantity [time] on the horizontal axis would be measured in seconds, while that of the quantity [flow rate] on the vertical axis would be measured in gallons per second. Both students expressed verbal awareness of the quantities represented in the vertical axes of their graphs as changing (i.e., decreasing in the case of Task 1 and increasing in the case of Task 2) with time, thus exhibiting evidence of engaging in the second level of covariational reasoning (i.e., direction). For example, while commenting on the domain of the graph of the function given in Task 1, Ada remarked that "as time progresses, the volume of water in the tank decreases." The students, especially Dean, plotted a few points that they initially joined with contiguous line segments prior to using smooth curves to join the points, thus providing evidence of engaging in levels 3, 4, and 5 of covariational reasoning. Specifically, the plotting of points is evidence of engaging in the third level of covariational reasoning also known as the quantitative coordination level. Furthermore, the fact that the students initially joined the points they plotted using "contiguous secant lines" is evidence of engaging in the fourth level of covariational reasoning (also known as the average rate level). Additionally, the students engaged in the highest level of covariational reasoning (i.e., the fifth level of covariational reasoning, also known as the instantaneous rate level). For instance, when asked to explain the meaning of the function f' in Task 2, Dean remarked that "the number of infected mice is increasing rapidly at the very beginning days and then the number of infected mice increases slower." He added, "there are fewer mice being infected after the eighth day of the experiment." I argue that Dean's attention to how the number of mice infected grows rapidly initially and then slows down after the eighth day of the experiment demonstrates the student's awareness of the instantaneous rate of change of the function considered in the task. Moreover, the student's graph [Figure 2] is a smooth curve with clear indications of concavity changes, which is yet another evidence of engaging in the highest level of covariational reasoning.

The rest of the students, at best, engaged at the second level of covariational reasoning while working on the two tasks. Nate's graph (shown in Figure 3) is representative of students' graphs that I coded as showing evidence of engaging in the first or second level of covariational reasoning in Task 1, while Cooper's graph (shown in Figure 4) is representative of students' graphs that I coded as showing evidence of engaging in the first or second level of covariational reasoning in Task 2.

Figure 3

Nate's First-Derivative Graph in Task 1

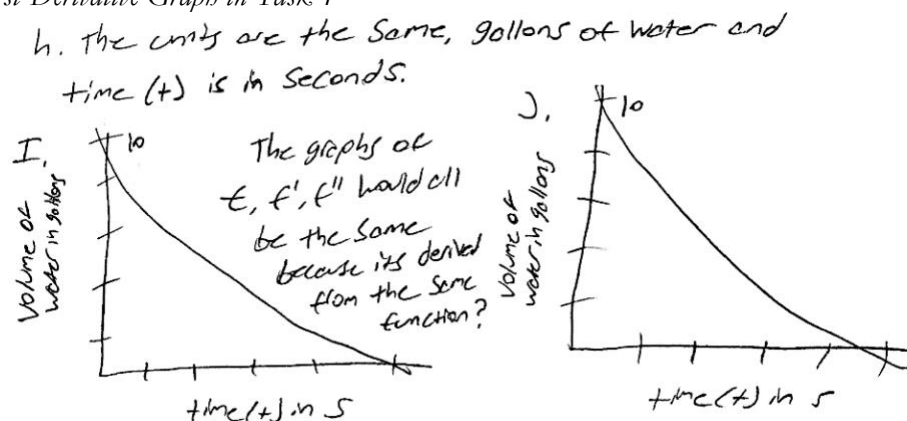
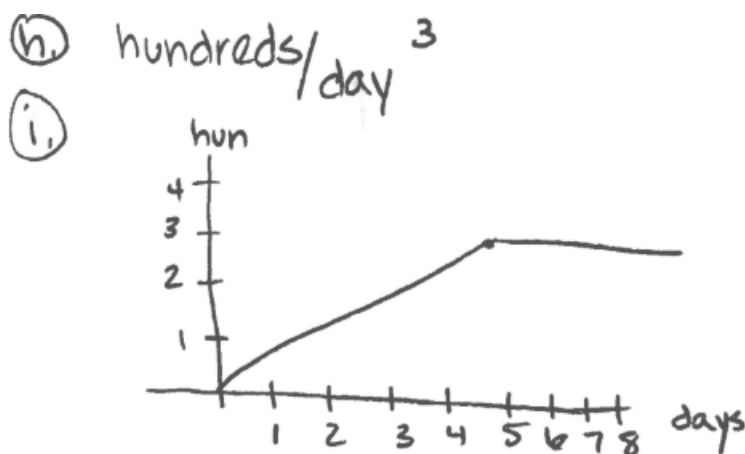


Figure 4*Cooper's First-Derivative Graph in Task 2*

The fact that both Nate and Cooper created and labeled horizontal and vertical axis for their graphs demonstrates an awareness that two quantities are changing in tandem in each task (e.g., the volume of water and time in Task 1). This is evidence of engaging in the first level of covariational reasoning. Since Nate's first-derivative graph (i.e., the graph on the left and labeled with the variable I in Figure 3) appears to be the same as the graph that is presented in the task (i.e., the original function), I judged his graph as not exceeding the first level of covariational reasoning. On a related note, I note that Nate's first-derivative graph has the same units of measure as that of the original graph, further suggesting that his first-derivative graph is simply a reproduction of the original graph. While not relevant, I note that the graph on the left in Figure 3 is Nate's second-derivative graph which also appears to be a reproduction of the original graph. When prompted to talk about the challenges he encountered while working on Tasks 1 and 2, Nate repeatedly remarked on the difficulty of sketching the graphs of f' and f'' in the aforementioned tasks:

Researcher: What was the difficult part when solving this problem [Task 1]? Explain.

Nate: The difficult part was not having an equation for $f(x)$, and so it's hard to state what the context or relationship between f , f' , and f'' would be and what the graphs of $f'(x)$ and $f''(x)$ would look like.

Researcher: What was the difficult part when solving this problem [Task 2]? Explain.

Nate: The difficult part was that there was no equation for $f(x)$, so it was difficult to determine the sketches of the graphs for $f'(x)$ and $f''(x)$, so I assumed they [graphs of $f'(x)$ and $f''(x)$] were the same [as the graph of $f(x)$].

A close examination of Nate's difficulties as it relates to sketching the graphs of f' and f'' in each task reveals that his difficulties were two-fold. First, Nate repeatedly remarked on not having an equation for the functions f . Second, Nate was not able to make sense of how the functions f , f' , and f'' were all related in the context of each of the tasks, especially in Task 1. This is indicated by his

claim, “so it’s hard to state what the context or relationship between f , f' and f'' would be” when he commented on how the graphs of f' and f'' would look like in Task 1.

When asked to sketch the graph of the second derivative in each task, a majority of the students either reproduced the graph they sketched for the first derivative, reproduced the graph of the original function given in each task, or they stated that they “did not know how to sketch the graph of the second derivative.” One of the students, Claire, claimed that she struggled with envisioning how the graph of f'' would look like. To substantiate her claim, she reasoned, “we have not spent enough time in class graphing them [graphs of f''].” I considered Cooper’s use of two connected line segments in Figure 4 (one increasing and another constant) as evidence of engaging in the second level of covariational reasoning. Specifically, the increasing line segment demonstrates that the student recognized that at the beginning of the outbreak (within the first four and a half days of the outbreak per the evidence contained in Figure 4), the number of sick mice would increase. This is evidence of engaging in the direction of change level (i.e., the second level of covariational reasoning).

Students’ reasoning about open intervals where a function is positive, negative, or zero

I begin this section by noting that the domain of each of the functions (f , f' , and f'') in Task 1 is considered to be the open interval $(0,20)$, while the domain of each of the functions (f , f' , and f'') in Task 2 is considered to be the open interval $(0,8)$. I further note that with regard to Task 1, the functions f and f'' are considered to be positive on the open interval $(0,20)$, while the function f' is considered to be negative on the open interval $(0,20)$. With regard to Task 2, the functions f and f' are considered to be positive on the open interval $(0,8)$, while the function f'' is considered to be positive on the open interval $(0,2.5)$ and negative on the open interval $(2.5,8)$. Furthermore, the function f'' in Task 2 is the only function that is considered to be zero (neither positive nor negative) in some part of its domain. In particular, this happens when the input variable t takes on the value of 2.5 (i.e., when $t = 2.5$). I remark that only four students correctly determined the input value of $t = 2.5$ as a location in the graph of the function f'' where the same function would be considered to be zero.

Analysis of students’ understanding of intervals over which the functions f , f' , and f'' in each task are positive, negative, or zero revealed that at least half of the students correctly determined the open intervals over which the functions f are positive or negative in each task, and that a majority of the students were not able to correctly identify open intervals over which the functions f' and f'' are positive or negative in each task. Specifically, seven students correctly determined the open interval over which the function f is positive or negative in Task 1, compared to five students who correctly determined the open interval over which the function f is positive or negative in Task 2. Additionally, six students correctly determined the open interval over which the function f' is positive or negative in Task 1, compared to only four students who correctly determined the open interval over which the function f' is positive or negative in Task 2. Furthermore, only 4 students correctly determined the open interval over which the function f'' is positive or negative in Task 1, compared to only three students who correctly determined the open intervals over which the function f'' is positive or negative in Task 2. The following excerpt highlights some of the challenges that students cited as stumbling blocks in their attempt to determine open interval where the functions f , f' , and f'' in each task are positive or negative.

- Researcher: What was the easiest part when solving this problem [Task 1]? Explain.
- Kiara: Determining the relationship between f , f' , and f'' to the volume of water and time because I was able to look at the graph of what is happening [Referring to the graph of the function f that is given as part of the task].
- Researcher: What was the difficult part when solving this problem [Task 1]? Explain.
- Kiara: Determining the parts of the domain corresponding to f [the function given on the task] being negative, positive, or zero because there was no equation given.
- Researcher: What was the easiest part when solving this problem [Task 2]? Explain.
- Kiara: Relating the number of sick mice to the number of days past because it was represented by the graph [i.e., understanding how the number of sick mice changed over time based on the information given on the graph].
- Researcher: What was the difficult part when solving this problem [Task 2]? Explain.
- Kiara: Drawing the graphs for f' and f'' because no exact numbers were given to help with the graph [the graph given in the task].

Based on the evidence presented in the preceding excerpt, it would seem that making sense of the relationship between the quantities represented by the input and output variables for the function f was straightforward for Kiara in each task. Kiara identified “determining the parts of the domain corresponding to f being negative, positive, or zero” as being challenging in Task 1. That is, determining open intervals where the function f is positive or negative was particularly challenging for Kiara. She went on to identify sketching the graphs of the functions f' and f'' as problematic in Task 2, an experience shared by several other students as well.

When prompted to comment on difficulties they had with determining open intervals where the functions f , f' , and f'' are positive or negative in each task, at least five students remarked on: (1) the challenge of not being given an algebraic formula/equation for the functions f' and f'' in each task, (2) the challenge of not being given a graph for the functions f' and f'' in each task, (3) the challenge of not knowing what it means for a function to be positive or negative in its domain, or (4) they cited the lack of labels on the vertical and horizontal axes of the graphs given in Tasks 1 and 2 as something that limited them from being able to correctly determine the open intervals where the functions f , f' , and f'' in each task are positive or negative. It should be noted that even though the vertical and horizontal axes for the graphs given in Tasks 1 and 2 are not labeled, the information needed to label the aforementioned axes is provided in the opening statement of each task. In fact, one student (Claire) explicitly stated that as a consequence of the aforementioned challenges, she relied on “guessing” in her effort to determine open intervals where the functions f , f' , and f'' in each task are positive or negative.

Discussion and Implications

This section presents a discussion of the key findings from the current study. Specifically, these findings are discussed considering the literature reviewed earlier. First, while several students correctly

interpreted the quantity represented by the function f' in each task, a majority of the study participants not only found interpreting the quantity represented by the function f'' in each task to be problematic, but they also incorrectly interpreted this quantity. In essence, this finding suggests that while interpreting real-world non-kinematics quantities that can be represented using first derivatives is seemingly straightforward for students, the same cannot be said with regard to interpreting similar quantities that can be represented using second derivatives. A few other studies have reported on calculus students' difficulties related to interpreting the second derivative (Carlson et al., 2002; Jones, 2019; Ovodenko & Kouropatov, 2019; Tsamir & Ovodenko, 2013). Furthermore, students' reasoning about units of measure for quantities represented by the functions f , f' , and f'' in each task suggests that a majority of the students confused amount quantities with other amount quantities, or they confused amount quantities with rate quantities. Similar results have been reported by other researchers (Jones, 2017; Mkhathshwa, 2020, 2023; Mkhathshwa & Doerr, 2018; Prince et al., 2012; Rasmussen & Marrongelle, 2006). Taken together, the study participants exhibited limited quantitative reasoning abilities as evidenced by the challenges they reported when tasked with interpreting real-world non-kinematics quantities represented by second derivatives or making sense of units of measure for these quantities. From an instructional point of view, I recommend that calculus instructors be intentional about incorporating classroom activities or assignments (the teaching unit on making sense of inflection points that was proposed by Ovodenko and Kouropatov, 2019) that can help students develop strong quantitative reasoning skills, especially as it relates to interpreting quantities represented by first or second derivatives, in addition to analyzing units of measure for these quantities. This is especially important because calculus instruction in the United States tends to focus more on helping students develop strong calculational/algebraic knowledge (e.g., calculating derivatives) and less on important aspects of quantitative reasoning such as interpreting quantities represented by derivatives or the units of measure for such quantities.

Second, a majority of the students in the present study exhibited limited covariational reasoning skills. Specifically, only three students engaged at the highest level of covariational reasoning when tasked with creating first-derivative graphs in the two tasks. Other research has reported on students' difficulties with engaging in covariational reasoning when solving calculus tasks (Jones, 2017, 2019; Mkhathshwa, 2023). Recent findings by Engelke-Infante (2021) suggests that the creation and use of diagrams during calculus instruction could help students develop strong covariational reasoning skills. Considering that covariational reasoning is an important mode of reasoning that is essential for students hoping to acquire a solid understanding of calculus ideas such as inflection points (Carlson et al., 2002), I urge that calculus instructors and calculus textbook authors should strive to include, in their classes or textbooks as the case may be, assignments that are geared toward supporting students in enhancing their covariational reasoning abilities. Recent studies have reported that opportunities to engage in covariational reasoning that are included in widely used calculus textbooks in the United States are not only minimal but are often limited to the lowest levels of covariational reasoning (Mkhathshwa, 2022, 2025b). Increasing covariational reasoning opportunities in calculus textbooks is paramount as research indicates that the content that instructors teach in the classroom and how they teach it is often determined by course textbooks (Begle, 1973; Reys et al., 2004; Wijaya et al., 2015).

Third, a majority of the students struggled with determining open intervals where the functions f , f' , and f'' are positive or negative in each of the two tasks they worked on. Among other difficulties, students reported that not having access to the algebraic or graphical forms of the aforementioned functions as a source of their difficulty in determining the open intervals where these functions are positive or negative. Previous studies have reported on students' proclivity to working with functions in algebraic form (Habre & Abboud, 2006; Herman, 2007; Ibrahim & Rebello, 2012; Thompson, 1994b). Additionally, other students in the present study cited not knowing what it means for a function to be positive or negative as the main reason they failed to determine open intervals where

the functions f , f' , and f'' are positive or negative. Taking into consideration that the study participants were calculus students and that finding open intervals where a function is positive or negative is content typically covered in courses (e.g., college algebra or precalculus) that are considered prerequisite to the study of calculus, findings of the present study suggest that students' struggles with making sense of open intervals where a function is positive or negative persists through at least the study of calculus. Thus, it might be important for calculus instructors not to assume that this content is well understood by students prior to the study of calculus. In essence, I recommend that calculus instructors include this concept among other concepts (e.g., the concept of function) they typically review with students prior to teaching concepts that are specifically about the study of calculus.

Major Contributions of this Study to the Literature

In summary, and as a major contribution to the literature, findings of the present study suggest that students' difficulties with quantitative reasoning or covariational reasoning are not only limited to kinematics contexts (e.g., Carlson et al., 2002, Mkhathswa, 2020), but that these difficulties also extend to non-kinematics contexts such as the contexts of volume of water in a tank and the spread of a disease, respectively, used in the present study. In particular, findings of this study indicate that interpreting quantities represented by the second derivative (f'') in non-kinematics contexts, an important aspect of quantitative reasoning, is particularly challenging for first-semester calculus students. Furthermore, findings of this study show that engaging in higher levels of covariational reasoning is not only particularly challenging for students when working with kinematics contexts (e.g., Carlson et al., 2002), but also when they are working with non-kinematics contexts.

In terms of moving the field forward, the observed widespread students' difficulties as it relates to making sense of derivatives in motion and non-motion contexts generally points to weaknesses in students' conceptual understanding of the concept of the derivative. Specifically, these weaknesses point to a need to focus on effective instructional approaches (Mkhathswa, 2024) in the teaching of the concept of the derivative. As detailed in Mkhathswa (2024), these effective instructional approaches often involve integrating educational technologies (e.g., Desmos and GeoGebra) and problem-solving strategies in the teaching of the concept of the derivative. From a research standpoint, the students' weaknesses observed in the present study point to a need for calculus education researchers to conduct design-based research (DBR). In particular, the goal of the DBR would be to develop teaching interventions (and assessing the effectiveness, or lack thereof, of these teaching interventions) that could be adopted by calculus instructors to help students overcome the weaknesses observed in the present study.

Considering that there is a dearth of research that has examined students' thinking about open intervals where a function is negative, zero, or positive, findings of this study as it relates to the aforementioned line of research provide valuable insight for both researchers and calculus instructors. Specifically, findings of this study provide a baseline for other researchers who may be interested in further research as it relates to students' thinking about open intervals where a function is negative, zero, or positive. For calculus instructors, findings of this study provide an insight on the nature of students' difficulties as it relates to open intervals where a function is negative, zero, or positive. Consequently, calculus instructors may want to consider helping students develop a solid understanding of these precalculus ideas, prior to covering calculus-specific topics such as those dealing with finding open intervals where a function is increasing or decreasing.

Limitations

I remark on two limitations of the current study. First, the study participants were recruited from the same institution, something that limits the study's generalizability (i.e., the study's external

validity). Taking into consideration that they were recruited from three sections of a Calculus I course taught by two experienced professors, I argue that the severity of this limitation is somewhat alleviated. Second, the sample size is relatively small, another potential threat to the study's external validity. I note that in conducting this exploratory study, the goal was not so much to extend the results to other settings, but rather to gain insight on calculus students' thinking about the concept of function through the lens of quantitative and covariational reasoning. That is, the goal of the study was to determine which aspects of the concept of the derivative are more or less challenging to calculus students. Future research may utilize a larger sample size and study participants from several institutions to improve the present study's external validity.

Acknowledgements

Task 1 and Task 2 were adapted from a Calculus I writing assignment designed by Tom Farmer, Emeritus Professor at Miami University.

The authors received no financial support for the research, authorship, and/or publication of this manuscript

Thembinkosi Peter Mkhathshwa (mkhatstp@miamioh.edu) is an Associate Professor of Mathematics at Miami University-Middleton, Ohio, United States. He teaches mathematics and statistics courses. His research generally spans three areas of study, namely student' thinking about fundamental ideas/concepts in calculus, experts' views on effective instructional approaches in the teaching of fundamental ideas/concepts in calculus, and opportunities to learn about fundamental ideas/concepts provided by widely used calculus textbooks in the United States.

References

- Altindis, N. (2025). Networking theories of quantitative reasoning and mathematical reasoning to explore students' understanding of functions. *The Journal of Mathematical Behavior*, 80, 101276. <https://doi.org/10.1016/j.jmathb.2025.101276>
- Begle, E. G. (1973). Some lessons learned by MSG. *Mathematics Teacher*, 66(3), 207–214.
- Beichner, R. (1994). Testing student interpretation of kinematics graphs. *American Journal of Physics*, 62(8), 750–755.
- Biza, I., González-Martín, A. S., & Pinto, A. (2022). 'Scaffolding' or 'filtering': A review of studies on the diverse roles of calculus courses for students, professionals and teachers. *International Journal of Research in Undergraduate Mathematics Education*, 8(2), 389–418. <https://doi.org/10.1007/s40753-022-00180-1>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101.
- Burns-Childers, A., & Vidakovic, D. (2018). Calculus students' understanding of the vertex of the quadratic function in relation to the concept of derivative. *International Journal of Mathematical Education in Science and Technology*, 49(5), 660–679. <https://doi.org/10.1191/1478088706qp0630a>

- Carlson, M. P. (1998). A cross-sectional investigation of the development of the function concept. *CBMS Issues in Mathematics Education*, 7, 114-162.
- Carlson, M., Jacobs, S., Coe, E., Larsen, S., & Hsu, E. (2002). Applying covariational reasoning while modeling dynamic events: A framework and a study. *Journal for Research in Mathematics Education*, 33(5), 352-378. <https://doi.org/10.2307/4149958>
- Çetin, N. (2009). The ability of students to comprehend the function-derivative relationship with regard to problems from their real life. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 19(3), 232-244. <https://doi.org/10.1080/10511970701686987>
- Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches* (4th ed.). Thousand Oaks, CA: Sage Publications.
- Dominguez, A., Barniol, P., & Zavala, G. (2017). Test of understanding graphs in calculus: Test of students' interpretation of calculus graphs. *EURASIA Journal of Mathematics, Science and Technology Education*, 13(10), 6507-6531. <https://doi.org/10.12973/ejmste/78085>
- Dubinsky, E., & Wilson, R. T. (2013). High school students' understanding of the function concept. *The Journal of Mathematical Behavior*, 32(1), 83-101. <https://doi.org/10.1016/j.jmathb.2012.12.001>
- Engelke-Infante, N. (2021). Helping students think like mathematicians: Modeling-related rates with 2 diagrams. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 31(7), 749-759. <https://doi.org/10.1080/10511970.2020.1712669>
- García-García, J., & Dolores-Flores, C. (2021). Pre-university students' mathematical connections when sketching the graph of derivative and antiderivative functions. *Mathematics Education Research Journal*, 33, 1-22. <https://doi.org/10.1007/s13394-019-00286-x>
- Goldin, G. A. (2000). A scientific perspective on structured, task-based interviews in mathematics education research. In A. E. Kelly & R. A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 517-545). Mahwah, NJ: Laurence Erlbaum Associates.
- Habre, S., & Abboud, M. (2006). Students' conceptual understanding of a function and its derivative in an experimental calculus course. *The Journal of Mathematical Behavior*, 25(1), 57-72. <https://doi.org/10.1016/j.jmathb.2005.11.004>
- Haciomeroglu, E. S., Aspinwall, L., & Presmeg, N. C. (2010). Contrasting cases of calculus students' understanding of derivative graphs. *Mathematical Thinking and Learning*, 12(2), 152-176. <https://doi.org/10.1080/10986060903480300>
- Herman, M. (2007). What students choose to do and have to say about use of multiple representations in college algebra. *Journal of Computers in Mathematics and Science Teaching*, 26(1), 27-54.
- Ibrahim, B., & Rebello, N. S. (2012). Representational task formats and problem solving strategies in kinematics and work. *Physical Review Special Topics-Physics Education Research*, 8(1), 010126(1-19). <https://doi.org/10.1103/PhysRevSTPER.8.010126>

- Jones, S. R. (2019). Students' application of concavity and inflection points to real-world contexts. *International Journal of Science and Mathematics Education*, 17, 523-544. <https://doi.org/10.1007/s10763-017-9876-5>
- Jones, S.R. (2017). An exploratory study on student understandings of derivatives in real-world, non-kinematics contexts. *The Journal of Mathematical Behavior*, 45, 95-110. <https://doi.org/10.1016/j.jmathb.2016.11.002>
- Kertil, M. (2020). Covariational reasoning of prospective mathematics teachers: How do dynamic animations affect?. *Turkish Journal of Computer and Mathematics Education*, 11(2), 312-342. <https://doi.org/10.16949/turkbilmat.652481>
- Kertil, M., & Dede, H. G. (2022). Promoting prospective mathematics teachers' understanding of derivative across different real-life contexts. *International Journal for Mathematics Teaching and Learning*, 23(1), 1-24. <https://doi.org/10.4256/ijmtl.v23i1.361>
- Lobato, J., Hohensee, C., Rhodehamel, B., & Diamond, J. (2012). Using student reasoning to inform the development of conceptual learning goals: The case of quadratic functions. *Mathematical Thinking and Learning*, 14(2), 85-119. <https://doi.org/10.1080/10986065.2012.656362>
- Malambo, P., van Putten, S., Botha, H., & Stols, G. (2018). Dysfunctional functions: The case of Zambian mathematics education students. *Eurasia Journal of Mathematics, Science and Technology Education*, 15(1), em1651. <https://doi.org/10.29333/ejmste/99510>
- Martínez-Planell, R., & Gaisman, M. T. (2012). Students' understanding of the general notion of a function of two variables. *Educational Studies in Mathematics*, 81(3), 365-384. <https://doi.org/10.1007/s10649-012-9408-8>
- Melhuish, K., Lew, K., Hicks, M. D., & Kandasamy, S. S. (2020). Abstract algebra students' evoked concept images for functions and homomorphisms. *The Journal of Mathematical Behavior*, 60, 100806. <https://doi.org/10.1016/j.jmathb.2020.100806>
- Mkhatshwa, T. (2020). Calculus students' quantitative reasoning in the context of solving related rates of change problems. *Mathematical Thinking and Learning*, 22(2), 139-161. <https://doi.org/10.1080/10986065.2019.1658055>
- Mkhatshwa, T. (2022). Quantitative and covariational reasoning opportunities provided by calculus textbooks: The case of the derivative. *International Journal of Mathematical Education in Science and Technology*. <https://doi.org/10.1080/0020739X.2022.2129497>
- Mkhatshwa, T. (2023). A quantitative and covariational reasoning investigation of students' interpretations of partial derivatives in different contexts. *International Journal of Mathematical Education in Science and Technology*, 54(4), 511-533. <https://doi.org/10.1080/0020739X.2021.1958941>
- Mkhatshwa, T. (2024). Best practices for teaching the concept of the derivative: Lessons from experienced calculus instructors. *Eurasia Journal of Mathematics, Science and Technology Education*, 20(4), 1-17. <https://doi.org/10.29333/ejmste/14380>

- Mkhatshwa, T. (2025a). An examination of calculus students' quantitative and covariational reasoning in the context of differentials. *Teaching Mathematics and Its Applications: An International Journal of the IMA*. <https://doi.org/10.1093/teamat/hraf003>
- Mkhatshwa, T. (2025b). A qualitative document analysis of quantitative and covariational reasoning opportunities provided by calculus textbooks: The case of related rates problems. *The Electronic Journal for Research in Science & Mathematics Education*, 29(2), 57-82.
- Mkhatshwa, T., & Doerr, H. (2018). Undergraduate students' quantitative reasoning in economic contexts. *Mathematical Thinking and Learning*, 20(2), 142-161. <https://doi.org/10.1080/10986065.2018.1442642>
- Monk, G. S. (1994). Students' understanding of functions in calculus courses. *Humanistic Mathematics Network Journal*, 1(9), 21-27.
- Monk, S. (1992). Students' understanding of a function given by a physical model. In G. Harel & E. Dubinsky (Eds.), *The concept of function: Aspects of epistemology and pedagogy*, MAA Notes, Vol.25 (pp. 175-193). Washington, DC: Mathematical Association of America.
- Monk, S., & Nemirovsky, R. (1994). The case of Dan: Student construction of a functional situation through visual attributes. *CBMS Issues in Mathematics Education*, 4, 139-168.
- Nagle, C., Tracy, T., Adams, G., & Scutella, D. (2017). The notion of motion: covariational reasoning and the limit concept. *International Journal of Mathematical Education in Science and Technology*, 48(4), 573-586. <https://doi.org/10.1080/0020739X.2016.1262469>
- National Council of Teachers of Mathematics. (2000). *Principles and standards for school mathematics*. Reston, VA: Author.
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, DC: Authors. <http://www.corestandards.org/Math/>
- Nguyen, A. T. T., & Tran, D. (2025). Quantitative reasoning as a lens to examine changes in modelling competencies of secondary preservice teachers. *Mathematics Education Research Journal*, 37(1), 27-54. <https://doi.org/10.1007/s13394-023-00481-x>
- Oehrtman, M. C., Carlson, M. P., & Thompson, P. W. (2008). Foundational reasoning abilities that promote coherence in students' understandings of function. In M. P. Carlson & C. Rasmussen (Eds.), *Making the connection: Research and practice in undergraduate mathematics* (pp. 27-42). Washington, DC: Mathematical Association of America.
- Orhun, N. (2012). Graphical understanding in mathematics education: Derivative functions and students' difficulties. *Procedia-Social and Behavioral Sciences*, 55, 679-684. <https://doi.org/10.1016/j.sbspro.2012.09.551>
- Ovodenko, R., & Kouropatov, A. (2019). The use of digital tools to confront errors during advanced calculus learning: the case of the inflection point. *Mathematics in Computer Science*, 13(1), 217-236. <https://doi.org/10.1007/s11786-018-0365-1>

- Rasmussen, C., & Marrongelle, K. (2006). Pedagogical content tools: Integrating student reasoning and mathematics in instruction. *Journal for Research in Mathematics Education*, 37(5), 388-420. <https://doi.org/10.2307/30034860>
- Reys, B. J., Reys, R. E., & Chavez, O. (2004). Why mathematics textbooks matter. *Educational Leadership*, 61(5), 61-66.
- Smith, J. P. J., & Thompson, P. W. (2007). Quantitative reasoning and the development of algebraic reasoning. In J. J. Kaput, D. W. Carraher & M. L. Blanton (Eds.), *Algebra in the early grades* (pp. 95-132). New York: Erlbaum.
- Spall, S. (1998). Peer debriefing in qualitative research: Emerging operational models. *Qualitative Inquiry*, 4(2), 280-292.
- Swidan, O., Schacht, F., Sabena, C., Fried, M., El-Sana, J., & Arzarello, F. (2019). Engaging Students in Covariational Reasoning within an Augmented Reality Environment. In T. Prodromou (Ed.), *Augmented Reality in Educational Settings* (pp. 147-167). Brill Sense. https://doi.org/10.1163/9789004408845_007
- Thompson, P. W. (1993). Quantitative reasoning, complexity, and additive structures. *Educational Studies in Mathematics*, 25(3), 165-208.
- Thompson, P. W. (1994b, April). *Bridges between mathematics and science education*. Paper presented at the Research Blueprint for Science Education Conference. New Orleans, LA.
- Thompson, P. W. (1994c). Students, functions, and the undergraduate curriculum. *Research in Collegiate Mathematics Education*, 1, 21-44.
- Thompson, P. W. (1994d). The development of the concept of speed and its relationship to concepts of rate. In G. Harel & J. Confrey. Eds., *The development of multiplicative reasoning in the learning of mathematics*, 179-234. Albany: SUNY Press.
- Thompson, P. W. (2011). Quantitative reasoning and mathematical modeling. In L. L. Hatfield, S. Chamberlain & S. Belbase (Eds.), *New perspectives and directions for collaborative research in mathematics education*. WISDOMe Monographs (Vol. 1, pp. 33- 57). Laramie, WY: University of Wyoming.
- Trujillo, M., Atarés, L., Canet, M. J., & Pérez-Pascual, M. A. (2023). Learning difficulties with the concept of function in maths: A literature review. *Education Sciences*, 13(5), 495. <https://doi.org/10.3390/educsci13050495>
- Tsamir, P., & Ovodenko, R. (2013). University students' grasp of inflection points. *Educational Studies in Mathematics*, 83(3), 409-427. <https://doi.org/10.1007/s10649-012-9463-1>
- Ubuz, B. (2007). Interpreting a graph and constructing its derivative graph: Stability and change in students' conceptions. *International Journal of Mathematical Education in Science and Technology*, 38(5), 609-637. <https://doi.org/10.1080/00207390701359313>
- Van den Eynde, S., Van Kampen, P., Van Dooren, W., & De Cock, M. (2019). Translating between graphs and equations: The influence of context, direction of translation, and function

type. *Physical Review Physics Education Research*, 15(2), 020113.
<https://doi.org/10.1103/PhysRevPhysEducRes.15.020113>

Wijaya, A., van den Heuvel-Panhuizen, M., & Doorman, M. (2015). Opportunity-to-learn context-based tasks provided by mathematics textbooks. *Educational Studies in Mathematics*, 89(1), 41–65. <https://doi.org/10.1007/s10649-015-9595-1>

Yerushalmy, M., & Swidan, O. (2012). Signifying the accumulation graph in a dynamic and multi-representation environment. *Educational Studies in Mathematics*, 80(3), 287-306. <https://doi.org/10.1007/s10649-011-9356-8>

Examining Middle School Mathematics Teachers' Burnout and Self-Efficacy During the Pandemic

Sinem Elif Elma 

Middle East Technical University

Ekin Balci 

TED University

Mine Işiksal-Bostan 

Middle East Technical University

ABSTRACT

The purpose of this study was to investigate the differences among middle school mathematics teachers' burnout and perceived self-efficacy levels in terms of gender, weekly lesson hours, school type, years of experience, marital status, and parental status during online education due to the COVID-19 pandemic. To examine the differences, a survey research design was used, and the data was collected from 316 middle school mathematics teachers in Turkey through Maslach's Burnout Inventory and Self-Efficacy Belief in Distance Education Scale. Findings revealed that the characteristics of gender, weekly lesson hours, school type, and marital status had no effect on the teachers' burnout. However, experienced teachers felt higher burnout than novice teachers. Also, compared to the teachers who had one or more children, the teachers with no children had less burnout. Considering the self-efficacy of the teachers, similar to the burnout, experienced teachers had a lower level of self-efficacy. Additionally, the variable of school type had a significant effect on teachers' self-efficacy for the benefit of the private school teachers. Yet, the other characteristics such as gender, weekly lesson hours, marital status, and having children did not affect the teachers' self-efficacy.

Keywords: Burnout, Teacher Self-Efficacy, Online Education, Covid-19

Introduction

The pandemic presented challenges in every aspect of life, particularly in education (OECD, 2021). In March 2020, Turkey's Ministry of National Education (MoNE) informed that face-to-face education at all levels was initially suspended for three weeks. However, due to the course of the pandemic, face-to-face education, similar to that in many countries around the world, could not continue for the entire semester. Consequently, online education was adopted for the rest of the semester through online platforms created by MoNE to conduct the lessons. However, many teachers working at K-12 levels were unfamiliar with and unprepared for conducting their classes online, and many of them had challenges such as access to the internet (Marshall et al., 2020; Simmons et al., 2020), adjusting their materials and lessons to online education (Honigsfeld & Nordmeyer, 2020) and having emotional and psychological problems (Presley, 2021). Additionally, depending on the teachers' familiarity with the online platforms or their experience, teachers had difficulties with online

classroom management (Monroy-Gómez-Franco et al., 2022). Already, “teaching is widely considered physically and psychologically challenging and stressful work” (Yerdelen et al., 2016, p. 148); therefore, the new challenges brought by the pandemic required examining teachers’ self-efficacy and burnout levels in online education during the pandemic. Although there were many studies investigating self-efficacy and burnout levels of teachers in terms of several characteristics, there is still room and a need to investigate teachers’ self-efficacy and burnout while teaching online during the COVID-19 pandemic. Since online teaching offers an alternative for education and educators when there is an unexpected and unfortunate situation, such as a pandemic or an earthquake, it is important to investigate the teachers’ self-efficacy and burnout levels in regard to different characteristics. In other words, by identifying the critical factors related to teachers’ self-efficacy and burnout, we could propose alternative solutions to the challenges brought by online education. Therefore, in light of this information, the current study aimed to investigate middle school mathematics teachers’ burnout and perceived self-efficacy levels in terms of gender, marital status, parental status, years of experience, weekly lesson hours, and school type during the online education due to the COVID-19 pandemic.

Theoretical Framework

This section will provide theoretical backgrounds and related studies on burnout and self-efficacy.

Burnout

“Burnout has most commonly been defined by a symptom pattern” (Savicki & Cooley, 1983, p. 229). Karakaş and Bilen (2014) described burnout as the state of an employee experiencing loss of power and energy due to failure, depreciation, and overloading or depletion of one’s inner sources and energy due to unmet desires. Burnout is a serious issue that can significantly affect individuals’ physical and psychological health, ranging from minor symptoms like headaches to severe outcomes such as depression and even suicide (Hendrickson, 1979). According to Elflein (2020), before COVID-19, global perfect mental health was 45.6%, and the percentage of people who had nonfunctional or extremely challenged mental health was 6.8%. However, after COVID-19, the percentage of global perfect mental health decreased by 13.8%, and the percentage of global extremely challenged mental health increased by 7.6%. Burnout is typically categorized and measured under three dimensions according to the Maslach Burnout Model: emotional exhaustion, depersonalization, and lack of personal accomplishment (Maslach & Schaufeli, 1993). This study employs the Maslach Burnout Inventory (MBI) to measure burnout levels across the dimensions of emotional exhaustion, depersonalization, and lack of personal accomplishment, as detailed in the methodology section.

Emotional Exhaustion

Maslach and Jackson described emotional exhaustion as “feelings of being emotionally overextended and drained by one’s contact with other people” (1984, p. 134). The main reasons for this emotional fatigue could be excessive workload, mobbing at work, occupational stress, problems in social relations, imbalance in the workload and leisure time, economic concerns, and so on (Weber & Jaekel-Reinhard, 2000). The COVID-19 pandemic has exacerbated these issues for teachers, introducing additional stressors such as fear of the virus, role ambiguity, and challenges with remote learning (Pressley, 2021). Studies have shown a significant increase in emotional exhaustion among teachers during this period, with factors like inadequate technical equipment and high levels of extraversion contributing to increased risk (Sokal et al., 2020). Moreover, individuals experiencing emotional exhaustion often report persistent frustration, stress, and anxiety about their daily work

tasks (Sağlam Arı & Çına Bal, 1983). These symptoms affect job performance and spill over into personal life, leading to a pervasive sense of depletion. For teachers during the pandemic, this has manifested as decreased teaching enthusiasm, increased likelihood of considering leaving the profession, and negative impacts on psychological well-being, including higher levels of depression and anxiety (Collie, 2021; Kim & Asbury, 2020).

Depersonalization

Depersonalization is characterized by an intense feeling of indifference to other people's needs and negative attitudes toward them (Maslach & Jackson, 1981). Furthermore, these negative attitudes and behaviors were more permanent and resistant to change even when workplace settings and resources were altered compared to other dimensions of burnout (Savicki & Cooley, 1983). In general, having persistent feelings of emotional exhaustion in the long term leads to depersonalization and inhuman behaviors (Maslach & Goldberg, 1998). The COVID-19 pandemic has exacerbated these issues for teachers. Factors such as fear of COVID-19, role ambiguity, and challenges with remote learning have contributed to increase depersonalization among teachers (Sokal et al., 2020). As a result of depersonalization, individuals may behave in rude, injurious, and arrogant ways towards not only their clients (or students) but also their co-workers, which can further deteriorate workplace relationships and overall morale. This is particularly concerning in the educational context, as higher levels of depersonalization have been associated with negative impacts on psychological well-being and an increased likelihood of teachers considering leaving the profession (Collie, 2021).

Lack of Personal Accomplishment

The last dimension of burnout is reduced personal accomplishment. Individuals set personal and professional goals, and their sense of accomplishment is influenced by their perceived success in achieving them. According to Maslach and Jackson (1981), if a person's perception of personal accomplishment decreases or their expectations remain unmet, they would experience a diminished sense of personal achievement. The COVID-19 pandemic has further complicated this issue for teachers, with some studies indicating a decrease in feelings of personal accomplishment during this period (Weißenfels et al., 2021). Factors such as perceived infectability and role ambiguity have been significant predictors of reduced personal accomplishment (Ptacek et al., 2022). Conversely, a heightened feeling of reduced personal accomplishment and work inadequacy can reinforce a sense of incapability, often referred to as a self-fulfilling prophecy (Özler, 2015). This lack of accomplishment can lead to demotivation and a loss of enthusiasm for the job, further impacting teaching quality and student engagement (Pressley, 2021).

The impact of COVID-19 on teachers' mental well-being and burnout has been examined in various studies. For instance, Pressley (2021) investigated the factors contributing to teacher burnout during COVID-19 to inform stakeholders. At the end of the study, three variables, current teaching anxiety, anxiety communicating with parents, and administrative support, significantly predicted teachers' burnout. Similarly, Sokal et al. (2020) aimed to detect the factors related to teacher burnout and endurance during the first wave of COVID-19 and made empirically backed recommendations to better support teachers during the next wave. Decreasing demands, providing appropriate sources of support, giving importance to administration and parental backing, and promoting physical and emotional self-care practices were suggested as essential strategies to prevent teacher burnout. Like Sokal et al. (2020), who emphasized the importance of reducing demands, providing adequate support, and promoting self-care practices to prevent teacher burnout during the first wave of COVID-19, Amri et al. (2020) investigated these dynamics among 125 primary Moroccan teachers. Their findings revealed that high levels of burnout were closely linked to increased workload, work-

family role conflicts, limited ICT skills, and insufficient social support, ultimately highlighting the urgent need for targeted interventions to protect teachers' mental well-being during and after lockdown. Building upon these findings, it is evident that teacher burnout represents a pervasive and multifaceted issue within the educational sector, exacerbated by unforeseen circumstances such as the COVID-19 pandemic. The rapid shift to online teaching modalities has introduced additional stressors that have significantly impacted teachers' mental health around the world. Numerous studies have investigated the impact of these transitions on teacher well-being, highlighting the multifaceted nature of stressors and their implications for educators' mental health. For instance, in a study conducted in India, Kumawat (2020) examined the prevalence of perceived stress and burnout among teachers engaged in online teaching during the pandemic. Findings revealed that a significant proportion of teachers experienced moderate to high levels of stress and emotional exhaustion, underscoring the need for interventions to mitigate these adverse effects. Despite these challenges, many teachers demonstrated resilience, maintaining high levels of personal accomplishment. Similarly, Shlenskaya et al. (2020) explored burnout among university teachers in Russia transitioning to online learning modalities during the pandemic. Through a survey using the Maslach Burnout Inventory, the study identified factors contributing to teachers' emotional exhaustion, depersonalization, and diminished personal accomplishment. Variations in burnout levels across these dimensions emphasized the diverse experiences of educators during this period. Extending this line of inquiry beyond the primary and secondary education context, the abrupt shift to online teaching prompted by the COVID-19 pandemic has also been examined in broader contexts, including higher education institutions worldwide. Mosleh et al. (2022) examined the impact of this transition on the stress and burnout levels of faculty and course instructors working remotely. Through a cross-sectional design involving 278 participants, results showed moderate stress levels, with those having 7–10 years of online teaching experience reporting higher stress. Elevated stress and lower satisfaction were linked to increased burnout, particularly among married participants with school-age children. These findings emphasize the need for supportive professional development to address faculty burnout in remote teaching. Similarly, Daumiller et al. (2021) explored faculty members' attitudes toward online teaching and their associations with underlying motivations, burnout levels, and student learning outcomes. Results highlighted the importance of faculty goals and attitudes, with learning approach goals positively linked to perceiving the transition as a positive challenge that enhanced teaching quality.

In addition to contextual and psychological factors, recent research has highlighted the importance of different demographic characteristics in shaping teachers' experiences of burnout. Further supporting the literature, Yavuz (2009) claimed that gender significantly affects burnout, with male teachers showing higher levels of burnout than females. On the other hand, Ozamiz-Etxebarria et al. (2021), Sokal et al. (2020), and Pressley (2021) studied burnout and found that female teachers reported higher levels of burnout. Similarly, Sánchez-Pujalte et al. (2021) explored burnout levels among Spanish teachers during COVID-19, and found significant differences associated with gender and years of experience, with higher burnout levels reported by women and less experienced teachers. Similarly, considering the year of experience, Yerdelen et al. (2016) found that teachers with more experience tend to have higher personal accomplishment but lower emotional exhaustion and depersonalization.

Furthermore, marital and parental status has been identified as a critical demographic characteristic shaping teachers' experiences of burnout. For instance, several studies highlighted that single teachers were more susceptible to burnout, pointing to marital status as key predictors of burnout risk (Başol & Altay, 2009; Karanfil, 2019; Zadok Gurman et al, 2021). On the contrary of these findings, Mosleh et al. (2022) suggested that married teachers experienced higher burnout than single ones. Regarding parental status, some studies suggest that teachers with children have lower levels of emotional exhaustion and depersonalization and higher levels of personal accomplishment

(Karabıyık Özipek, 2006; Yerdelen et al., 2016), others claim that teachers without children report low levels of emotional exhaustion (Mosleh et al., 2022).

Moreover, various research showed that the type of school and weekly lesson hours also play a role in burnout, with Karanfil (2019) finding that teachers in public schools experience more emotional exhaustion and fewer personal accomplishments than their private school counterparts. On the other hand, Ozamiz-Etxebarria et al. (2021) found that private school teachers had a higher burnout during COVID-19. Additionally, the number of weekly lesson hours has been shown to correlate with increased burnout, particularly in terms of emotional exhaustion and depersonalization (Yerdelen et al., 2016).

Collectively, these studies underscore the importance of understanding the diverse factors contributing to teacher burnout, including demographic variables, work conditions, and the unprecedented challenges posed by the COVID-19 pandemic. This understanding can inform targeted interventions to support teacher well-being and enhance the overall quality of education in both traditional and online settings.

Teacher Self -Efficacy

Bandura (1977a) defines self-efficacy in his social learning theory as “the belief in one’s capabilities to organize and execute the courses of action required to manage prospective situations.” (p.3). In this view, self-efficacy is influenced by four major sources, which are mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective responses (Bandura, 1997). If people have successful experiences, it helps them believe in the strength of their efficacies in similar tasks. By contrast, an unsuccessful experience leads them to doubt their competencies (Bandura, 1994). It is suggested that self-efficacy develops through investigating other people’s experiences. When people cannot determine their failure or success because of their insufficient experiences, they generally tend to compare themselves with others and make an inference about their possible outcome (Bandura, 1997). People can boost or diminish their self-efficacy by being persuaded by others’ verbal responses. When people are encouraged to think they can achieve, they tend to put in more effort. In terms of physiological and affective responses, an individual’s body responses to the conditions at hand affect self-efficacy beliefs. People use physiological and affective responses such as stress and think these reactions will lead to failure. Further, when people need to perform physically, their states, such as tiredness and aches, lead them to consider their deficiency in their physical skills. Finally, this affects their interpretation of the capabilities they possess (Bandura, 1994, 1997).

According to Bandura (1977b), people with a sense of self-efficacy recover easily when they encounter unsuccessful situations and think about managing them without worrying needlessly. Moreover, they tend to think that failures are related to inadequate abilities and knowledge, which can be enhanced. This kind of belief about failures helps them decrease their anxiety levels (Bandura, 1994). Besides, people with low self-efficacy refrain from taking responsibility and trying due to thinking of failure since they are convinced that this failure stemmed from their insufficiencies, not because of insufficient effort or knowledge about people with high self-efficacy. These people cannot evaluate their competence of knowledge and abilities; hence, they have trouble through their learning process. Also, they give up easily or have a high level of anxiety when they have difficulties (Bandura, 1994, 1997).

Since having a high self-efficacy level is critical, investigating the teachers’ self-efficacy beliefs gain importance. Following Bandura’s definition of self-efficacy, teacher self-efficacy can be perceived as what the teachers think about their capabilities “to plan, organize, and carry out activities required to attain given educational goals” (Skaalvik & Skaalvik, 2010, p. 1059). Tschannen-Moran and Hoy (2001) claimed that teacher self-efficacy is a criterion that has the potential to have either a

positive or negative effect on education. Research reveals that teachers who believe in their abilities work more and are more adept at organizing and planning (Pajares, 1992). Furthermore, teacher self-efficacy plays an essential role in students' success and motivation, and there is a positive relationship between these two concepts (Caprara et al., 2006; Lazarides et al., 2018).

Similar to burnout, some demographics have a relationship with teacher self-efficacy. Considering the gender variable, whereas many researchers expressed that gender did not affect teachers' self-efficacy beliefs (Robinia & Anderson, 2010), some researchers suggested that it had a significant effect. While some studies showed that female teachers had higher levels of self-efficacy beliefs (Atta et al., 2012), some advocated that male teachers' self-efficacy levels were higher than females' (Klassen & Chiu, 2010).

Years of experience have consistently been identified as a significant factor influencing teachers' self-efficacy, with more experienced teachers generally demonstrating higher self-efficacy levels (Atta et al., 2012; Ávalos & Valenzuela, 2016; Hoy & Wollfok, 1993; Reyhing & Perren, 2021). Similarly, research examining the relationship between self-efficacy and weekly lesson hours has revealed a positive relationship and suggested that teachers with a greater number of weekly lesson hours tend to have higher self-efficacy levels (Reyhing & Perren, 2021). Furthermore, school type has also been shown to have a critical role, as a study conducted by Semerci and Uyanik Balat (2018) revealed. According to the study, private school teachers often report higher self-efficacy levels compared to public school teachers.

Although limited, some research has explored the relationship between marital status and teachers' self-efficacy (Batoool et al., 2020). However, studies specifically investigating this demographic characteristic within the context of COVID-19 warrant further attention. Transforming to online education and teaching during the pandemic brought new challenges for teachers. The studies indicated that in online teaching, teachers faced difficulties in communicating with the students due to the separation between the parties (Putri et al., 2020) and spending more time to be familiar with online teaching and motivating students (Scull et al., 2020). Based on these findings, it is evident that the abrupt transition to remote teaching has the potential to increase teachers' burnout levels while at the same time undermining their self-efficacy, which may ultimately compromise the overall quality of education.

Furthermore, the literature highlights the necessity of examining how demographic characteristics influence teachers' experiences of burnout and self-efficacy. Investigating these relationships becomes critical as it can inform the development of support strategies and interventions for teachers, and enable stakeholders to address specific needs more effectively.

In particular, marital status has been identified as a demographic characteristic which affects both burnout and self-efficacy, and this relationship may be further moderated by parental status. Exploring the independent and interactive effects of these variables is therefore essential to gain an understanding of teachers' psychological resilience and professional functioning during crisis situations such as the COVID-19 pandemic. By addressing these gaps, this research aims to contribute valuable insights for policymakers and educational leaders seeking to promote teacher well-being and sustain educational quality under challenging circumstances.

Thus, in this research, we aimed to investigate the following questions:

1. Do gender, weekly lesson hours, years of experience, and the school type significantly affect burnout and perceived self-efficacy of middle school mathematics teachers during online education due to the COVID-19 pandemic?
2. Do marital status and parental status, considered independently and interactively, significantly affect burnout and perceived self-efficacy of middle school mathematics teachers during online education due to the COVID-19 pandemic?

Methodology

In the study, the quantitative research was implemented using a survey research design since survey research is useful “to collect large amounts of survey data from a representative sample of individuals sampled from the targeted population” (Kalaian, 2008, p. 728). To represent the targeted population, 316 middle school mathematics teachers from all regions of Türkiye participated in the study. In Türkiye, middle school mathematics teachers are responsible for teaching middle school students from 5th grade to 8th grade. In selecting the sample through convenience sampling, online forums and groups regarding mathematics education were identified. Within forums and groups, middle school mathematics teachers were invited to participate voluntarily in the online survey. The decision to employ convenience sampling was pragmatic, aiming to streamline participant recruitment and facilitate data collection within a limited timeframe and resource constraints during the COVID-19 pandemic. Convenience sampling allowed for practical access to middle school mathematics teachers who were readily available and willing to participate. Ethical considerations were rigorously upheld throughout the study, with informed consent obtained from all participating teachers and measures implemented to protect their privacy and confidentiality (Franken et al., 2012). Some demographic information regarding the sample is given below.

Sample

Looking at the demographic information, the participants of the study were 316 (223 females and 93 males) middle school mathematics teachers. The years of experience in teaching ranged from 7 months to 33 years ($M = 11.41$ & $SD = 7.60$), and this variable was separated into three groups. The group of 0 to 10 years for years of experience had the highest percentage (52.8%), the percentage of the group of 10 to 20 years was 32.3%, and the lowest percentage belonged to the group of 20+ years (14.6%). The variable of weekly lesson hours ranged from 11 to 40, including after school lesson hours ($M = 25.30$ & $SD = 6.41$), into four groups. For this variable, the information was that 54.1% of the sample belonged to the group of 16-25 hours, 40.8% of them were 26-40, and the lowest percent (1.9%) was the group of 41 hours and above. The average number of students in a class ranged from 5 to 75 in four groups. For the variable of the number of students in a class, while the highest percentage belonged to 16 to 30 (58.5%), the lowest percentage belonged to 46 and above (1.6%). The percentages of 31 to 45 and 15 and below groups were 20.9% and 17.4%. Moreover, 64.9% of the teachers were married, while 35.1% were single. 78.8% of the teachers had bachelor's degrees, 20.6% had master's degrees, 0.3% had doctoral degrees, and 0.3% were high school graduates. 81.3% of the teachers worked in public schools, and 18.7% worked in private schools. Also, 26.6% of the teachers had one child, 27.8% had two or more children, whereas 45.6% had no child. Lastly, while 70.6% of the teachers teach 8th graders, 57.3% teach 5th graders. Additionally, the percentage of the teachers in terms of grade level exceeded 100 because the teachers taught more than one grade level at the same time.

In addition to the sample's demographic characteristics, the information on the teachers' educational infrastructure is indicated below. 34.5% of the teachers conducted their lessons at home, while 63.3% taught at school and home. Additionally, 53.5% of the teachers did not receive any in-service training, course, or seminar related to online teaching, and 93% had a personal device belonging to themselves. As for the sufficiency of the internet connectivity, 44% of the teachers stated that their connectivity was poor, and what followed this statement was acceptable connectivity with 22.8%; however, poor connectivity and very poor connectivity were 44% and 25%, respectively. It was understood that the good and very good connectivity was very low, at 5.7% and 2.5%, respectively. For the type of electronic devices used during classes, 88.6% of the teachers said they used a computer, 40.8% said a smartphone and 29.7% said a tablet. 90.5% of the teachers used

written documents, 67.4% used EBA (an educational network developed by the Ministry of National Education), and 31% used dynamic geometry applications.

When the education infrastructure information of the sample was checked, most of the teachers used Zoom as a virtual classroom application, with 94.6% and 60.8% using EBA for this goal. 53.8% of the teachers preferred mixed type of lesson conduction, and 44.6% preferred synchronous classes. 17.4% of the teachers stated that 20% and below of their students could access online education. However, only 30.4% of the teachers indicated that 81% and above of their students could access online education. As the teachers stated, the other students' dispersion was 17.4% of them, 20% and below, 15.8% of them, 21% to 40%, 16.8% of them, 41% to 60% connectivity for their students' access to online courses. 21.5% of the teachers did not conduct any assessment during online education, and only 17.7% assessed their students more than once a week. Among the teachers who did assessments during online education, 84.7% chose to give homework to assess students, 54% used online quizzes for assessment, and 23.4% assigned students a project.

Data Collection Process and Procedure

Data was collected through an online survey using the link on the form sent out to teachers. The online survey consisted of four sections. The first section included nine demographic questions regarding gender, marital status, number of children, experience in years, weekly lesson hours, type of school they worked in, last degree they received, the average number of students in their classes, and the class levels they worked with.

The second section is dedicated to the education infrastructure and includes eleven questions. Related to how teachers conduct online education and what they used to do. The questions sought information about the technological tools, materials, applications, medium used for teaching, classroom management, access rate to lessons of the students, and teachers' assessment methods. There were eleven questions, ten of which were multiple choice and one of which was a 5-point Likert-type question. The third section consisted of the Self-Efficacy Belief in Distance Education Scale developed by Yıldız (2015). The scale consisted of 5–5-point Likert type ten items ranging from “1 = I am not sufficient at all” to “5 = I am very sufficient”. The scale was divided into three subscales: learning management, technology management, and virtual classroom management. The learning management dimension refers to the ability to prepare the environment, course materials, teaching activities, measurement, and assessment activities for effective learning. An example sample item for this dimension was “I can prepare effective learning environments.”. The second dimension, which was technology management, included the effective use of school technologies and the ability to solve problems related to the use of technology. An example sample item for this dimension was “I can integrate the necessary technologies for effective learning into my lessons.”. The last dimension was virtual classroom management, which refers to the ability to use virtual classroom applications and manage the classroom successfully while using them. An example sample item for this dimension was “I can perform classroom management in the online classes.”. The reliability values of the scale were calculated for this study, and the values for the study ranged from 0.843 to 0.922. The first dimension, learning management, had four items in the test, and the Cronbach Alpha reliability coefficient was 0.852. The second dimension, technology management, had four items in the test, and the Cronbach Alpha reliability coefficient was 0.900. Then, the last dimension, virtual classroom management, had two items in the test, and the Cronbach Alpha reliability coefficient was 0.843.

The last section is dedicated to Maslach's Burnout Inventory. This scale was developed by Maslach and Jackson (1986) and included twenty-two 5-point Likert-type questions. The items ranged from “1 = Never” to “5 = Always”. The MBI scale was divided into three subscales: emotional exhaustion, depersonalization, and personal accomplishment. The emotional exhaustion

subscale consisted of nine questions related to an individual's emotional exhaustion because of the job, and the Cronbach Alpha reliability coefficient was 0.899. An example sample item for this dimension was "I feel emotionally drained from my work.". The second subscale was depersonalization and included five questions with a Cronbach Alpha reliability coefficient of 0.783. An example sample item for this dimension was "I worry that this job is hardening me emotionally". Depersonalization refers to a teacher's emotionless behaviors towards students. The last subscale was a personal accomplishment with eight questions and referred to a teacher's feelings of professional inadequacy. An example sample item for this dimension was "In my work, I deal with emotional problems very calmly.". The Cronbach Alpha reliability coefficient of the personal accomplishment was 0.723. The reliability value of the whole scale was calculated for this study as 0.903.

Data Analysis

After data were collected from 316 middle school mathematics teachers, quantitative analysis using IBM SPSS Statistics 26.0, was conducted to examine the differences among the dependent variables, teachers' perceived self-efficacy and burnout, regarding independent variables: gender, marital status, parental status, years of experience, weekly working hours, and school type. Multivariate analysis of variance (MANOVA) allows for the examination of significant differences in the means of multiple dependent variables across several groups, while taking into account the relationships among the dependent variables (Tabachnick & Fidell, 2019). Prior to hypothesis testing, preliminary analyses were conducted to ensure that the dataset met the required assumptions for MANOVA.

For all analyses, assumption checks included the evaluation of sample size adequacy, normality, linearity, homogeneity of variance and covariance matrices, multicollinearity, and the presence of univariate or multivariate outliers (Pallant, 2016). Skewness and kurtosis values for key variables were within acceptable ranges (between -2 and +2). Visual inspections of histograms, Q-Q plots, and boxplots also supported the normality assumption. The presence of multivariate outliers was assessed using Mahalanobis distance, and no extreme values requiring exclusion were detected. Linearity was confirmed through scatterplot matrices, and multicollinearity was ruled out based on correlation coefficients below the critical threshold ($r < .80$). The assumptions of homogeneity of variance and covariance were assessed using Levene's test and Box's M test, respectively, and were found to be satisfied across the groups.

To address the first and second research questions, descriptive analysis and inferential statistics were conducted. As inferential analysis, both one-way and two-way MANOVAs were conducted. One-way MANOVA was used to examine whether perceived self-efficacy and burnout levels differed significantly by gender, years of teaching experience, weekly working hours, and school type. A two-way MANOVA was conducted to explore the effects of marital status and parental status on the same dependent variables and on each other. Post-hoc tests were performed when significant multivariate effects were found to identify specific group differences.

All statistical tests were evaluated at a significance level of $\alpha = .05$, and effect sizes (partial eta squared for MANOVA) were reported to assess the practical significance of the findings.

Results

This research investigated the differences among the middle school teachers' perceived self-efficacy and burnout regarding gender, marital status, parental status, years of experience, weekly working hours, and school type. One-way MANOVAs were conducted for gender, years of experience, weekly working hours, and school type, while a two-way MANOVA was conducted for marital and parental status. The descriptive statistics and results are presented according to the

research questions.

The effect of gender, weekly lesson hours, years of experience, and school type

The first research question deals with the effect of gender, weekly lesson hours, school type, and years of experience on burnout and perceived self-efficacy. Participants' descriptive statistics of self-efficacy and burnout related to gender, years of experience, weekly working hours, and school type were shown in the table below. As mentioned before, the range of both the MBI Scale and the Self-Efficacy Belief in Distance Education Scale is 1 to 5. Thus, for each variable, scores could range from 1 as the lowest self-efficacy and the lowest burnout to 5 as the highest self-efficacy and the highest burnout. In other words, when one of the scores increases, the respective score of a dependent variable also increases.

Table 1

Descriptive Statistics of Burnout and Self-Efficacy Related to Gender, Weekly Working Hours, School Type, and Years of Experience

Variables	Groups	<i>n</i>	Burnout		Self-Efficacy	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Gender	Male	93	3.80	.68	4.14	.65
	Female	223	3.69	.60	4.07	.60
Weekly Lesson Hours	0-15	10	4.07	.45	3.88	.54
	16-25	171	3.76	.60	4.07	.62
	26-40	129	3.64	.66	4.10	.61
	41+	6	3.92	.44	4.65	.37
School Type	Public	257	3.72	.62	4.03	.62
	Private	59	3.75	.64	4.36	.54
Years of Experience	0-10	167	3.61	.60	4.15	.58
	11-20	102	3.77	.41	4.11	.66
	21+	46	4.04	.55	3.84	.58

Table 1 shows there were 223 female teachers and 93 male teachers. These teachers were categorized into four groups based on weekly lesson hours, into two groups based on the school type where they worked, and into three groups based on their years of experience. For burnout scores, the table shows male teachers ($M = 3.80$, $SD = .68$) felt more burnout than female teachers ($M = 3.69$, $SD = .60$). For the variable weekly lesson hours of teachers, teachers who have less than 16 hours ($M = 4.07$, $SD = .45$) felt more burnout than teachers who have 16 – 25 hours ($M = 3.76$, $SD = .60$), those who have 26 – 40 hours ($M = 3.64$, $SD = .66$) and those who have more than 40 hours ($M = 3.92$, $SD = .44$). This revealed teachers who have less than 16 hours and more than 40 hours feel more burnout than the other teachers in between. Also, the table shows the teachers who worked in private schools ($M = 3.75$, $SD = .64$) felt more burnout than public school teachers ($M = 3.72$, $SD = .62$). Lastly, teachers with more than 20 years of experience ($M = 4.04$, $SD = .55$) felt more burnout than teachers who have 11 – 20 years of experience ($M = 3.77$, $SD = .41$) and those who have less than 10 years of experience ($M = 3.61$, $SD = .60$). This indicated experienced teachers felt more burnout than novice teachers in online education.

Additionally, as it could be understood from the table, male teachers ($M = 4.14$, $SD = .65$) showed higher self-efficacy levels than female teachers ($M = 4.07$, $SD = .60$). For the weekly lesson

hours, teach more than 40 lesson hours ($M = 4.65$, $SD = .37$) had higher levels of self-efficacy than the teachers with less than 15 hours ($M = 3.88$, $SD = .54$), those with 16 – 25 hours ($M = 4.07$, $SD = .62$) and those with 26- 40 hours ($M = 4.10$, $SD = .61$). It could be concluded that as the number of weekly lesson hours increased, teachers' self – efficacy increased. Concerning the school type, the table implied that teachers who worked in private schools ($M = 4.36$, $SD = .54$) had more self-efficacy levels than the teachers in public schools ($M = 4.03$, $SD = .62$). Furthermore, concerning years of experience, self – efficacy mean scores of teachers with less than 10 years of experience ($M = 4.15$, $SD = .58$) showed higher self – efficacy levels than teachers with 11 – 20 years of experience ($M = 4.11$, $SD = .66$) and those with more than 20 years of experience ($M = 3.84$, $SD = .58$). This indicated experienced teachers had lower self – efficacy levels than the novice teachers in online education.

To test whether the given mean differences in burnout and self-efficacy scores were statistically significant in terms of gender, weekly lesson hours, and years of experience, one-way MANOVAs were conducted. Before MANOVAs, the assumptions such as independence observation, sample size, normality, linearity, univariate and multivariate outliers, multicollinearity and singularity, homogeneity of variance-covariance matrices, and homogeneity of variance were checked (Pallant, 2016). Firstly, for independent observation, since the instruments were applied online, there was no interaction between observations. Secondly, Tabachnick and Fidell (2007) gave a formula ($N > 50+8m$) to check the sample size assumption by considering the number of independent variables (IVs) that would be used. “m” was the number of IVs and N was the number of participants. For this study, m was 5; therefore, $50+8*5 = 90$. The observation number was 316 (N), and due to the $316 > 90$, the sample size assumption was met. Thirdly, regarding the normality assumption, both univariate and multivariate normality should be checked. Furthermore, according to Pallant (2016), in descriptive statistics, when the sample size for each cell of every variable included at least 30 participants, the violations of normality and equality of variance assumptions did not create so much problem. Thus, except for two subgroups of the variable, which was weekly lesson hours, there were no potential violations, but the assumptions were still checked in detail. For univariate normality, all skewness and kurtosis values of self-efficacy, and burnout regarding three independent variables, gender, weekly lesson hours, and years of experience were between -1.17 and 1.32, and they fitted the desired range of -2 and +2 (Pallant, 2007). In addition to checking skewness and kurtosis values, the histograms, normal q – q plots, detrended normal q – q plots, and boxplots were checked, and they met the desired normality. Considering the sample size is less than 50, normality for subgroups of 0-15 hours and 41+ hours would be checked using Shapiro-Wilk (Shapiro & Wilk, 1965). The significance values for the test of normality for burnout of subgroups of 0-15 hours ($p = .61 > .05$) and 41+ hours ($p = .28 > .05$) and self-efficacy of subgroups of 0-15 hours ($p = .36 > .05$) and 41+ hours ($p = .95 > .05$) were met. Thus, no violation was found for univariate normality. Furthermore, to check multivariate normality, the Mahalanobis distance was used and calculated as 14.21 for just one case. In contrast, others were smaller than the critical value for two dependent variables, which was 13.82. (Pallant, 2007). According to Pallant (2007), there was no need to take action when there was a small number of outliers and a reasonable Mahalanobis value. Thus, the multivariate normality assumption was also met because the data set had one outlier, and the Mahalanobis distance value was not extreme. While the normality assumption was checked, the fourth assumption, outliers, were also checked with boxplots and Mahalanobis values, and there was no violation, either.

The fourth assumption, linearity, seemed to be met by considering the linear relationship of each pair of dependent variables regarding independent variables on the scatterplot matrices. The other assumption was multicollinearity and singularity. Since the Pearson correlation was significant and the correlation coefficient ($r = .23$) did not exceed .8 or .9 (Pallant, 2007), multicollinearity was not violated. The last assumptions to be checked were homogeneity of variance-covariance matrices

and homogeneity of variance. As the Box's M Test of Equality of Covariance Matrices showed, gender ($p = .18 > .05$), weekly lesson hours ($p = .68 > .05$), and years of experience ($p = .45 > .05$) satisfied the homogeneity of variance-covariance matrices assumption. Moreover, according to the results in Levene's tests of equality of variance table, the homogeneity of variance assumption of burnout and self-efficacy in terms of gender ($p_{burnout} = .52 > .05$) ($p_{self-efficacy} = .64 > .05$), weekly lesson hours ($p_{burnout} = .30 > .05$) ($p_{self-efficacy} = .28 > .05$), and years of experience ($p_{burnout} = .15 > .05$) ($p_{self-efficacy} = .41 > .05$) was also met. Thus, all assumptions were met for one-way MANOVAs.

According to the multivariate test for gender, regarding the first research question, gender ($p = .28 > .05$) did not significantly affect burnout of middle school mathematics teachers and self-efficacy in online education. In other words, there was no difference between male and female combined burnout and self-efficacy scores. Moreover, there was an insignificant difference between subsamples of the variable weekly lesson hours ($p = .06 > .05$). On the other hand, school type ($p = .00 < .05$) and years of experience ($p = .00 < .05$) affected combined burnout and self-efficacy scores significantly.

Furthermore, while one of the dependent variables, burnout, was not affected by school type [$F(1, 315) = .11, p > .025$; Wilks' Lambda = .95; multivariate effect size = .00], self-efficacy was affected [$F(1, 315) = 14.98, p < .025$; Wilks' Lambda = .95; multivariate effect size = .05]. Due to the significant effect of school type on self-efficacy, the post-hoc test was conducted. The result of the post-hoc analysis showed that the teachers who worked in private schools had a higher self-efficacy rating than those who worked in public schools.

Besides, the dependent variable burnout [$F(2, 313) = 9.75, p < .025$; Wilks' Lambda = .89; partial eta squared = .06; multivariate effect size = .06] and self-efficacy [$F(2, 313) = 4.75, p < .025$; Wilks' Lambda = .89; partial eta squared = .03; multivariate effect size = 0.06] were significantly affected by years of experience. Further post hoc analysis revealed that teachers who had less than 10 years of experience and teachers who had between 10 and 20 years of experience felt less burnout than those who had more than 20 years. Considering the self-efficacy levels, teachers who had less than 10 years of experience and teachers who had between 10 and 20 years of experience had more self-efficacy levels in online education than those who had more than 20 years. Since the partial eta squared values of the test of years of experience for burnout .06 ($\eta^2 = .06$), and for self-efficacy .03 ($\eta^2 < .06$), the effect sizes of the study were small and medium, respectively (Pierce et al., 2004).

The effect of marital status and parental status

The second research question deals with the effect of marital status and parental status on burnout and perceived self-efficacy.

Regarding the second research question, a two-way MANOVA was conducted, and the participants' descriptive statistics of burnout and self-efficacy related to marital status and parental status are shown in the table below.

Table 2*Descriptive Statistics of Burnout and Self-Efficacy Related to Marital Status and Parental Status*

Variables	Groups	<i>n</i>	Burnout		Self-Efficacy	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Marital Status	Married	205	3.80	.63	4.09	.64
	Single	111	3.59	.59	4.09	.57
Parental Status	No Child	144	3.80	.68	4.14	.65
	One or More Children	172	3.70	.60	4.06	.60

For the variable of marital status, there were 205 married and 111 single teachers. Also, for the variable of parental status, there were 144 teachers without a child and 172 teachers with one or more children. With respect to burnout, single teachers ($M = 3.59$, $SD = .59$) felt less burnout than married teachers ($M = 3.80$, $SD = .63$). Additionally, teachers who had one or more children ($M = 3.70$, $SD = .60$) felt less burnout than ones who did not have any ($M = 3.80$, $SD = .68$). In terms of the teachers' self-efficacy, there was no difference in self-efficacy between teachers who were married ($M = 4.09$, $SD = .64$) and who were single ($M = 4.09$, $SD = .57$). For parental status, teachers who did not have children ($M = 4.14$, $SD = .65$) had higher self-efficacy than ones who had ($M = 4.06$, $SD = .60$).

Table 3*Descriptive Statistics of Burnout and Self-Efficacy Related to Interaction of Marital Status and Parental Status*

Variables	Groups	<i>n</i>	Burnout		Self-Efficacy	
			<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Marital Status – Parental Status	Married with No Child	42	3.42	.64	4.29	.49
	Married with One or More Children	163	3.89	.59	4.04	.67
	Single with No Child	102	3.57	.60	4.11	.55
	Single with One or More Children	9	3.77	.52	3.93	.76

Table 3 shows that there were 42 married teachers who had no child, 163 married teachers who had one or more children, 102 single teachers who had no child, and 9 single teachers who had one or more children. Regarding burnout, teachers with the highest burnout scores were the ones who were married and had one or more children ($M = 3.89$, $SD = .59$), whereas teachers with the lowest burnout scores were the ones who were married and with no child ($M = 3.42$, $SD = .64$). In terms of the teachers' self-efficacy, the highest self-efficacy level for teachers was the ones who were married and had no child ($M = 4.29$, $SD = .49$), while the lowest self-efficacy level for teachers was the ones who were single and had one or more children ($M = 3.93$, $SD = .76$). Before conducting one – way MANOVA, the assumptions were checked.

According to the multivariate test, there was no statistical difference between subsamples of the interaction of marital status ($p = .44 > .05$) and parental status ($p = .39 > .05$) on the combined dependent variables. Although there was an insignificant difference between subsamples of marital

status ($p = .08 < .05$) concerning dependent variables, there was a significant difference between subsamples of parental status ($p = .01 < .05$) concerning dependent variables. Since the significance is high, between-subject effects are analyzed to identify the difference among the dependent variables.

Table 4

Tests of Between-Subject Effects for Parental Status

Variables	Burnout				Self-Efficacy			
	<i>df</i>	<i>F</i>	<i>p</i>	Partial Eta Squared	<i>df</i>	<i>F</i>	<i>p</i>	Partial Eta Squared
Parental Status	1	8.16	.005	.03	1	3.16	.08	.01

Table 4 shows that parental status did not significantly affect the self-efficacy variable; however, the burnout variable was significantly affected by parental status [$F(1, 314) = 8.16$, $p < .025$; Wilks' Lambda = .95; partial eta squared = .03; multivariate effect size = 0.05]. Post-hoc analysis showed that teachers with no children felt more burnout than those with one or more children. Besides, the self-efficacy variable [$F(1, 314) = 3.16$, $p > .025$; Wilks' Lambda = .95; partial eta squared = .01; multivariate effect size = 0.05]. Since the partial eta squared of the study was .03 ($\eta^2 < .06$), the effect size of the study was small (Pierce et al., 2004).

Discussion

The purpose of the study was to investigate the effects of gender, weekly lesson hours, years of experience, school type in which the participants worked, marital status, and parental status on burnout and self-efficacy in terms of online education during the pandemic. Considering the first research question, even though the study results pointed out a mean difference between female and male teachers in which female teachers experienced burnout less than male teachers, this difference was not significant. This result was inconsistent with some of the studies in the literature (Ozamiz-Etxebarria et al., 2021; Pressley, 2021; Sokal et al., 2020). The lack of a significant gender difference in the present study might be explained by the extraordinary stressors of COVID-19 pandemic, which might overshadow the gender-based differences regarding burnout. Similarly, according to the current findings, considering the gender, self-efficacy scores of female and male teachers were not found to be significantly different from each other. This result was also consistent with the literature, which stated no difference between male and female teachers regarding their self-efficacy beliefs (Corry & Stella, 2018; Robinia & Anderson, 2010). Although Freeman and He (2010) stated that men had a higher self-efficacy score in the technological context than women, this situation was not valid in this research. In this study, the demographic about taking any seminar or course about online education, was answered positively mostly by women. Thus, this knowledge about the participants may have equalized or decreased the difference between genders and caused insignificant results.

Additionally, when teachers' burnout related to weekly lesson hours was investigated, there was no significant difference between teachers; however, the analysis suggested a mean difference, which indicated that teachers who had less than 16 hours and more than 40 hours felt more burnout than the teachers in between. However, the results have some inconsistencies with the study conducted by Yerdelen and colleagues (2016), which states that burnout levels increase when the weekly lesson hours increase. This result could be attributed to the pandemic when teachers conducted

their lessons mostly at home, and it may not make any difference for them to have fewer or more lesson hours since they were already at home. Regarding self-efficacy, in descriptive statistics, similar to the literature, it could be easily seen that the self-efficacy of teachers increased as weekly lesson hours increased. In contrast to the findings reported in the literature, where the variable of weekly working hours had a significant positive effect on self-efficacy (Reyhing & Perren, 2021), the results were found to be insignificant in this study. Future studies could include a larger sample size to examine the potential effect more deeply.

The findings of the study also demonstrated that there is a significant difference between more experienced and less experienced teachers and demonstrated that teachers with higher experience had higher burnout levels. This finding was inconsistent with the literature, stating that high experience led to low burnout (Yerdelen et al., 2016). Maslach and Jackson (1981) stated that burnout occurred more at the beginning of the profession. Although Maslach and Jackson's idea (1981) and the study demonstrated that novice teachers experienced burnout as psychological self-defense as a consequence of unmet expectations in the workplace (Edelwich & Brodsky, 1980), the reason why experienced teachers felt more burnout in this study could be related to online education and incompetence in technology. In addition to those factors, the pandemic might also have triggered burnout because of the limited connection of students through online education. In terms of self-efficacy, the current study also showed that there is a significant difference in terms of years of experience. According to the post-hoc analysis, teachers who have less than 10 years of experience and teachers who have between 10 and 20 years of experience had higher levels of self-efficacy in online education than the teachers who have more than 20 years of experience, unlike the studies which found that the more years of experience led to the higher self-efficacy in teachers (Atta et al., 2012; Ávalos & Valenzuela, 2016; Reyhing & Perren, 2021). The reason for the difference between the results of the literature and this study could be that the studies in the literature worked on self-efficacy in a general term; however, in this study, the aim was to investigate the effect of the independent variables on self-efficacy during online education in the COVID-19 pandemic. Thus, the necessary competence of online education, which is the ability to use technological tools such as computers and programs such as Zoom, EBA, and Web 2.0, might have affected why the study showed different results from the literature.

According to the Turkish Council of Higher Education (YÖK, 1998), the Instructional Technology and Material Development course was added to the teacher education programs to enhance teachers' competence in technology. This decision meant that teachers who graduated before 2002 did not have the privilege to take this course because it was not in the curriculum. Thus, it could cause novice teachers to have more self-efficacy than experienced ones in online education. Furthermore, the descriptive analysis supported the idea that the teachers who had taken seminars about online education before the survey had higher self-efficacy beliefs than the others.

The variable of school types in which the teachers worked was found to be insignificant for burnout. Unlike the findings of the study in terms of school type, Karanfil (2019) stated that the teachers in private schools had a higher level of burnout. The fact that school type doesn't seem to affect teacher burnout might be due to how the usual differences between schools disappeared during COVID-19. While earlier studies, such as Karanfil (2019), reported higher burnout among private school teachers due to factors like increased workload and parental pressure, these unique stressors were likely overshadowed by the universal challenges due to the pandemic. Online teaching, emotional strain, and systemic uncertainty could have affected teachers across all school types and it may neutralize previously observed differences. Therefore, it may not lead to any significant change for teachers. In terms of self-efficacy, the variable of school types in which the teachers worked was found to be significant regarding self-efficacy for the benefit of private school teachers (Semerci & Uyanık Balat, 2018). The reason why these teachers had higher levels of self-efficacy than others might depend on the provided opportunities, such as purchased programs and obligatory in-service

education by private schools. Moreover, as Karanfil (2019) stated, this might have occurred because of long working hours, a competitive environment, high expectations from teachers, motivated students, and cooperation with parents during a pandemic.

Considering the second question, the finding of the study displayed that the teachers who did not have any children significantly experienced more burnout than the teachers who had one or more children. This result was consistent with the literature, which said that having children made teachers feel less burnout (Karabıyık Özipek, 2006; Yerdelen et al., 2016). As opposed to the current study, the existing literature revealed that having children provided parents with emotional support, reducing the teachers' burnout level. As regards the pandemic, having children brought emotional support for the parents in a difficult time like the pandemic. The effect of parental status also showed that they were insignificant regarding the effect on self-efficacy. These findings were also supported by the studies in the literature (Bonett, 1994).

As for the marital status of the teachers, the results did not indicate any significant difference between married and single teachers. This result was inconsistent with the literature (Başol & Altay, 2009; Karanfil, 2019; Zadok Gurman et al, 2021) which stated that single teachers have higher levels of burnout than married one. This result may be attributed to issues related to working remotely. Also, when the pandemic suddenly hit the world and people started to work at home, partners sharing resources, such as the Internet, workspace, or computers, may have caused unexpected friction. This friction between married couples of teachers could have decreased the difference between single and married teachers. The effect of marital status also showed that they were insignificant regarding the effect on self-efficacy. These findings were also supported by the study of Batool and colleagues (2020). One possible reason for this finding is that, during the pandemic, all teachers, regardless of marital status, had to abruptly adapt to new teaching methods, such as online platforms and digital tools. This might have created a common learning curve that placed everyone in similar professional circumstances, and it may have minimized the potential effect of personal factors like marital status. Additionally, while marriage may typically provide emotional support, the unique stressors of the pandemic, such as financial uncertainty, health concerns, and increased household responsibilities, may have diminished these benefits. As a result, marital status may have had little influence on teachers' perceived self-efficacy during COVID-19.

Further, when looking at the interaction of marital status and parental status, it could be seen that this difference was insignificant, considering burnout levels. A possible explanation for the insignificant interaction between marital and parental status on burnout may lie in the shared crisis environment created by the COVID-19 pandemic. The widespread uncertainty, health concerns, and disruption to daily life may have affected people across all family structures and created a collective sense of stress and emotional exhaustion. In such a context, personal circumstances like being married or having children may have had less impact on burnout levels, as all teachers, regardless of their marital or parental status, were having similar psychological and professional challenges. In terms of self-efficacy, the findings of the interaction of marital status and parental status were insignificant. Similar to the insignificant difference between single or married teachers in terms of self-efficacy levels, one possible reason for the insignificance of the interaction between marital and parental status on self-efficacy is that all teachers, regardless of their family circumstances, were exposed to similar professional challenges during the COVID-19 pandemic. The sudden transition to online education required every teacher to adapt quickly to unfamiliar technologies, manage the classroom, and maintain instructional quality under uncertain conditions. These shared professional stressors may have had a stronger influence on teachers' sense of self-efficacy than differences in personal life and it may have led to a leveling effect across marital and parental categories.

Implications for Research and Practice

In the current study, significant differences were found in years of experience and parental status for burnout levels of teachers. Also, there were significant differences in years of experience and school type with respect to the self-efficacy levels of teachers. Based on these findings, more targeted and practical strategies can be recommended. These include:

- Providing more seminars, courses, and in-service training on stress management, coping skills, and innovative teaching methods to help decrease burnout and improve self-efficacy.
- Organizing collaboration sessions where teachers interact with each other, exchange ideas about teaching strategies and online education materials, and receive social and professional support to reduce anxiety.
- Improving the quality, characteristics, and opportunities of public schools so that no differences exist among them regarding factors that influence teachers' self-efficacy levels.
- Supporting teachers who have children by reducing their workload when possible and ensuring access to reliable institutions or programs where their children can safely attend, thereby minimizing the impact of parental status on burnout.

These targeted and practical strategies can help create a healthier, more supportive working environment and ultimately improve teachers' well-being and instructional quality.

The use of a convenience sample in this study represents an important limitation that should be acknowledged. Since participants were not randomly selected, there is a risk that the sample may not accurately represent the broader population of middle school mathematics teachers in Türkiye. Consequently, the findings may not be fully generalizable to all teachers and should be interpreted with caution. Regarding the choice of sampling, the sample can be generated from a targeted population that has similar characteristics to the sample. Future studies using random sampling or more diverse sampling strategies would strengthen the external validity of the results and improve the replicability of the study.

The authors received no financial support for the research, authorship, and/or publication of this manuscript

Sinem Elif Elma (sinemelifelma@gmail.com) is a master's degree graduate in Mathematics Education at Middle East Technical University. She has worked as a mathematics teacher in a public school. Her research focuses on statistics education research, educational psychology, and teacher education. She has experience in a TÜBİTAK-funded research project that has published in national and international conferences.

Ekin Balci (ekin.balci@tedu.edu.tr) is a PhD student in Mathematics Education at Middle East Technical University and a Research/Teaching Assistant at TED University. Her research focuses on mathematics teacher education, teacher knowledge, high-level cognitive tasks, and reasoning. She has experience in national TÜBİTAK-funded research projects and has published in international conferences and journals.

Mine İşıksal-Bostan (misiksal@metu.edu.tr) is a mathematics education professor at Middle East Technical University (METU). Her research focuses on a wide range of topics including teacher education, curriculum development, and technology in math education. She has been involved in leading national and international projects, notably with TÜBİTAK, Erasmus+, and Horizon 2020, aiming to improve math education. She has mentored numerous graduate students and received multiple awards from METU.

References

- Amri, A., Abidli, Z., Elhamzaoui, M., Bouzaboul, M., Rabea, Z., & Ahami, A. O. T. (2020). Assessment of burnout among primary teachers in confinement during the COVID-19 period in Morocco: A case study in Kenitra. *The Pan African Medical Journal*, 35(92.2, Suppl.), <https://doi.org/10.11604/pamj.supp.2020.35.2.24345>
- Atta, N., Ahmad, N., Ahmed, M., & Ali, Z. (2012). Role of gender and teaching experience on teachers' self-efficacy. *Language in India*, 12(9), 251–262.
- Ávalos, B., & Valenzuela, J. P. (2016). Education for all and attrition/retention of new teachers: A trajectory study in Chile. *International Journal of Educational Development*, 49, 279–290. <https://doi.org/10.1016/j.ijedudev.2016.03.012>
- Bandura, A. (1977a). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295x.84.2.191>
- Bandura, A. (1977b). *Social learning theory*. Englewood Cliffs, N.J: Prentice Hall. <https://doi.org/10.1177/105960117700200317>
- Bandura, A. (1994). Self-efficacy. In V.S. Ramachandran (Ed.), *Encyclopedia of human behavior* (Vol.4, pp.71–81). New York: Academic Press. (Reprinted in H. Friedman (Ed.), *Encyclopedia of mental health*. San Diego: Academic Press, 1998).
- Bandura, A. (1997). *Self-efficacy in changing societies*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511527692>
- Başol, G., & Altay, M. (2009). Examining occupational burnout levels of educational administrators and teachers. *Educational Administration: Theory and Practice*, 15(58), 191–216.
- Batool, S., Atta, M., & Riaz, N. (2020). Impact of self-efficacy on job stress in teachers: The role of marital status. *Journal of Research in Social Sciences*, 8(2), 46–55. <https://doi.org/10.7176/EJBM/13-6-23>
- Bonett, R. M. (1994). Marital status and sex: Impact on career self-efficacy. *Journal of Counseling and Development*, 73(2), 187–190. <https://doi.org/10.1002/j.1556-6676.1994.tb01734.x>
- Caprara, G. V., Barbaranelli, C., Steca, P., & Malone, P. S. (2006). Teachers' self-efficacy beliefs as determinants of job satisfaction and students' academic achievement: A study at the school level. *Journal of School Psychology*, 44(6), 473–490. <https://doi.org/10.1016/j.jsp.2006.09.001>
- Collie, R. J. (2021). COVID-19 and teachers' somatic burden, stress, and emotional exhaustion: Examining the role of principal leadership and workplace buoyancy. *AERA Open*, 7(1), 1–15. <https://doi.org/10.1177/2332858420986187>
- Corry, M., & Stella, J. (2018). Teacher self-efficacy in online education: a review of the literature. *Research in Learning Technology*, 26. <https://doi.org/10.25304/rlt.v26.2047>
- Daumiller, M., Rinas, R., Hein, J., Janke, S., Dickhäuser, O., & Dresel, M. (2021). Shifting from face-to-face to online teaching during COVID-19: The role of university faculty achievement goals for attitudes towards this sudden change, and their relevance for burnout/engagement and student evaluations of teaching quality. *Computers in Human Behavior*, 118. <https://doi.org/10.1016/j.chb.2020.106677>
- Edelwich, J., & Brodsky, A. (1980). *Burnout: Stages of disillusionment in the helping professions*. New York: Human Sciences Press. <https://doi.org/10.1177/019263658106544520>
- Elflein, J. (2020). Global mental health of workers before and after COVID-19 outbreak April 2020 [Dataset]. Retrieved from <https://www.statista.com/statistics/1169810/covid-mental-health-of-workers-in-select-countires/>
- Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (8th ed.). New York: McGraw Hill.
- Freeman, L. A., & He, J. (2010). Are men more technology-oriented than women? The role of gender

- on the development of general computer self-efficacy of college students. *Journal of Information Systems Education*, 21(2), 203–212.
- Hendrickson, B. (1979). Teacher Burnout: How to Recognize It? What to Do about It. *Learning*, pp. 7, 37–39.
- Kalaian, S. A. (2008). Research Design. In Lavrakas P. J. (Ed.), *Encyclopedia of Survey Research Methods*. Thousand Oaks, California: SAGE Publications. 725-732. <https://doi.org/10.4135/9781412963947>
- Karabıyık Özipek, A. (2006). *The Reasons and the Level of Professional Burnout in Teachers Working in Secondary Schools*. Unpublished master's thesis, Trakya University, Edirne.
- Karakaş, A., & Bilen, A. (2014). Muhasebe meslek mensuplarının tükenmişlik düzeylerinin işgören performansına etkisi. *Bingöl University Journal of Social Sciences Institute*, 8(4), 139-154.
- Karanfil, F. (2019). *Burnout levels and reasons in English teachers working in state and private school in Turkey: A comparative study*. [Unpublished master's thesis]. Uludağ University, Bursa.
- Kim, L. E., & Asbury, K. (2020). 'Like a rug had been pulled from under you': The impact of COVID-19 on teachers in England during the first six weeks of the UK lockdown. *British Journal of Educational Psychology*, 90(4), 1062-1083. <https://doi.org/10.1111/bjep.12381>
- Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, 102(3), 741–756. <https://doi.org/10.1037/a0019237>
- Kumawat, K. (2020). Perceived stress and burnout in online teaching in teachers in India during pandemic COVID-19. *Indian Journal of Health & Wellbeing*, 11.
- Maslach, C., & Goldberg, J. (1998). Prevention of burnout: New perspectives. *Applied and Preventive Psychology*, 7(1), 63-74. [https://doi.org/10.1016/S0962-1849\(98\)80022-X](https://doi.org/10.1016/S0962-1849(98)80022-X).
- Maslach, C., & Jackson, S. E. (1981). The measurement of experienced burnout. *Journal of Occupational Behavior*, 2, 99-113. <https://doi.org/10.1002/job.4030020205>
- Maslach, C., & Jackson, S. E. (1984). Burnout in organizational settings. *Applied Social Psychology Annual*, 5, 133-153.
- Maslach, C., & Jackson, S. E. (1986). *Maslach burnout inventory manual* (2nd ed.). Palo Alto, CA: Consulting Psychologists Press.
- Maslach, C., & Schaufeli, W. B. (1993). Historical and conceptual development of burnout. In W. B. Schaufeli, C. Maslach, & T. Marek (Eds.), *Professional burnout: Recent developments in theory and research* (pp. 1-16). Philadelphia: Taylor & Francis.
- Monroy-Gómez-Franco, L., Vélez-Grajales, R., & López-Calva, L. F. (2022, May). The potential effects of the COVID-19 pandemic on learnings. *International Journal of Educational Development*, 91, 102581. <https://doi.org/10.1016/j.ijedudev.2022.102581>
- Mosleh, S. M., Kasasbeha, M. A., Aljawarneh, Y. M., Alrimawi, I., & Saifan, A. R. (2022). The impact of online teaching on stress and burnout of academics during the transition to remote teaching from home. *BMC Medical Education*, 22(1), 475.
- OECD. (2021). *The State of School Education: One Year into the COVID Pandemic*. OECD Publishing: Paris. <https://doi.org/10.1787/201dde84-en>
- Ozamiz-Etxebarria, N., Idoiaga Mondragon, N., Dosil Santamaria, M., & Picaza Gorrotxategi, M. (2021). Stress, anxiety, and burnout levels in the first phase of the COVID-19 outbreak in a sample of Spanish teachers. *European Journal of Education and Psychology*, 14(1), 95-110. <https://doi.org/10.32457/ejep.v14i1.1307>
- Özler, D. (2015). *Örgütsel Davranışta Güncel Konular*. İstanbul: Ekin Basım Yayın Dağıtım.
- Pallant, J. (2016). *SPSS Survival Manual: A Step By Step Guide to Data Analysis Using SPSS* (6th ed.). London, UK: McGraw-Hill Education.
- Pierce, C. A., Block, R. A., & Aguinis, H. (2004). Cautionary note on reporting eta-squared values

- from multifactor anova designs. *Educational and Psychological Measurement*, 64(6), 916-924. <https://doi.org/10.1177/0013164404264848>
- Pressley, T. (2021). Factors contributing to teacher burnout during COVID-19. *Educational Researcher*, 50(5), 325–327. <https://doi.org/10.3102/0013189X211004138>
- Ptacek, R., Ptackova, H., Martin, A., & Stefano, G. B. (2022). Teacher burnout syndrome during COVID-19. *Medicina*, 58(5), 638.
- Putri, R. S., Purwanto, A., Pramono, R., Asbari, M., Wijayanti, L. M., & Hyun, C. C. (2020). Impact of the COVID-19 pandemic on online home learning: An explorative study of primary schools in Indonesia. *International Journal of Advanced Science and Technology*, 29(5), 4809-4818.
- Reyhing, Y., & Perren, S. (2021). Self-efficacy in early childhood education and care: What predicts patterns of stability and change in educator self-efficacy? *Frontiers in Education*, 6. <https://doi.org/10.3389/feduc.2021.634275>
- Robinia, K. A., & Anderson, M. L. (2010). Online teaching efficacy of nurse faculty. *Journal of Professional Nursing*, 26(3), 168–175. <https://doi.org/10.1016/j.profnurs.2010.02.006>
- Sağlam Arı, G., & Çına Bal, E. (1983). Tükenmişlik Kavramı: Birey ve Örgütler Açısından Önemi. *Yönetim ve Ekonomi Dergisi*, 15(1), 131-148.
- Sánchez-Pujalte, L., Mateu-Martínez, O., Cabezas-González, M., & Rubio-Hurtado, M. J. (2021). Teachers' burnout during COVID-19 pandemic in Spain: Trait emotional intelligence and socioemotional competencies. *Sustainability*, 13(13). <https://doi.org/10.3390/su13137259>
- Savicki, V., & Cooley, E. (1983). Theoretical and research considerations of burnout. *Children and Youth Services Review*, 5(3), 227–238. [https://doi.org/10.1016/0190-7409\(83\)90028-2](https://doi.org/10.1016/0190-7409(83)90028-2)
- Scull, J., Phillips, M., Sharma, U., & Garnier, K. (2020). Innovations in teacher education at the time of COVID-19: an Australian perspective. *Journal of Education for Teaching*, 46(4), 497–506. <https://doi.org/10.1080/02607476.2020.1802701>
- Semerci, D., & Uyanık Balat, G. (2018). Examining the relationship between preschool teachers' classroom management skills and self-efficacy Perceptions. *Inonu University Journal of the Faculty of Education*, 19(3), 494-519.
- Shapiro, S. S., & Wilk, M.B. (1965). An analysis of variance test for normality (Complete Samples). *Biometrika*, 52, 591-611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Shlenskaya, N., Karnaukhova, A., Son, L., & Lapteva, E. (2020, July). Teachers' burnout in online university courses in the time of pandemic. In *2020 the 4th international conference on education and multimedia technology* (pp. 95-99).
- Skaalvik, E. M., & Skaalvik, S. (2010). Teacher self-efficacy and teacher burnout: A study of relations. *Teaching and Teacher Education*, 26(4), 1059–1069. <https://doi.org/10.1016/j.tate.2009.11.001>
- Sokal, L., Trudel, L. E., & Babb, J. (2020). Canadian teachers' attitudes toward change, efficacy, and burnout during the COVID-19 pandemic. *International Journal of Educational Research Open*, 1, 100016. <https://doi.org/10.1016/j.ijedro.2020.100016>
- Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th ed.). Allyn & Bacon/Pearson Education.
- Weber, A., & Jaekel-Reinhard, A. (2000). Burnout syndrome: a disease of modern societies? *Occupational Medicine*, 50(7), 512–517. <https://doi.org/10.1093/occmed/50.7.512>
- Weißenfels, M., Klopp, E., & Perels, F. (2021). Changes in teacher burnout and self-efficacy during the COVID-19 pandemic: Interrelations and variables related to change. *Frontiers in Psychology*, 12, 705484.

- Yavuz, M. (2009). An investigation of burnout levels of teachers working in elementary and secondary educational institutions and their attitudes to classroom management. *Educational Research and Reviews*, 4(12). 642-649
- Yerdelen, S., Sungur, S., & Klassen, R. M. (2016). Relationship between Turkish elementary science teachers' occupational well-being and some contextual and demographic characteristics: A multivariate analysis. *Education and Science*, 41(183). <https://doi.org/10.15390/eb.2016.4257>
- Yıldız, M. (2015). *The Relationships Among Distance Education Instructors' Knowledge, Belief and Practices Towards Distance Education*. [Unpublished master's thesis]. Hacettepe University, Ankara.
- Yükseköğretim Kurumu [YÖK]. (1998). *Eğitim Fakültesi Öğretmen Yerleştirme Programları*. Ankara. <https://www.yok.gov.tr/Documents/Yayinlar/Yayinlarimiz/egitimbakultesi-ogretmen-yetistirme-lisans-programlari-mart-1998.pdf>
- Zadok-Gurman, T., Jakobovich, R., Dvash, E., Zafrani, K., Rolnik, B., Ganz, A. B., & Kaplan, Y. (2021). COVID-19 pandemic: A nationwide study of burnout and depression among Israeli teachers during the school reopening period. *International Journal of Environmental Research and Public Health*, 18(8), 4029. <https://doi.org/10.3390/ijerph18084029>

High School Students Explore Science Identity Through Participatory Science in Informal Education Program

Charlie K. Blake 

Southern Illinois University Edwardsville STEM Center

Andreia F. Dexheimer 

Southern Illinois University Edwardsville STEM Center

Carol E. Colaninno 

Washington University in St. Louis

Candice L. Johnson 

Saint Louis University

Ben K. Greenfield 

Public Health Program, Muskie School of Public Service, University of Southern Maine

Lillian C. Ward

Southern Illinois University Edwardsville STEM Center

Ozaveshe Paul Amune 

Southern Illinois University Edwardsville STEM Center

Adriana E. Martinez 

Southern Illinois University Edwardsville STEM Center

Sharon M. Locke 

Southern Illinois University Edwardsville STEM Center

Georgia Bracey 

Southern Illinois University Edwardsville STEM Center

ABSTRACT

Science identity is central to students' desire to continue their STEM learning, yet many students struggle to see themselves as scientists. To create learning environments where all students can see themselves in science, we need a comprehensive understanding of the elements that contribute to student science identity. We developed and implemented a community-focused participatory science curriculum, examined how well Carlone and Johnson's (2007) three-part model of science identity—based on performance, competence, and recognition—captured students' experiences, and investigated novel elements of science identity that emerged from data voiced by the students who participated in our program. We employed mixed-methods analysis of pre- and post-program surveys, a repeated measures survey, and interviews from our ethnically and socioeconomically

diverse high school student population. We suggest additional elements of science identity that emerged from our data and show how participation in our program impacted students' science identity. We found that students' definitions of a scientist can interfere with their ability to recognize role models or see themselves as scientists, highlighting the importance of countering stereotypes and emphasizing that performing scientific tasks is a key component to becoming a scientist. Our findings are relevant to educators seeking to encourage science identity development. We suggest that educators should support students to broaden their definitions of a scientist into definitions that could include themselves.

Keywords: citizen science, curriculum, environmental health, science identity

Introduction

Science identity is central to students' desire to continue their STEM learning. Science identity is defined as the aspect of the self that relates to science, i.e. the part that wants to be, and be perceived as, a “science person” (Carlone & Johnson, 2007; Gee, 2014; McDonald et al., 2019). Creating opportunities for science identity formation in all students can be an important way to foster a diverse future STEM workforce and a scientifically literate society. However, many students struggle to connect with STEM fields or to see themselves as scientists, especially marginalized students (Andersen et al., 2014). To create environments where all students can see themselves in science, we need a comprehensive understanding of student science identity and the factors that influence its development.

Carlone and Johnson (2007) posited that the prototypical science person would have a certain level of competence, be able to demonstrate it, and be acknowledged for it by others, resulting in a model of identity based on performance, competence, and recognition. In this study we examined Carlone and Johnson's (2007) science identity model through an out-of-school program that integrated science identity exploration with a student curriculum centered on empowering high school students as scientists. Bringing together concepts of place-based education, participatory science, and science identity in education, we conducted a mixed-methods study to address the following research questions:

- How well does Carlone and Johnson's (2007) model capture our students' experiences?
- In what ways does engagement in community-focused environmental participatory science interact with science identity among high school students?
- How do individual levels of performance, competence, and recognition relate to a student's sense of being a scientist?

Place-based Education and Exposure Science

Place-based education situates the learning experience in the place where learners live. The underlying premise is that students find greater meaning and thus learn more effectively when they closely examine and contribute to improving their own communities and ecosystems (Sobel, 2004). Smith and Sobel (2010) argued that by examining and understanding their own local environment, learners develop academic skills, a sense of relevance, connection to their environment, and strengthened ties to their community. In developing our program and curriculum, we applied place-based education with a focus on the fields of environmental health and exposure assessment. Exposure science, a subfield of environmental health, evaluates and quantifies the level of exposure to physical, chemical, or biological agents in the environment that can affect human health (Barr, 2006;

Lioy, 2010). Exposure science intersects well with place-based education given the emphasis on exposure to one's local environment, the potential health consequences of that exposure, and the ability to alter their exposure through behavioral or environmental modifications in response to monitoring results (Sobel, 2004, pg. 12). In our curriculum, we combined these concepts by giving students opportunities to explore their own exposures in their environments using sensor-based monitoring tools.

Participatory Science

Participatory science¹ describes the practice of people, often volunteers or nonprofessionals who may not have traditional scientific credentials, participating in scientific research. Modern participatory science programs in the United States were popularized by the Cornell Lab of Ornithology in 1995 with bird-watching programs where volunteers assisted scientists with research (Bonney et al., 2015). The 1990s and early 2000s saw significant growth in organized participatory science programs, along with increasing recognition of data collected by non-professional community scientists as valuable for scientific research (Kosmala et al., 2016; McKinley et al., 2017; Sullivan et al., 2014).

We use participatory science in our curriculum to introduce students to the idea that anyone who conducts scientific inquiries is doing science, and that doing science is not reserved only for people with academic or professional credentials. Our students engaged with ongoing national participatory science/citizen science projects, as well as collecting environmental health data in their own community. This approach, embedded within our place-based curriculum, ensured students engaged in genuine and authentic scientific research experience. STEM subjects are more engaging to students when they can connect what they are learning to real-world issues that are relevant to them and the places they care about (Honey et al., 2014; Smith & Sobel, 2010). Further, ample evidence shows how participatory science can play a significant role in effective and community-engaged science education, exposing participants not just to content knowledge from a specific citizen science project but also foundational connections to the nature and practice of science (Pandya & Dibner, 2018). By incorporating participatory science in our curriculum, we hoped to help students begin to see themselves in the role of scientist, and as members of the scientific community.

Science Identity

The development of a science identity involves a multifaceted process, and a series of science identity models have been proposed and tested. Based on their interviews with women of color in science academic programs or professions, Carlone and Johnson (2007) posited that the prototypical science person would have a certain level of competence, be able to demonstrate it, and be acknowledged for it by others, resulting in a model of identity based on *performance*, *competence*, and *recognition*. Performance refers to social performances of relevant scientific practices such as ways of talking and using tools. Competence refers to knowledge and understanding of science content, which may be less publicly visible than performance. Recognition refers to recognizing oneself and getting recognized by others as a science person (Carlone & Johnson, 2007; Hazari et al., 2010; Roberts & Hughes, 2022). Roberts and Hughes (2022) also demonstrated how an informal learning environment, like a strategically designed science summer camp, can provide increased external recognition, an important component of science identity, especially for marginalized or underrepresented students.

¹ Our curriculum materials produced during earlier years of our program used the term citizen science. We use the term participatory science in this paper following the recent shift towards more inclusive language (Citizen Science Association, 2023).

Carlone and Johnson's (2007) model also includes race, ethnicity, and gender identities as an important contextual factor influencing the construction of science identity.

Other researchers have built upon and modified Carlone & Johnson's (2007) model. Hazari et al. (2010) added a fourth component to science identity, *science interest*, which they define as the desire and curiosity to think about and understand science, which relates to external and internal motivation. Hazari et al. (2020) and others have also proposed that performance and competence can be combined into one construct, at least in some learning contexts, which is supported with quantitative and qualitative data from undergraduate students in physics, computer science, and engineering (Godwin et al., 2016; Große-bölting et al., 2023; Mahadeo et al., 2020). McDonald et al. (2019) simplified science identity into one component and used a single-item survey to measure self-recognition. Roberts & Hughes (2022) used Carlone & Johnson's (2007) model as their theoretical framework and focused only on recognition, separating internal and external recognition. Using empirical data, Estrada et al. (2011) proposed a different model for integration of underrepresented students in the scientific community composed of science identity, self-efficacy, and scientific values. Several of these studies have agreed that science identity is a key element for both intentions to pursue a career in science and also motivation to persist (Estrada et al. 2011). We approached the variations in the existing literature as different lenses that could be used to understand science identity, and the best lens to use may vary depending on the specific groups of students and their learning context. Considering these differing versions of science identity models, we sought to explore which most important elements apply in our context, an informal environmental science program with a diverse group of high school students.

Theoretical Framework

We used Carlone and Johnson's (2007) science identity model as a theoretical framework to guide our research, with awareness of the later modifications proposed by other researchers discussed above. This framework shaped our study instruments and our codebook as we conducted analysis, and we were attuned to look for ways in which data from our student population did, or did not, fit this model. Our study focused on testing the role of, and relationship among, student perceptions of performance, competence, and recognition as they relate to science identity in the context of our student program.

Materials and Methods

Program Design

We developed an engaging, community-focused, environmental science summer program for high school students with the aim to foster participant explorations of science identity and implemented the program along with high school teachers hired as co-instructors. We designed the program to give students the opportunity to participate in authentic science in a real-world context, interact and communicate with their community and professional scientists, act as and be seen by others as scientists, and begin to see themselves as potential members of the scientific community. The Southern Illinois University Edwardsville (SIUE) STEM Center implemented this curriculum in an out-of-school summer program to situate the activities more directly in the broader community setting, leveraging the flexibility, community engagement, and authenticity that these learning environments often provide (Bell et al., 2009; National Research Council, 2009).

In summer 2021 we held a summer program in-person at a local high school. A limited number of students (seven) were recruited to keep the class size small and allow for social distancing during the COVID-19 pandemic. We offered a similar program in the summer of 2022 based at the SIUE STEM Center on campus, and expanded our enrollment to 30 students. We partnered with an Upward

Bound Math and Science program that served students from low-income families and families in which neither parent holds a bachelor's degree. Both summer programs contained similar content but summer 2021 was more compressed with nine days of programming over two weeks while summer 2022 was 13 days of programming over three weeks.

Student Curriculum

We developed a modular out-of-school curriculum in environmental sciences where high school students designed, implemented, and communicated findings from technology-rich, participatory environmental science research projects. Because our objective for the curriculum was to position students as scientists, we intentionally sequenced activities that included each step of the scientific process. The curriculum included topics with a specific focus on environmental health including air pollution, noise pollution, soil pollution, and the natural and built environment, with connections throughout to participatory science projects like Purple Air, iNaturalist, and the Noise Project (*iNaturalist*, 2025; *Noise Project*, 2025.; *Real-Time Air Quality Monitoring by PurpleAir*, 2025.). Our team, along with a group of experienced high school teachers, hired as co-instructors, then guided students in research study design, data collection, analysis, and communication of findings. The students also had many opportunities to interact with scientists, including guest speakers, university partners, and authors. With instruction from scientists and educators, students developed scientific research questions and considered what kinds of questions they could answer with the available equipment and time. As we helped students develop research projects, we emphasized personal connections and using science to understand and help their own community and environment. Students worked in small teams to create research posters and presented them at a research symposium to community members, peers, and school administrators (Table 1).

Table 1

An Overview of the Student Curriculum from Summer 2022

Summer Program	Description
Day 1: Citizen science and the nature of science	Students tour campus and learn about campus life. They learn about citizen science, its role in the scientific community, and submit data to iNaturalist. Students write blog posts about everyday interactions with science.
Day 2: Environmental health and soil pollution	Students learn about types of pollution and health impacts on people. Students collect soil samples, test soil chemistry, and map impacts of soil lead pollution. A guest scientist visits from the biology department.
Day 3: Air and noise pollution	Students learn about air quality, citizen science mapping of air pollution with PurpleAir, and graph data from their region. They learn about sound waves, noise pollution, and the Noise Project. Students analyze local noise data alongside Google Earth data to identify patterns of noise pollution and collect their own data with decibel meters.
Day 4: Humans across the landscape and introduction to research projects	Students learn about mapping land cover and impacts of human-modified landscapes. They learn about all of the tools they can use to design their own study, such as soil probes, X-ray fluorescence analyzer, and drones. Guest scientists visit from a nearby university.

Day 5: Research project planning and intro to data collection	Students practice collecting data and testing hypotheses for a short practice study. Students begin planning their own research studies in groups.
Day 6: Research design and planning	Students learn about sampling from a population with M&Ms, and about observational studies vs. experiments. Students visit three environmental scientists' labs on campus. They finalize research plans.
Days 7 and 8: Data collection	Students collect data for their research project.
Day 9: Understanding variables and using tables	Students learn about explanatory and response variables. They use noise and soil data collected by previous students near a local steel mill to practice using Google Sheets. The students make their own data tables.
Day 10: Data visualization and analysis	Students learn about what type of graph is best for which type of data, with examples of environmental data collected and graphed by previous students. Students make graphs of data from their own research.
Day 11: Science communication	Students learn about effective science communication and use a template to start creating their own scientific poster to share their findings.
Day 12: Scientific posters	Students continue working on creating scientific posters or can work on blog post or lab report.
Day 13: Poster symposium	Students practice presenting their posters with peers and then present posters at symposium for invited guests (parents, university community, teachers).

Our approach to crafting the curriculum employed the backwards design framework (Wiggins & McTighe, 2005), where we identified our expected learning outcomes and designed activity plans that aligned with those outcomes. We also employed the backwards design method at the whole curriculum level with elements of science identity in mind. For example, we positioned the poster symposium as our culminating activity because we knew we wanted students to be able to perform and be recognized for knowing and doing science through presenting their scientific poster. The curriculum underwent iterative development cycles consistent with the principles of design research (Reeves et al., 2005). Although the implementation described here was an informal summer program setting, the curriculum is also aligned for use in formal classrooms with U.S. national Next Generation Science Standards (NGSS), in an effort to make the curriculum useful to other teachers and educators who may be required to meet these standards within a school context. The curriculum partially fulfilled ESS3.C: Human Impacts on Earth Systems, ESS3.D: Global Climate Change, HS-ESS3-3: Create a computational simulation to illustrate the relationships among the management of natural resources, the sustainability of human populations, and biodiversity; HS-ESS3-4: Evaluate or refine a technological solution that reduces impacts of human activities on natural systems, HS-ESS3-6: Use a computational representation to illustrate the relationships among Earth systems and how those relationships are being modified due to human activity, PS4.A: Wave Properties, and HS-PS4-1: Develop a model of waves to describe patterns in terms of amplitude and wavelength and that waves can cause objects to move. Additionally, the curriculum incorporated most of the NGSS cross-cutting

concepts, and covered all of the NGSS science and engineering practices (NGSS Lead States, 2013). We also used Carlone and Johnson's three science identity components (competence, performance, and recognition) to guide the design of our curriculum to support student identity exploration. For example, we designed activities to give students ample opportunity to perform scientific skills and receive recognition for their expertise. The final version of the full curriculum and supplemental lesson materials are freely available to download at: <https://stemcenter.siu.edu/resources/curriculum/>

Environmental and Sociocultural Setting

The students we worked with live in communities situated in the Metro East region of Illinois, United States, in a predominately postindustrial area directly across the Mississippi River from the larger metropolitan area of St Louis, Missouri. This region has a rich history in music and Black culture and has long served as an important hub of manufacturing and transporting industrial products. Following the ebb of the railroad economy and postindustrial decline, portions of the Metro East experienced large decreases in population in the later part of the 1990s and were left with a legacy of environmental contamination and a distressed economy (Nunes & Theising, 2011). An authentic place-based context for environmental learning is particularly important for the students in our focus area, because these communities have faced a long history of environmental injustices, with disproportionate exposure to industrial pollution and environmental contamination (Abel, 2008; Colten & Samsel, 1988; Kaminski & Landsberger, 2000; Vermillion et al., 2005). The population we worked with also included some students from rural, agricultural communities, and some from exurbs favored by people who work in the larger urban center of St. Louis across the river. The history of the Metro East shapes the environmental context and sociocultural background of the students we worked with, as well as the content of the place-based curriculum we developed.

Student Demographics

We obtained parental consent and participant assent for all 37 students who attended the 2021 and 2022 summer programs, which gave us an ethnoracially diverse student sample (Table 2). Our sample of students is fairly representative of the student population in the area, with a slightly higher number of Black, Hispanic/Latinx, and multiracial students and a lower number of White students as compared to the district as a whole. The demographics of the nearest school district are 52% White, 14% Black, 27% Hispanic, 6% multiracial, and the remainder belong to racial/ethnic groups not sampled in our study (Illinois Report Card, 2023). We believe our findings to be representative of, and generalizable to, other diverse student populations that are from low-income families and/or would be the first in their families to attend college.

Table 2

Student Demographics

“I am...”		“I am...”		“What grade are you starting?”	
Male	54%	Latinx or Hispanic	32%	9th	30%
Female	38%	White	32%	10th	10%
Nonbinary	3%	Multiple	14%	11th	27%
Not answered	5%	African American or Black	14%	12th	23%
		Not answered	8%	Not answered	7%

Note: From surveys of 37 total students in (7 in summer 2021, and 30 in summer 2022). Percentages are rounded to the nearest whole number.

Research Design

We employed a concurrent triangulation mixed methods approach using both quantitative responses to surveys and qualitative data derived from student interviews. We chose a mixed methods approach as the most appropriate given our sample size and research questions, with the qualitative data providing vital nuance and corroboration when quantitative data were inconclusive or difficult to interpret on their own (Creswell et al., 2003). We collected and analyzed each data type separately and integrated the data when developing our interpretations. During analysis, we considered how the theoretical model of science identity and the conceptual frameworks of participatory science and place-based learning intersected in the participants' experience and in the program activities. The survey data provided us with a means to measure the students' science identity as they participated in the program and their responses to the interview questions allowed us to explore what led students to feel the way they did about their science identity. Throughout our analysis, we focused on direct student perspectives and prioritized the qualitative data in order to develop a rich, contextualized understanding. This study protocol was approved by the Southern Illinois University of Edwardsville Institutional Review Board.

Instrument Development and Data Collection

Three researchers (GB, CC, and SL) developed the semi-structured interview protocol informed by Brinkmann & Kvale's (2015, p.131) concept of thematizing the interview study. The approximately 15-minute interview consisted of 12 questions that were based on Carlone and Johnson's (2007) science identity model, with an emphasis on how students perceive their ability to perform science-related tasks, their perceived ability to perform these tasks with competence, and how they view their own and others' recognition of their science-related skills (Supplementary Materials).

We interviewed the student participants during the first two days of the program, and again in the final two days of the program.

We used two different types of surveys in our quantitative data collection, a long pre- and post-program survey given at the beginning and end of the program, and a shorter repeated measures survey that was given six times throughout the program. The pre- and post-program survey had identical items, each containing 33 questions with a 5-point Likert scale (Strongly disagree, Disagree, Neither agree nor disagree, Agree, Strongly agree). The questions were divided into five blocks: Competence (seven questions), Performance (seven questions), Recognition (seven questions), Science Identity (five questions), and Future/Careers (seven questions). The first three blocks of the long pre- and post-program survey (Competence, Performance, and Recognition) were adapted from the 5-point Science Self Efficacy survey (Estrada et al., 2011). The Science Identity block of questions was adapted from the 5-point Scientific Identity survey (Estrada et al., 2011). The final block (Future/Careers) was adapted from the 5-point Career Interest Questionnaire (Tyler-Woods et al., 2010). For all sections of the long pre- and post-program survey, we adjusted the language to fit our high school student audience and to reflect our focus on environmental sciences. At the end of each survey, students answered demographic questions (Table 2). Surveys were only administered in 2022, and we collected 29 complete surveys of the 30 student participants. We administered pre-program surveys on the first two days and post surveys on the last two days of the 2022 summer program (Supplementary Materials).

In addition to the longer pre-program and post-program surveys described above, researchers administered a short, repeated measures, 8-item science identity survey twice a week during the 2022 summer program, for a total of six implementations. The short survey was intended to measure potential changes in science identity during the program that might not be captured on the longer pre- and post-program surveys. The first seven items were adapted from Trujillo & Tanner's (2014) science identity instrument, which uses a 5-point Likert scale. The final item was from McDonald and colleagues' (2019) validated instrument. For this item, respondents selected from a series of five Venn diagrams showing overlap between two circles labeled "Me" and "Science Person" ranging between no overlap (lowest value) to complete overlap (highest value) between the two circles (Supplementary Materials).

Qualitative Analysis

After interviews were recorded, we assigned pseudonyms and used an artificial intelligence transcription service to generate draft transcripts from the audio files (Otter.ai). Once generated, one researcher reviewed and corrected the transcript draft using the original audio file. We developed an *a priori* coding framework based on Carlone and Johnson's (2007) model, as well as an initial emergent coding of interviews collected during the 2021 program implementation. These researchers then applied the incipient coding frame to a representative sample of pre- and post-program interviews (i.e., four pre-program interviews and three post-program interviews) to refine the coding frame. While applying the incipient coding frame, the researchers remained open to divergence and subtleties in the data that they perceived as potential emergent codes (Saldana, 2021). After an initial application of the coding frame, the researchers met to review their perceived emergent codes, discuss the coding frame, and further examine excerpts that seemed relevant to the research questions but lacked alignment with any of the *a priori* codes. This process resulted in a codebook that included 14 primary and 29 secondary codes (Supplementary Materials).

Following the development of the codebook, six researchers applied the codebook to 68 interviews from 2021 and 2022 (of our 37 total students, 36 completed pre-program interviews and 32 completed post-program interviews), meeting weekly to review and discuss the application of the codebook. Two researchers independently coded each interview with a more experienced researcher

providing a third analysis of seven interviews (10%) to ensure that all researchers applied the codebook with fidelity and consistency. During our initial coding, we coded the interviews for the presence of codes without indicating the direction of association of the particular quote with the code. One researcher (CB) did a second round of coding where they coded each previously identified quote according to whether the student's comment was positively or negatively related to the concept. For example, a clear positive comment about science identity was coded as strong positive science identity, while a clear negative statement about science identity was coded as weak science identity. We conducted a similar process for identity, recognition, and competence, separating the students' quotes into negative association, moderate or unclear, or strong positive association with the main code (supplemental Table S1). After we finalized the coding process, we then held weekly meetings over the course of five weeks to further condense the codes into themes that informed our science identity conceptual model.

Quantitative Analysis

For the long surveys, our analysis included responses from the 29 students who completed both pre-program and post-program surveys before and after the program. We calculated the mean score of student responses for each research question block (i.e., Performance, Competence, Recognition, Identity, and Career/Future) and compared pre-program and post-program survey scores using Wilcoxon's matched-pairs signed-ranks test.

For the short surveys, our analysis only includes the 18 students who completed all six repetitions of the short survey during the program. We conducted two separate analyses on the short survey, considering items 1-7 as a separate instrument from item 8. Data from the short surveys were analyzed using a Friedman's rank test as a non-parametric repeated measures analysis tool. Prior to analyzing the short survey results for trends in science identity, we examined empirical correlations among the questions using Spearman Rank Correlation Coefficient (ρ). The nonparametric Friedman's test was performed on the short survey results to check whether student responses varied overall among the six sampling dates ($N = 18$ students). All quantitative data analyses were conducted in R version 4.1.2 (R Core Team, 2023).

Methodological Triangulation

To integrate the quantitative and qualitative data, two members of the research team (CB and AD) met frequently during the analysis phase. CB triangulated qualitative findings from student interviews with the quantitative trends found in those same students' surveys, for example examining if students who spoke enthusiastically about the impact of the program in interviews also showed changes in their survey response from pre- to post-. AD steered the quantitative analysis and interpretation based on the qualitative findings. During triangulation meetings, the researchers considered areas of agreement with qualitative and quantitative data and corroborated these areas in the context of the science identity theoretical model with attention to the participatory science and place-based framework. All researchers worked together to integrate the qualitative and quantitative datasets until they reached consensus on their interpretations of these data.

Results

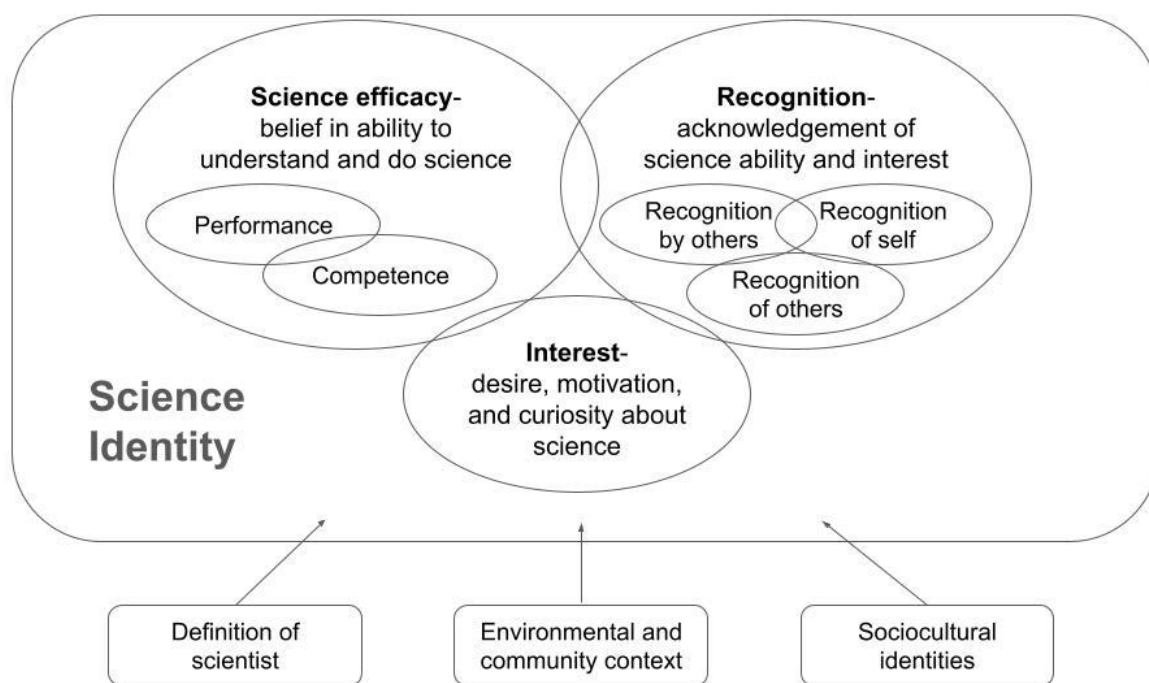
Qualitative Findings

Model of Science Identity for our Students

Through our analysis of the student interview data, we developed a model that modifies and adds to existing models of science identity (Figure 1). We started our research with Carlone and Johnson's (2007) model describing three dimensions of science identity: competence, performance, and recognition. Our model for science identity modifies and adds to those dimensions suggested by earlier scholars and has three main dimensions: *science efficacy*, *recognition*, and *interest*, all of which are influenced by contextual factors, the student's *definition of a scientist*, *environmental and community context*, and *sociocultural identities*. All of these elements are expected to change over time, and our data collection points represent snapshots of each student's relationship to these elements of science identity. Each of these dimensions is explained in more detail in the sections below, with examples of how our data supports the way we have constructed this model.

Figure 1

Our Model of Science Identity



Note: Our model includes elements of earlier models with several important additions based on our analysis of student data from our program.

Science Efficacy

In our proposed model, both competence and performance are grouped under the umbrella of science efficacy. Science efficacy in our model encompasses both scientific knowledge and understanding, as well as the skill to perform hands-on science tasks. We define science efficacy as the student's self-perceived ability to successfully know and enact science. We found it difficult to distinguish competence (knowledge and understanding of science content) from performance (social performances of relevant scientific practices) in our conversations with students. In fact, most of the

over 150 interview excerpts coded for performance were double coded as competence. In contrast to Carlone and Johnson (2007), we did not collect GPA or assessment test scores to directly evaluate performance or competence, but rather asked students to describe their self-perceived scientific competence and performance in the interviews. When coding, we often could not disentangle student's perceptions of their competence and performance when they spoke to their belief in their science ability. For example, in his 2022 pre-program interview when we asked Ethan, "*What would you do to help in studying the environment?*" he replied, "*I could take more classes. You know, learn more about it so I can help contribute, to help fix anything that's going on. I think that's what I can do. Take more classes.*" This interview question was intended to ask students about what they could do, as in actions they could take to conduct environmental research. However, we saw students had varying interpretations of what it meant to "study" the environment. Even when we discussed this quote as researchers, it was difficult for us to determine Ethan's meaning. It is possible he was conflating acquiring knowledge through classes (building competence), with performance of scientific studies because he did not consider performance and competence to be appreciably different components of his science efficacy. Or, he may have described taking classes as a way to build scientific competence because he interpreted the interviewer's question as pertaining to competence instead of performance, or science efficacy in general. However, it is possible Ethan meant that taking more classes (i.e., acquiring greater competence) was a necessary step that would precede taking action to study the environment (performance).

In addition to the quote above, many other students did not make a distinction between scientific knowledge and the skill to perform scientific tasks when responding to the interview questions, and/or interpreted questions intended to be about one of these aspects of efficacy as pertaining to either or both competence or performance. Students also often mentioned that they enjoyed learning science and doing science without making a strong distinction between the two, and that for them learning science has been their way of doing science. The difficulty of making the distinction between performance and competence in both crafting the interview questions and in interpreting student responses further justifies the way we present the relationship between competence and performance in our model. Moreover, because we designed the program to be student-led and hands-on, our approach had little division between content and practice, which may have contributed to the blurring of these two concepts in student responses to the interview questions. Given these findings, we grouped performance and competence under science efficacy in our model.

Science Interest

Our model incorporates *science interest*, which we defined as a desire and/or curiosity to think about and understand science, including external and internal motivation (Carlone & Johnson, 2007; Hazari et al., 2010). Based on our definition, a person can be interested in science without perceiving that they are good at science. Some students came into the program with a preexisting interest in science. This interest was often tied to their school experience, as seen when we asked Mia in her pre-program interview what makes her feel like a scientist or a science person:

Mia: Well, I've always liked science. And so like, whenever in school, or here in the summer program, we do summer science projects, I really liked that. I think science is something that I really love to do more than anything else. Really interesting and fun.

Science interest varied among students – some expressed no interest in science, while others had a more nuanced view, like Alex when asked about being a science person in his pre-program interview:

Alex: Because I find some of the stuff interesting. Like physics and chemistry, that stuff is really cool. Being able to, like, mix two things together, and get something completely different...I find that pretty interesting...geology and that kind of stuff, I don't find that interesting.

Alex's interest in specific STEM disciplines but not others lead him to express a more complex identification with the term science person, illustrating the role of interest in science identity. Further, we also saw evidence of students expressing science interest that they connected to future potential careers. When asked if she feels like a science person, Camila talked about her identity as a science person in relation to her interests, including an interest in the healthcare field:

Camila: *Well, I think it's always been my favorite subject. I like learning about living things, or just things around us, like what we've just been learning. And then I like the healthcare field, and that has a lot to do with science."*

Camila's expression of a specific career interest suggests an intention to reengage, which is considered by Hidi and Renninger (2006) to be an important characteristic of the ultimate phase of well-developed individual interest. Although we did not measure future reengagement directly, we saw in our interviews that the intention to reengage and independently pursue an interest in science factored into how students talked about their science identity.

Recognition

Our model divides the recognition component into recognition of self, recognition by others, and recognition of others (Figure 1). Recognition of self and recognition by others are similar to Carlone and Johnson's (2007) differentiation between self-recognition and recognition from meaningful others. We did not differentiate between meaningful and non-meaningful others when approaching how a student perceives their identity. In Carlone and Johnson's (2007) model, the inclusion of recognition by others is especially important for people who are from marginalized groups. We also added a third type of recognition, which we named recognition of others, which refers to a student's recognition of the experts we exposed them to as scientists.

Overall, we found that many students had positive associations with the idea of being recognized as a science person. Students mentioned that others thinking of them as a science person would make them feel smart, curious and fun, or proud. In her pre-program interview, when asked how it would feel to be seen as a scientist, Camila put it this way:

Camila: *I think that would make me feel just proud of myself. It shows what I have accomplished, and people can actually recognize me for things that I do.*

INTERVIEWER: *Awesome. Okay. How does it make you feel to think that your family thinks you're a science person?*

Camila: *Um, I think it's really cool, really interesting just because there's not really...a lot of my family...like, my parents are immigrants. A lot of my family comes from Mexico. So there's not really much...not a lot of education, I guess. So, I know that my parents, they didn't go to school for too long. So just knowing that I'm gonna go higher.*

Camila expresses that recognition by others as a science person is something that she considers important and seems to indicate that this recognition is something she wants to achieve and would be proud of. In our study, we found meaning in what we believe is an equally important factor to consider – a student's ability or inability to recognize scientific role models even when presented with them. The students were given the opportunity to interact with many scientists through the course of the program, including guest speakers and scientists who helped organize the program (some of whom are authors on this manuscript). However, students who interacted with scientists in the program had trouble identifying these professionals as scientists, especially if these people did not easily align with the student's notion of what a scientist looks like or does. One student illustrated her lingering lack of clarity by asking the researcher conducting her post interview "Um, are you guys declared scientists?" This question seems to imply not only that the student had not received our intended message that the researchers she was interacting with in the program were scientists, but also that there was some unnamed external authority that "declared" someone a scientist, rather than defining anyone who does

science as a scientist. In another example, Camila did not immediately identify participating in archaeological field work with one of our researchers as working with a scientist:

INTERVIEWER: *Have you ever met or worked with any scientists before?*

Camila: *Um, I don't think so.*

INTERVIEWER: *What did you do last week?*

Camila: *Hmm?*

INTERVIEWER: *What did you do last week?*

Camila: *Oh! I did anthropology with you.*

In the above example the interviewer is a scientist with expertise in ecological archaeology as well as science education research and is a white woman. It is possible that these characteristics did not immediately call to mind the idea of a scientist for this student. For students who responded to the question about knowing or working with scientists with an affirmative answer, it was more common for them to reference a different scientist they met in the program, Dr. Ben. This other scientist, Dr. Ben, is a white, male, environmental scientist, with glasses, in his 50s. Despite all program staff making efforts to contradict these stereotypes, students seemed to have perceived Dr. Ben as more easily fitting into their ideas of what being a scientist, and doing science looks like. For example, Camila, the same student quoted above, who had apparently forgotten her interviewer was herself a scientist, later stated in her post interview *"I think it's really cool to work with actual, real scientists. And you can see their different perspectives and insights, and then Ben how he's like an actual college professor."* The two women (one white, one Latina) who conducted Camila's pre-program and post-program interviews respectively were also scientists, and both held faculty positions at the university. Our reading of the interviews indicated that this perception and recognition of other scientists as scientists, especially any limits or barriers to this recognition, can be an important factor in the development of students' own science identity.

Definition of Scientist

Connected to the question of who students recognize and identify as scientists, is our inclusion in the model of the students' definition of scientist. This new construct is a key part of the context in which students form their science identity, and the student's definition of a scientist or science person can influence the other elements that make up science identity. That definition, which may change over time, is based on students' ideas about who a scientist is, what they look like, and what they do. As previously discussed in describing students' recognition of others, a student's exposure to a professional scientist as a potential role model does not guarantee that the student will receive the message intended by educators. We found that many students carried a strong preexisting definition of a scientist that seemed to be influencing their conception of their science identity. During our interviews, we asked students about what a scientist looks like and what scientists do. Here is one example of a student's mental image of a scientist from her pre-program interview:

INTERVIEWER: *So what about in like TV's and movies? What kinds of scientists do you usually see, like on television, and what are they doing?*

Olivia: *I don't really... Well, people think of chemists and things like that. Like making things. You know like, with the...*

INTERVIEWER: *Yeah, like, what do you picture when you picture a chemist?*

Olivia: *Kind of like, I don't really know what they're called, like, the glass thing?*

INTERVIEWER: *Oh, yeah. Like a glass beaker?*

Olivia: *Yeah, beakers, yeah things like that. They're, like, mixing stuff and then maybe it will explode. You have to wear safety goggles and stuff like that.*

Like Olivia, other students also showed they were aware of the stereotypical scientist image, some bringing up lab coats and glasses or goggles. However, we also had several students in pre-program

interviews who specifically articulated that they did not want to stereotype who a scientist could be. Some students said things like: *“Just depends on the field...it can be...anyone can be a scientist, whoever wants to be a scientist.”*

Students' experiences of the learning opportunities in our program seemed to be strongly mediated by who they considered a scientist to be, and whether this definition aligned with the activities they themselves were engaging in. As we explored students' perceptions of their science identity, we asked students both about being a “science person” and about being a “scientist.” Most students, even those who expressed a strong sense of themselves as a “science person” made an important distinction between “science person” and “scientist.” Rachel explained the difference this way in her post interview:

INTERVIEWER: *Do you feel like a scientist?*

Rachel: *I mean, scientist and science person are two different things if you think about it. Like, a scientist is somebody, they want to further their career in science, they want to be a scientist, work with other scientists, that sort of thing. Now, science person is more of somebody who is like, ‘I wouldn't mind doing this’, ‘I wouldn't mind being a scientist’. It's like, they're considering being a scientist. And a scientist is more like the person who is an esteemed scientist, like they've made their choice. Science person is somebody who's venturing into that choice, but isn't fully there yet. So do I feel like a scientist? Maybe. Maybe I could, I could see myself being a scientist.*

Seven other students expressed that to be a scientist one would need a degree or advanced training, would need to be a professional, would need to be an adult, or would need to have money and resources. One student replied when asked in his post interview if he was a scientist, *“Last time I checked, I don't have thousands of dollars and a lab yet...”* Many students expressed the idea that they had the potential to become a scientist in the future but did not currently feel like a scientist, although they had conducted scientific research in our program. These sentiments again reveal the importance of the students' definitions of what it means to be a scientist in the development of students' science identity. Even when we explained to students the concept that in participatory science, people without a degree can be scientists in their own community and guided them to engage in conducting scientific research themselves, some students retained a more restricted definition of what it means to be a scientist.

In post interviews some students expressly stated that their view of scientists had changed through participating in our program: *“Before this program, I was saying, like, Albert Einstein, that's my scientists. Or even Dr. Frankenstein, just, like, movie scientists. But yeah, I think now it's just anyone.”* We posit that these evolving views of who a scientist is and what a scientist does are important in how a student understands themselves in relation to science. If a student has a mental image of a scientist that is dissimilar from themselves in appearance, or in the type of actions they are doing, it will be harder for the student to see themselves included in their definition of a scientist, even when they have actively engaged in scientific inquiry. This definition of a scientist can potentially inform all the other aspects of science identity. For example, if a student does not consider their interest in veterinary medicine to be an interest that aligns with being a scientist based on their current definition of scientist, this would limit the impact of their interest in veterinary medicine on forming their science identity.

Environmental and Community Context

One key component of our model, the *environmental and community context*, reflects our place-based learning approach and the students' sense of responsibility and agency in environmental stewardship and the connections within their community. Our program content placed a strong emphasis on establishing connections between scientific research and students' everyday lives. We found that this curricular content was particularly effective in creating personally powerful experiences for students.

Many of our students made connections between the people in their communities and their non-human environment. During the interviews, we defined community as *the people and places where you live, like your neighborhood, school, street you live on*. During pre-program interviews, we directly asked students what they would like to learn about the places where they live and if they could think of ways they could study the environment to help their community. Students struggled with conceptualizing their community during pre-program interviews and often made general statements, for example about the effects of global change or how more people should care about the environment. After the program, students still emphasized environmental issues, but following the program students also made direct connections between the broader environment and their personal experiences, including experiences of performing science. When asked what she liked about the program during her post interview, Valentina responded:

Valentina: *I like, you know, my friends. And then doing a research project, I thought that was pretty fun because we got to do our own research by ourselves. I liked testing out the different equipment, and the soil testing, and the air monitor, that was pretty interesting...*

INTERVIEWER: *Can you think of some ways you could study the environment to help people in places where you live? Thinking about your house, your school, parks, wherever you go?*

Valentina: *I think we could put like, because with the air quality thing, I think we could put air quality testers in more places. So that it can be a resource to everyone. Like so that when you, let's say you walk into a restaurant, you can see what the air quality is. Just in case like, you know you have a health problem or something, you can see it and know that maybe you shouldn't stay there that long. And then you can find ways to improve it in that area. So, you can maybe find what's making the air quality bad.*

During post interviews, some students made direct connections between their communities and their specific research projects. For example, Justin mentioned that his neighbors would be interested in learning more about recycling and would be shocked to learn about how much recyclable material ends up in common landfills and would want to do something to help. This was a direct connection between his group's research project and his community. Students also talked about using knowledge and skills they gained at the program in their day-to-day lives, such as Valentina's quote above, mentioning the idea of placing air quality monitors in restaurants. When asked about what she liked about the program, Leah emphasized student agency and choice:

Leah: *I liked that we got to do different things. [...] And then we got to pick our own project, instead of you guys assigning one to us. So it's like I could do something I actually liked, instead of always being told I had to do it like in school.*

As we can see from these quotes, students' science identity was influenced by the program's content. Students were able to make more direct connections with their environment and their community by applying their knowledge of environmental sciences and by connecting their projects' topics and results to real-life situations. At the end of the program, students were also more aware of how to use science in their local community, which is a direct reflection of our emphasis on participatory science, and place-based learning which served as the foundation of the curriculum we developed and implemented.

Sociocultural Identities

Our partnership with the Upward Bound program ensured we reached a student population with high sociocultural diversity due to their recruitment of low-income and potential first-generation college-bound students (Table 2). We found that the influence of sociocultural identity inevitably surfaced as we discussed aspects of science identity with students. For example, in Camila's quote above, she speaks about a desire for recognition of her science identity by others, and that desire is connected to the fact that her parents are immigrants from Mexico who did not have much access to education. Her feelings about how recognition plays in to her identity as a science person are clearly

intertwined with her other identities. Although it was also not a part of our original research questions, we also saw that the sociocultural identities of the researchers, instructors, and visiting scientists also seemed to matter in how students experienced our program, such as the previously noted pattern of students like Camila more readily recognizing people as scientific role models if their sociocultural identities conformed to those stereotypically represented in science.

Associations Within the Model

Our analysis of the student interviews also yielded some intriguing associations among several elements within the model. We found a positive association between a strong science identity and recognition by others. In fact, all 17 students who expressed a strong sense of themselves as a science person in their interviews also expressed moderate to high levels of recognition by others (Table S1). Many of these students came into the program with a strong sense of their own science identity, shown here by Mia in her pre-program interview:

INTERVIEWER: *So do you think that other people think of you as a science person, like your friends and family?*

Mia: *I think my friends do because I do have some friends, like my closer friends in my science class. And they always tell me, 'Oh, ya know, how did you understand that? Like, what did you put together to know that information?' And my family? I think they do, because I'm always talking to them, 'Oh, I did this in science class, I really liked it.' You know, I really want to do something in STEM. I think science is my favorite subject. You know, I really like to do science.*

INTERVIEWER: *Yeah. So you talk to them about it. Do you help your friends out in science class?*

Mia: *Yes...all the time. [both laugh]*

INTERVIEWER: *I'm sure they love that. So you like being seen that way—you like that people think of you as a science person?*

Mia: *Yes.*

In contrast, students who lacked a strong science identity also reported that others did not think of them as science people, as described by Noah in his pre-program interview:

INTERVIEWER: *Okay, so, do you feel like you're a science person?*

Noah: *Not necessarily.*

INTERVIEWER: *Okay, so you think you have an interest?*

Noah: *Not in science, I appreciate it. But I'm not interested in a career in science.*

INTERVIEWER: *Okay, okay. So how would you describe your feelings about science?*

Noah: *It is interesting. I do, I do appreciate it. I do like learning about it. It's just not something I want to spend the rest of my life working with and doing.*

INTERVIEWER: *Okay. So do you think like your family members or like your friends see you as a science person?*

Noah: *Uhh, no I wouldn't say so.*

These two quotes illustrate a stark contrast in these students' experiences of how their science identity is perceived (or not) by others. While we cannot confirm that this association between external recognition and strong science identity is a causal relationship, our interpretation of our qualitative data suggests that recognition by others may contribute to development of a strong science identity.

Quantitative Findings

Survey Reliability

We found good internal reliability on all five blocks of our long survey. Reliability of the performance was $\alpha=0.89$, competence was $\alpha=0.93$, recognition was $\alpha=0.96$, science identity was

$\alpha=0.91$, and future plans was $\alpha=0.87$. We also found good internal reliability on all 6 blocks of the short survey (Cronbach's Alpha ranging from 0.84 to 0.93).

Long Survey

We conducted a Wilcoxon matched-pairs signed-ranks test on the five survey blocks (Performance, Competence, Recognition, Identity, and Career/Future) for the 29 students who completed both pre- and post-program surveys. Our quantitative results showed that the only component that was significantly different between pre- and post-program surveys was competence ($p = 0.016$). All other components were not different between pre- and post-program surveys ($p > 0.05$). However, see below for further discussion of science efficacy in our triangulation of qualitative and quantitative findings.

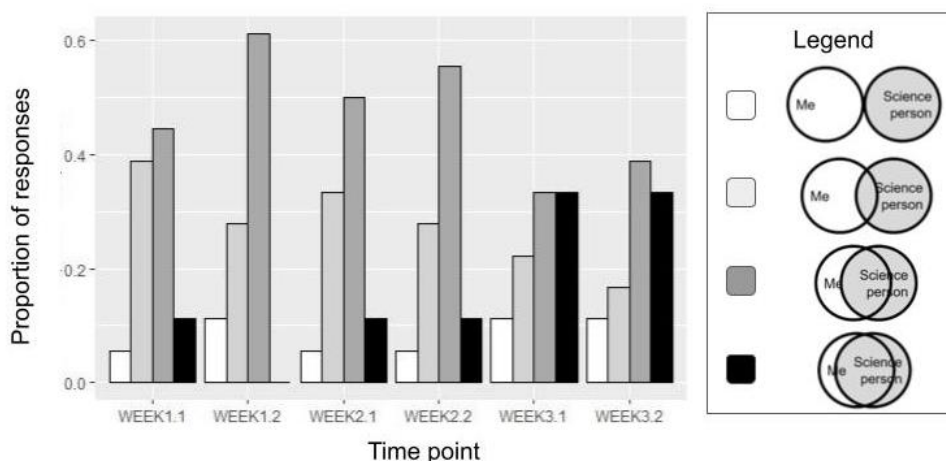
Short Survey

We used the short survey to get finer details on how students responded over time to each section of the three-week summer program. Likert responses to the short survey questions were all positively correlated with each other. The correlation between Question 5 ("I am a scientist") and the other questions was low to moderate, ranging from $\rho = 0.28$ to $\rho = 0.52$ (median $\rho = 0.39$). Correlations between the remaining questions were higher, ranging from 0.31 to 0.83 (median $\rho = 0.53$).

For the average of all eight questions, there was no significant change in student responses over the repeated measures (Friedman's $Q = 7.72$, $df = 5$, $p = 0.17$). A few participants of the total group ($N = 26$) exhibited a directional shift in their scores towards more extreme scores. In particular, for the average scores of Questions 1 to 7, Jeremiah's and Valentina's answers increased from 3.1 to 4.6 and from 4.0 to 4.9 between the first and last surveys. A trend was seen for Question 8 alone ($Q = 10.7$, $p = 0.06$), with a small effect size (Kendall's $W = 0.11$). Results for Question 8 are depicted in Figure 2. For Question 8, the number of students who selected the 4th option (out of 5 possible) for the Venn diagram overlap between "me" and "science person" increased from only a couple of students initially, to eight students in the last two surveys (Figure 2). This change shows an increase in identity as a science person during the last week of program participation for some of the students.

Figure 2

Trends In Student Survey Responses Over Time In Our Program.



Note: Students were asked to “mark the picture that best describes the overlap of the image you have of yourself and the image you have of a science person”. A fifth option with fully overlapped circles was also offered but not selected by any students.

Triangulation of Analysis and Findings

Throughout our analysis, our decisions and interpretations were informed by emerging results from both the qualitative and quantitative data. We sought to corroborate our conclusions with both our qualitative and quantitative data as much as possible, especially through considering individual students' interviews and survey results together to form a more complete picture of their individual experience. We also made use of methodological triangulation when we saw that the quantitative results indicated a weaker correlation between question five and the other survey questions. Question five asked students if they agreed with the statement “I am a scientist.” The first seven questions were adopted from the same instrument and expected to be significantly correlated, however, even students who responded more positively to the statement “I am a science person” did not always agree as strongly with the statement that they were “a scientist”. This quantitative result coincided with our emerging qualitative discovery that students made a definite distinction between being “a science person” and “a scientist” during their interviews. These parallel emerging results in turn strengthened our conviction that some of the survey questions should be analyzed separately in our statistical analysis. Taken together, these quantitative and qualitative findings support our claim that students' definition of scientist impacted their interpretation of the term scientist in both the surveys and interviews, and that this evolving definition is a key element in understanding student science identity.

Science Efficacy

Our qualitative and quantitative analysis revealed tension in how competence and performance should be framed within student identity, whether as two distinct components or related aspects of a more encompassing latent concept like science efficacy. While our quantitative data showed a significant increase in competence following our program, performance did not show a statistically significant increase when analyzed as a separate quantitative variable. Because our qualitative analysis suggested that students perceived competence and performance as intertwined, we also calculated a combined score from these sections of the survey to create a new quantitative variable *science efficacy*. We calculated a new Cronbach's alpha value for science efficacy from the 14 pre- and post-program survey questions which related to this combined variable ($\alpha = 0.90$). A Wilcoxon matched-pairs signed-ranks test comparing pre- and post-program survey scores of the new science efficacy variable showed a significant change following our program ($p = 0.013$). On average, participants increased in their science efficacy following the program by 0.27 points on the 5-point Likert scale. This suggests that a combined science efficacy variable could be a meaningful measure if student perspectives suggest they perceive performance and competence as inexorably related. Our final model shows performance and competence as two related concepts grouped within the category of science efficacy to accurately represent the nuance of both our qualitative and quantitative findings.

Shifts in Science Identity

Through considering quantitative and qualitative results together for individual students, we saw further evidence that exposure to the program supported students' exploration of their science identity. Five students expressed notable increases in describing their science identities in post interviews compared to pre-program interviews (Table 3), and some students showed a quantitative shift in their responses to the science identity questions on the short survey. For example, when Jack

started out in the program, he expressed uncertainty and an overall weak connection to science identity, “*Um...I mean, I don't see myself going to college for anything science related. But I don't particularly dislike science. I just like, I feel like most of the time when I'm in school, and they're teaching science, it's just always, you know, it's sooo... I guess they kind of killed it for me a little bit.*” However, following the program, Jack's responses to the Venn diagram science identity question on the short survey (question 8) increased by two full steps out of the five choices given, and similarly he showed a stronger sense of himself as a science person and scientist in his post interview responses:

INTERVIEWER: *Okay, do you feel like a science person?*

Jack: *Um...yeah.*

INTERVIEWER: *Yeah? Can you tell me why?*

Jack: *Well, I know a decent amount of science, about science. And we ran some tests. And, like, we collected data, observations. So I think I'm a science person.*

INTERVIEWER: *Nice. Yeah. That's all science for sure. Do you feel like you're a scientist?*

Jack: *Yeah.*

INTERVIEWER: *Yeah? Cool. Can you talk a little bit about why you feel like a scientist?*

Jack: *Well, because I've done some things that scientists probably do. Like I've collected data, and you know, have a hypothesis.*

INTERVIEWER: *Yeah. Cool.*

Jack: *Made observations.*

Jack is actually one of only a few students who expressed identification with the term “scientist.” Jack expressed that the student-led research experience made him feel like a scientist, which is a noticeable increase in his sense of science identity compared to his perspective before participating in the program just a few weeks earlier. He pointed to specific hands-on activities like collecting data that felt authentic to him, and to important concepts of scientific research like having a hypothesis. Valentina was another student whose science identity scores increased on the short survey over the course of the program. Valentina also described the student-directed group research projects as memorable: “*I thought that was pretty fun. Because we got to do our own research by ourselves.*” These personal experiences they described in interviews about conducting scientific activities, such as research question development, data collection, and data analysis, coincide with these two students' quantitative increase in their science identity responses on the short survey.

Table 3

Strength of Science Identity Self-described by Students in Pre- and Post-Program Interviews

Responses to the question: “Do you feel like a science person?”	Students in pre-program interview (n=36)	Students in post-program interview (n=32)
Negative	13.9%	12.5%
Mixed	41.7%	31.2%
Strong Positive	41.7%	56.2%

Discussion

Based on our findings, we created a model of science identity that builds on previous research and adds modifications that arose from the data provided by the students in our program. Our model has three main dimensions (*recognition, science efficacy, and science interest*) that are informed by the *environmental and community context*, the student's *sociocultural identities*, and their *definition of a scientist*. Our model builds upon Carlone and Johnson's (2007) three components, but we nested performance and competence within *science efficacy*. Considering competence either as a sole construct, or as a combined element of science efficacy that incorporates competence and performance, our qualitative and quantitative analysis showed that science identity strengthened for some students during the course of our program. Previous research has highlighted how engaging in authentic scientific tasks can be important to science identity formation for high school students (Chapman & Feldman, 2017). Like others, we found support for the benefits of participation in student-led research (e.g., Jacquez et al., 2013), and additionally, exposure to citizen science or participatory science may broaden students' mental images of a scientist to one that could include themselves.

Similar to Hazari et al. (2020) we found that students expressed convoluted feelings around performance and competence. While our quantitative survey data showed some differences in the patterns for competence and performance measures from surveys, our qualitative data suggested that these two elements could not be reliably identified as separate concepts in the way students discussed them in interviews. Other researchers have also supported that performance and competence be considered as a combined construct in studies with undergraduate students in physics (Godwin et al., 2016) and computer science (Mahadeo et al., 2020; Große-bölting et al., 2023). Our study reveals further insights into students' perceptions of what it means to be confident and/or effective in science, suggesting that students may not always make a strong distinction between disciplinary knowledge versus performing scientific skills, and/or that this distinction may be difficult to differentiate, at least in some learning contexts. This finding coincides with Hazari et al.'s (2020) finding that when students feel able to perform, they also feel more competent, and when they feel competent, they also feel able to perform. The blurring of competence and performance could also be related to our study population being younger high school students who were less experienced in science than the undergraduate science students in the Carlone and Johnson (2007), and we also depended solely on self-perceived measurements of performance and competence rather than collecting any learning assessment data as previous researchers working in formal learning contexts have done. We suggest that measuring both performance and competence may remain insightful with quantitative data collected through surveys or graded assessments, even if students themselves conflate the two variables, but that qualitative approaches may find it more justified (and less confusing to participants) to discuss science efficacy as encompassing aspects of both performance and competence. We hope future research will continue to explore when and how it is useful to make distinctions between performance and competence.

We posit that high school students' definition of a scientist is a vital consideration in understanding how they develop their science identity. When a student's narrow definition of a scientist prevents them from connecting the act of performing scientific activities to being a scientist, this can limit their ability to see themselves as a scientist. Similarly, a limited concept of a scientist can inhibit the benefits of interacting with role model scientists, because the students may not readily be able to recognize these people as scientists. Previous research has examined how students' mental image of a scientist is impacted by gender and racial stereotypes, as well as stereotypes about what scientists wear and what type of science they do (Finson, 2002; Laubach et al., 2012; Miller, David et al., 2018). In adults, social group membership has been shown to influence who is and is not perceived as a scientist, as well as the scientific field of the person in question, especially if they have a PhD in a "hard" science like biology or chemistry (Suldivsky et al., 2019).

Previous scholars have noted that science identity development is strongly influenced by students' sociocultural identities and social context (Carlone & Johnson, 2007; Hazari et al., 2013; Kim & Sinatra, 2018; Lane, 2016; Rodriguez et al., 2017), and that one's science identity interacts with and is affected by racial, ethnic, gender, and other identities (Banks et al., 2016; Carlone & Johnson, 2007). For example, other researchers have found that the rifts between 'science person', 'a person who likes science', and 'a scientist', may be deeper for Black girls than for students of other identities, who may feel it is easier to move from enjoying science to thinking of themselves as scientists (Wade-Jaimes et al., 2021). Our study also provided evidence for the impact of sociocultural identities and contexts on science identity. In fact, the gender identities of researchers and educators in our program seemed highly relevant to the way students perceived these individuals as scientists (or not), which coincides with previous work on the role of the science teacher's own identity (Avraamidou, 2016).

Under the recognition component, our data supports the need to differentiate between recognition of self and recognition by others that has been discussed by previous scholars (Carlone & Johnson, 2007; Hazari et al., 2010; Hughes et al., 2021; Roberts & Hughes, 2022). The component we have defined as *recognition of self* is tied to previous research on self-efficacy, because self-efficacy beliefs incorporate recognizing one's own capabilities. Marginalized students, in particular, may struggle with building self-efficacy and recognizing themselves as a scientist, especially if their primary idea of self is as a student rather than as a scientist (Avraamidou, 2020; Flowers & Banda, 2016). Research shows the importance of self-efficacy for science identity, and especially the negative impact that navigating STEM spaces as a marginalized person has on students' belief in their own abilities regardless of their actual academic accomplishments (MacPhee et al., 2013; Rittmayer & Beier, 2009). Previous scholars have noted the importance of thinking and acting like a scientist for students' recognition of themselves as science people (Rodriguez et al., 2019). While students can develop recognition of themselves through noticing when they are acting like a scientist, we showed that if they are unaware that their actions are scientist-like then this could hamper their development of self-recognition.

We also have added support for the idea that recognition by others in the student's life seemed to be associated with stronger science identity. In our study, every student who expressed a strong sense of themselves as a science person also expressed that others recognized them as a science person. Our findings coincide with previous research such as Dou & Cian (2022), who found evidence for the importance of interactions with others in the development of science identity, especially support or lack thereof from people in the students' home environment. External recognition can be particularly important for marginalized students, who often experience a disparity in support and recognition for their science interest and ability as compared to their peers who fit adults' stereotypical notions of who can be a scientist (Stake & Nickens, 2005). Others have noted that there can also be disparities among students in the way recognition from educators occurs and whether this recognition is delivered publicly in a way that elicits further recognition by peers or remains a privately communicated recognition of the student directly by an educator (Hughes et al., 2021). Our findings, along with the previous research on the subject, suggest there is still much to explore in the way the recognition by others combines with recognition of self to impact science identity.

Considering recognition further, we added to our model a novel component highlighting the importance of recognition of others as scientists. Previous research has highlighted the importance of identity matching in potential role models (Andersen et al., 2014), and we expanded on this idea by noting the initial recognition of others as scientists as an important precursor to being able to see oneself in those potential role models. Previous research also supports the idea that interacting with experts that share demographic characteristics can build self-concept in STEM (Stout et al., 2011), but we note that this process breaks down if the student does not recognize the potential role model as a STEM expert. For example, stereotypes about scientists and stereotypes about gender can combine to create a sense of role incongruity when presented with the idea of a successful woman scientist or can even contribute to an implicit belief that women are deficient in the qualities needed to be effective

scientists (Carli et al., 2016). Thus, expanding students' understanding of the traits that a scientist might possess and the types of tasks they might do, could allow students to imagine different types of people to fit into that role. Our finding about the role of recognition of others also underlines that providing students with exposure to potential role models that are meant to provide counterstereotypical examples of scientists from underrepresented groups needs to emphasize explicitly that these role models are successful scientists, and that their activities and traits exemplify what it means to be a scientist.

Additionally, we found evidence from our qualitative analysis to include *science interest* as a component of our model, in agreement with Hazari et al. (2010). Students in our program expressed science interest when describing disciplines or topics they enjoyed learning about, discussed means of science learning, and shared other things that they like related to science. Our interview protocol directly asked students, “*what type of career do you see yourself pursuing in the future?*”, and we saw several students’ express science interest in the context of their future plans and careers. Those ranged from a general interest in pursuing a scientific career, or interest in a specific profession related to science, to an acknowledgment that even non-science careers will benefit from an exposure to and understanding of science. Hidi & Renninger (2006)’s Four Phase Model of Interest Development considers that the highest level of interest (well-developed individual interest) is when a person chooses to continue to pursue a topic when given a choice. We argue that an interest in a science career or using science in their future careers shows that our students plan on reengaging with scientific topics even after they finish high school.

As we noted above in the findings, our analysis of the ways student’s defined scientists revealed that students perceived “scientist” and “science person” as two distinct identities. This perception persisted even after the program, despite our encouragement for students to think of themselves as scientists during activities. The students’ distinction between these two terms further highlights the importance of the student’s definition of scientist in their conception of science of identity. These results are similar to Archer’s (2010) findings with middle school students who distinguished between doing and enjoying science and being or seeking to become a professional scientist. During our program we used lesson material centered around participatory science projects and attempted to communicate that anyone who does science is a scientist. Our curriculum introduced the concept of participatory science, exposed students to the idea of citizen or community scientists, and offered the opportunity to become community scientists. We posit that identification first with an identity as a community scientist, or even a student scientist, may help students bridge the gap between their notions of science person and scientist, and expand their definition of scientist to encompass their own identity.

We also included in our model a dimension related to environment/community. Student science identity development is intricately linked to the context, topical content, and pedagogical choices of the learning environment. In crafting our curriculum, we sought to lead students to relate their scientific knowledge and skills to their own environment. Others have found that place-based education and authentic research are powerful tools in STEM education (Habig & Gupta, 2021; Sobel, 2004), and our results support a role for place-based education in science identity formation. We found that our curricular focus on place-based exposure science was especially helpful in inspiring students to connect their learning to the health of themselves and their families on a personal meaningful level. Further, our student-led participatory science approach coincides with Lave & Wenger (1991) concepts of situated learning and legitimate peripheral participation in which students engage in the same practices as experts but to a more limited degree. Science identity development does not happen in a vacuum, and we believe authentically incorporating real-life environmental context into our curriculum helped students absorb science identity into their understanding of themselves.

Conclusion

In returning to our initial research questions, our findings led us to several important modifications to Carlone and Johnson's science identity model (Figure 1). Our quantitative results showed an increase in students' perceived science efficacy through participation in our community-focused environmental participatory science program, and our qualitative analysis revealed the nuanced impacts of our program on other aspects of science identity as well. We also saw that participation in our program encouraged a sense of being a science person for some students, even though our efforts were somewhat curtailed by constrictive concepts of what it means to be a scientist. Lastly, we found that individual factors such as the amount of recognition by others a student has received correlates with the strength of that student's science identity.

Challenges and Future Directions

While our focused recruitment with specific schools and partners ensured that we reached a diverse student population (Table 2), our limited sample size of students with full data and IRB consent prevented us from drawing strong conclusions on the impact of individual students' sociocultural identities on their science identity. Interviews revealed some student experiences that were informed by the student's race and gender, and we hope future studies can explore these connections.

One of our key findings, that a student's definition of a scientist informed their understanding of science identity, calls for further investigation. We did not plan to focus on measuring students' definition of scientist in designing our qualitative or quantitative instruments, and we recommend that future research directly measures students' definitions of scientists and intentionally design programs that broaden students' understanding of who is (or can be) a scientist. Evaluating shifts in students' perceptions of scientists before and after program participation will provide valuable insights into the impact of a student's scientist definition on their science identity and can enhance science education program development. Future research could also continue to explore the impact of the distinction between a science person and a scientist, and how this gap may be limiting a student's ability to see themselves as a scientist.

Implications

We believe our results have implications for science educators interested in shaping a diverse scientific community and future STEM workforce. Educators can encourage science identity development by explicitly recognizing students' achievements, knowledge, and interest in science. Further, it is vital to uncover and challenge students' definitions of what it means to be a scientist. Actively redefining the term scientist and undertaking activities that broaden perspectives regarding who is a scientist, may allow students to more easily recognize potential role models as scientists, and eventually to see themselves as scientists. Our findings suggest that educators should emphasize how engaging in scientific skills and performing scientific activities are the crux of what it means to be a scientist. We also advocate for providing students with more opportunities to engage in participatory science activities with authentic research experiences that are connected to their own communities. We have provided our full curriculum and supplemental curriculum files for free use at the SIUE STEM Center website.

Acknowledgements

We thank all student participants, our teacher partners, and their school districts, especially Madison Community School District 21st Century Program (Yolanda Crochrell and Reneeka Gause-Stallings); Southern Illinois University Edwardsville East St. Louis Charter High School (Johnathan Tate); Collinsville High School Math + Science Upward Bound Program (Desiree Tyus and Melissa Bidinger), Marvin Allen, Richardo Wells, and Eleanor Schumacher. Additionally, we thank the many participatory science programs that our curriculum content draws from, especially The Noise Project, iNaturalist, and the PurpleAir Network. Several graduate students contributed to the implementation of the program in addition to those already listed as authors, including Samreena Acharya, Lisa Drennen, Christine Favilla, Jannatul Ferdous, Dwayne Ferguson, Josh Gifford, Alexander Kalna, Samantha Murphy, Cristal Campocasso, Danielle Fischer, Margaret Lingane, and Elizabeth Wolk. The National Science Foundation provided support for this research through the Innovative Technology Experiences for Students and Teachers (ITEST) program under award DRL #1850343. Any opinions, findings, conclusions, or recommendations expressed are those of the authors and do not necessarily reflect those of the National Science Foundation.

Declaration of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Charlie Blake (cblake@siue.edu) is an interdisciplinary scholar and earned a PhD from Texas State University focused on behavioral ecology. They received their BA from Earlham College in biology and studio art. Currently, Dr. Blake is a STEM education researcher at the Southern Illinois University STEM Center with a keen interest in bringing justice and equity to science education.

Andreia Dexheimer (afdexheimer@gmail.com) has a doctorate in Biology from the University of Missouri- St. Louis, where she studied animal learning, more specifically working with tropical bees in Florida and in collaboration with The Saint Louis Zoo. She also worked as a postdoctoral fellow and research faculty at the Center for STEM Research, Education, and Outreach at Southern Illinois University Edwardsville, leading and coordinating STEM Education research projects across the St. Louis Metro region.

Carol E. Colaninno (carolc@wustl.edu) has been actively involved in K-12 STEM curriculum development, education, and research for the past decade. She has a passion for translating science research, content, and practices into K-12 learning opportunities that support students to develop science interest, science identity, and a sense of belonging in STEM. She also studies contextual factors in field research settings that promote safe and inclusive science.

Candice Johnson (candi.johnson@slu.edu) is the former Program Coordinator of the Southern Illinois University Edwardsville STEM Center and is now Program Director of Research and Development at Saint Louis University. Candice's work focuses on advancing educational equity and she enjoys working to provide resources that support improved learning outcomes for all students.

Ben K. Greenfield (benjamin.greenfield@maine.edu) is an Associate Professor and the Department Chair in the Public Health Program of the Muskie School of Public Service at the University of Southern Maine. His research focuses on environmental exposure and risk assessment and community environmental justice. Ben's recent projects have examined mercury contamination in

fish, and antimicrobial drug resistance in aquatic environments, and integration of environmental health into STEM education.

Lillian Ward (lilliancw95@gmail.com) graduated with a Bachelor of Science degree in Geography from Aquinas College in 2018. Since 2019, she has worked as a fisheries field technician for the Illinois Natural History Survey Great Rives Field Station. In 2023, she completed her Master of Science degree in Environmental Biology at Southern Illinois University Edwardsville. When she's not working on the river, she trains and performs with a professional contemporary dance company, crochets, bakes bread and spends time with her four beloved chickens.

Ozaveshe Paul Amune (aozaveshe@gmail.com) is a former graduate research assistant at the SIUE STEM Center where he specialized in qualitative research. Currently he is a Field and Application Sales Engineer at Schneider Electric, specializing in delivering tailored energy solutions across Louisiana and Mississippi and helping to optimize energy operations through technical innovations and automation.

Adriana Martinez (admart@siue.edu) is a Full Professor at Southern Illinois University Edwardsville in the Department of Geography & GIS and the Department of Environmental Sciences. She studies the impact of border militarization on the Rio Grande through a mix of field and GIS methods. She also focuses her research on creating inclusive spaces in geoscience for undergraduate and graduate students.

Sharon Locke (slocke@siue.edu) earned her PhD in Geosciences from the University of Minnesota and has served as Director of the Southern Illinois University Edwardsville STEM Center since 2010. Dr. Locke is a former cluster coordinator in the Division of Research on Learning in Formal and Informal Settings at the National Science Foundation and has worked on STEM education projects around the world. Dr. Locke's research focuses on effective STEM education and career pathways.

Georgia Bracey (gbracey@siue.edu) recently retired from her position as a research assistant professor at the Southern Illinois University Edwardsville STEM Center. She earned a BA in elementary education and spent thirteen years teaching in public schools. She also has an MS in physics and a PhD in science education. She has published research on physics education, citizen science, and the integration of technology in teaching and learning.

References

- Abel, T. D. (2008). Skewed riskscape and environmental injustice: A case study of metropolitan St. Louis. *Environmental Management*, 42, 232–248. <https://doi.org/10.1007/s00267-008-9126-2>
- Andersen, H., Krogh, L., & Lykkegaard, E. (2014). Identity Matching to Scientists: Differences that Make a Difference? *Research in Science Education*, 44(3), 439–460. <https://doi.org/10.1007/s11165-013-9391-9>
- Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). “Doing” science versus “being” a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education*, 94(4), 617–639. <https://doi.org/10.1002/sce.20399>
- Avraamidou, L. (2016). *Studying Science Teacher Identity: Theoretical, Methodological and Empirical Explorations*. Brill. <https://doi.org/10.1007/978-94-6300-528-9>

- Avraamidou, L. (2020). Science identity as a landscape of becoming: Rethinking recognition and emotions through an intersectionality lens. *Cultural Studies of Science Education*, 15(2), 323–345. <https://doi.org/10.1007/s11422-019-09954-7>
- Barr, D. B. (2006). Expanding the role of exposure science in environmental health. *J Expo Sci Environ Epidemiol*, 16(2), 473. <https://doi.org/10.1289/ehp.0901634>
- Banks, J., Cochran-Smith, M., Moll, L., Richert, A., Zeichner, K., LePage, P., Darling-Hammond, L., Duffy, H., & McDonald, M. (2016). Teaching diverse learners. In J. Bransford & L. Darling-Hammond (Eds.), *Handbook of research on teaching* (pp. 593–638). American Educational Research Association Washington, DC.
- Bell, P., Lewenstein, B., Shouse, A. W., & Feder, M. A. (Eds.). (2009). *Learning Science in Informal Environments: People, Places, and Pursuits*. National Research Council. <https://doi.org/10.17226/12190>
- Bonney, R., Phillips, T., Ballard, H., & Enck, J. (2015). Can citizen science enhance public understanding of science? *Public Understanding of Science* (Bristol, England), 25. <https://doi.org/10.1177/0963662515607406>
- Brinkmann, S., & Kvale, S. (2015). *InterViews: Learning the craft of qualitative research interviewing* (3rd ed). Sage Publications Thousand Oaks.
- Carli, L. L., Alawa, L., Lee, Y., Zhao, B., & Kim, E. (2016). Stereotypes About Gender and Science: Women $\neq$ Scientists. *Psychology of Women Quarterly*, 40(2), 244–260. <https://doi.org/10.1177/0361684315622645>
- Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching*, 44(8), 1187–1218. <https://doi.org/10.1002/tea.20237>
- Chapman, A., & Feldman, A. (2017). Cultivation of science identity through authentic science in an urban high school classroom. *Cultural Studies of Science Education*, 12(2), 469–491. <https://doi.org/10.1007/s11422-015-9723-3>
- Citizen Science Association. (2023, July 14). Announcing a new name for this Association Return to Blog. <https://citizenscience.org/2023/07/14/announcing-a-new-name-for-this-association/>
- Colten, C. E., & Samsel, T. B. (1988). Historical Assessment of Hazardous Waste Management in Madison and St. Clair Counties, Illinois, 1890-1980. (Hazardous Waste Research and Information Center). Illinois State Water Survey Division. <https://hdl.handle.net/2142/1910>
- Creswell, J. H., Plano Clark, V. L., Gutmann, M. L., & Hanson, W. E. (2003). Advanced mixed methods research designs. In A. Tashakkori & C. Teddlie (Eds.), *Handbook of Mixed Methods in Social & Behavioral Research* (pp. 209-240). Sage Publications.
- Donovan, E. (2016). Learning to embrace our stories: Using place-based education practices to inspire authentic writing. *Middle School Journal*, 47(4), 23–31. <https://doi.org/10.1080/00940771.2016.1202657>
- Dou, R., & Cian, H. (2022). Constructing STEM identity: An expanded structural model for STEM identity research. *Journal of Research in Science Teaching*, 59(3), 458–490. <https://doi.org/10.1002/tea.21734>
- Estrada, M., Woodcock, A., Hernandez, P. R., & Schultz, P. W. (2011). Toward a Model of Social Influence that Explains Minority Student Integration into the Scientific Community. *Journal of Educational Psychology*, 103(1), 206–222. <https://doi.org/10.1037/a0020743>
- Finson, K. D. (2002). Drawing a Scientist: What We Do and Do Not Know After Fifty Years of Drawings. *School Science and Mathematics*, 102(7), 335–345. <https://doi.org/10.1111/j.1949-8594.2002.tb18217.x>
- Flowers, A. M., III, & Banda, R. (2016). Cultivating science identity through sources of self-efficacy. *Journal for Multicultural Education*, 10(3), 405–417. ProQuest One Business. <https://doi.org/10.1108/JME-01-2016-0014>

- Gee, J. P. (2014). *An introduction to discourse analysis: Theory and method*. routledge.
- Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2016). Identity, Critical Agency, and Engineering: An Affective Model for Predicting Engineering as a Career Choice. *Journal of Engineering Education*, 105(2), 312–340. <https://doi.org/10.1002/jee.20118>
- Große-bölting, G., Gerstenberger, D., Gildehaus, L., Mühling, A., & Schulte, C. (2023). Identity in Higher Computer Education Research: A Systematic Literature Review. *ACM Trans. Comput. Educ.*, 23(3), 35:1-35:35. <https://doi.org/10.1145/3606707>
- Habig, B., & Gupta, P. (2021). Authentic STEM research, practices of science, and interest development in an informal science education program. *International Journal of STEM Education*, 8(1), 1–18. <https://doi.org/10.1186/s40594-021-00314-y>
- Hazari, Z., Chari, D., Potvin, G., & Brewé, E. (2020). The context dependence of physics identity: Examining the role of performance/competence, recognition, interest, and sense of belonging for lower and upper female physics undergraduates. *Journal of Research in Science Teaching*, 57(10), 1583–1607. <https://doi.org/10.1002/tea.21644>
- Hazari, Z., Sadler, P. M., & Sonnert, G. (2013). The Science Identity of College Students: Exploring the Intersection of Gender, Race, and Ethnicity. *Journal of College Science Teaching*, 42(5), 82–91. https://doi.org/10.2505/4/jcst13_042_05_82
- Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching*, 47, 978–1003. <https://doi.org/10.1002/tea.20363>
- Hidi, S., & Renninger, K. A. (2006). The Four-Phase Model of Interest Development. *Educational Psychologist*, 41(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4
- Honey, M., Pearson, G., & Schweingruber, H. (Eds.). (2014). *STEM Integration in K-12 Education: Status, Prospects and an Agenda for Research*. National Research Council. <https://drive.google.com/drive/folders/1EZ0VvjFjBdd40nqAbOLDxelbWeRDL9i->
- Hughes, R., Schellinger, J., & Roberts, K. (2021). The role of recognition in disciplinary identity for girls. *Journal of Research in Science Teaching*, 58(3), 420–455. <https://doi.org/10.1002/tea.21665>
- Illinois Report Card. (2023). Illinois State Board of Education. <https://www.illinoisreportcard.com>
- iNaturalist. (2025). iNaturalist. <https://www.inaturalist.org/>
- Jacquez, F., Vaughn, L. M., & Wagner, E. (2013). Youth as Partners, Participants or Passive Recipients: A Review of Children and Adolescents in Community-Based Participatory Research (CBPR). *American Journal of Community Psychology*, 51(1), 176–189. <https://doi.org/10.1007/s10464-012-9533-7>
- Kaminski, M. D., & Landsberger, S. (2000). Heavy Metals in Urban Soils of East St. Louis, IL, Part I: Total Concentration of Heavy Metals in Soils. *Journal of the Air & Waste Management Association*, 50(9), 1667–1679. <https://doi.org/10.1080/10473289.2000.10464195>
- Kim, A. Y., & Sinatra, G. M. (2018). Science identity development: An interactionist approach. *International Journal of STEM Education*, 5(1), 51. <https://doi.org/10.1186/s40594-018-0149-9>
- Kosmala, M., Wiggins, A., Swanson, A., & Simmons, B. (2016). Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10), 551–560. <https://doi.org/10.1002/fee.1436>
- Larson, L. R., Cooper, C. B., Futch, S., Singh, D., Shipley, N. J., Dale, K., LeBaron, G. S., & Takekawa, J. Y. (2020). The diverse motivations of citizen scientists: Does conservation emphasis grow as volunteer participation progresses? *Biological Conservation*, 242. <https://doi.org/10.1016/j.biocon.2020.108428>
- Lane, T. B. (2016). Research environments as counterspaces? Examining spaces that inhibit and support science identity development for black students in STEM. *Urban Education Research & Policy Annuals*, 4(1).

- Laubach, T. A., Crofford, G. D., & Marek, E. A. (2012). Exploring Native American Students' Perceptions of Scientists. *International Journal of Science Education*, 34(11), 1769–1794. <https://doi.org/10.1080/09500693.2012.689434>
- Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.
- Lioy, P. J. (2010). Exposure science: a view of the past and milestones for the future. *Environmental health perspectives*, 118(8), 1081-1090. <https://doi.org/10.1289/ehp.0901634>
- Mahadeo, J., Hazari, Z., & Potvin, G. (2020). Developing a Computing Identity Framework: Understanding Computer Science and Information Technology Career Choice. *ACM Trans. Comput. Educ.*, 20(1), 7:1-7:14. <https://doi.org/10.1145/3365571>
- MacPhee, D., Farro, S., & Canetto, S. S. (2013). Academic Self-Efficacy and Performance of Underrepresented STEM Majors: Gender, Ethnic, and Social Class Patterns. *Analyses of Social Issues and Public Policy*, 13(1), 347–369. <https://doi.org/10.1111/asap.12033>
- McDonald, M. M., Zeigler-Hill, V., Vrabel, J. K., & Escobar, M. (2019). A Single-Item Measure for Assessing STEM Identity. *Frontiers in Education*, 4, 78. <https://doi.org/10.3389/feduc.2019.00078>
- McKinley, D. C., Miller-Rushing, A. J., Ballard, H. L., Bonney, R., Brown, H., Cook-Patton, S. C., Evans, D. M., French, R. A., Parrish, J. K., Phillips, T. B., Ryan, S. F., Shanley, L. A., Shirk, J. L., Stepenuck, K. F., Weltzin, J. F., Wiggins, A., Boyle, O. D., Briggs, R. D., Chapin, S. F., ... Soukup, M. A. (2017). Citizen science can improve conservation science, natural resource management, and environmental protection. *The Role of Citizen Science in Biological Conservation*, 208, 15–28. <https://doi.org/10.1016/j.biocon.2016.05.015>
- Miller, David, Nolla, Kyle, Eagly, Alice, & Uttal, David H. (2018). The Development of Children's Gender-Science Stereotypes: A Meta-analysis of 5 Decades of U.S. Draw-A-Scientist Studies. *Child Development*, 89(6), 1943–1955. <https://doi.org/10.1111/cdev.13039>
- National Research Council. (2009). *Learning science in informal environments: People, places, and pursuits*. National Academies Press.
- Next Generation Science Standards: For States, By States. (2013). National Academies Press. <https://doi.org/10.17226/18290>
- Noise Project. (2025). Noise Project. Retrieved August 11, 2025, from <https://noiseproject.org/>
- Nunes, B., & Theising, A. J. (2011). East St. Louis. Arcadia Publishing.
- R Core Team. (2023). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org/>
- Real-time Air Quality Monitoring by PurpleAir. (2025). PurpleAir. Retrieved August 11, 2025, from <https://www2.purpleair.com/>
- Reeves, T. C., Herrington, J., & Oliver, R. (2005). Design research: A socially responsible approach to instructional technology research in higher education. *Journal of Computing in Higher Education*, 16(2), 96–115. <https://doi.org/10.1007/BF02961476>
- Rittmayer, A. D., & Beier, M. E. (2009). Overview: Self-efficacy in STEM. In B. Bogue & E. Cady (Eds.), *Applying Research to Practice (ARP) Resources*.
- Roberts, K., & Hughes, R. (2022). Recognition Matters: The Role of Informal Science Education Programs in Developing Girls' Science Identity. *Journal for STEM Education Research*, 5(2), 214–232. <https://doi.org/10.1007/s41979-022-00069-3>
- Rodriguez, S. L., Cunningham, K., & Jordan, A. (2017). What a Scientist Looks Like: How Community Colleges Can Utilize and Enhance Science Identity Development as a Means to Improve Success for Women of Color. *Community College Journal of Research and Practice*, 41(4–5), 232–238. <https://doi.org/10.1080/10668926.2016.1251354>

- Rodriguez, S. L., Cunningham, K., & Jordan, A. (2019). STEM Identity Development for Latinas: The Role of Self- and Outside Recognition. *Journal of Hispanic Higher Education*, 18(3), 254–272. <https://doi.org/10.1177/1538192717739958>
- Saldana, J. (2021). *The Coding Manual for Qualitative Researchers*. SAGE Publications Ltd.
- Smith, G., & Sobel, D. (2010). Bring It On Home. *Educational Leadership: Journal of the Department of Supervision and Curriculum Development*, N.E.A, 68, 38–43.
- Sobel, D. (2004). Place-based education: Connecting classrooms & communities (2nd ed.). Orion Society.
- Stake, J. E., & Nickens, S. D. (2005). Adolescent Girls' and Boys' Science Peer Relationships and Perceptions of the Possible Self as Scientist. *Sex Roles: A Journal of Research*, 52(1), 1–11. <https://doi.org/10.1007/s11199-005-1189-4>
- Stout, J. G., Dasgupta, N., Hunsinger, M., & McManus, M. A. (2011). STEMing the tide: Using ingroup experts to inoculate women's self-concept in science, technology, engineering, and mathematics (STEM). *Journal of Personality and Social Psychology*, 100(2), 255–270. <https://doi.org/10.1037/a0021385>
- Suldozsky, B., Landrum, A., & Stroud, N. J. (2019). Public perceptions of who counts as a scientist for controversial science. *Public Understanding of Science*, 28(7), 797–811. <https://doi.org/10.1177/0963662519856768>
- Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A.A., Dietterich, T., Farnsworth, A., Fink, D., Fitzpatrick, J. W., Fredericks, T., Gerbracht, J., Gomes, C., Hochachka, W. M., Iliff, M. J., Lagoze, C., La Sorte, F. A., ... Kelling, S. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, 169, 31–40. <https://doi.org/10.1016/j.biocon.2013.11.003>
- Trujillo, G., & Tanner, K. D. (2014). Considering the Role of Affect in Learning: Monitoring Students' Self-Efficacy, Sense of Belonging, and Science Identity. *CBE—Life Sciences Education*, 13(1), 6–15. <https://doi.org/10.1187/cbe.13-12-0241>
- Tyler-Wood, T., Knezek, G., & Christensen, R. (2010). Instruments for assessing interest in STEM content and careers. *Journal of Technology and Teacher Education*, 18(2), 341–363.
- U.S. Census Bureau. (2022). 2018-2022 American Community Survey. census.gov
- Vermillion, B., Brugam, R., Retzlaff, W., & Bala, I. (2005). The sedimentary record of environmental lead contamination at St. Louis, Missouri (USA) area smelters. *Journal of Paleolimnology*, 33(2), 189–203. <https://doi.org/10.1007/s10933-004-4078-x>
- Wade-Jaimes, K., King, N., & Schwartz, R. (2021). “You could like science and not be a science person”: Black girls' negotiation of space and identity in science. *Science Education*, 105(5), 855–879. <https://doi.org/10.1002/sce.21664>

Supplementary

Student Interview Protocol

Hello! My name is _____. I want to learn a little more about you. Are you okay talking now?
[Interviewer note: If you have not met or worked with the student before, please make sure to introduce yourself and your role on the program. Try to make this informal, so the student feels more comfortable speaking to you. If the student has not participated in an interview before, review the process with them. For example, tell them there are 12 questions, with some having multiple parts and that the interview will take about 15 to 20 minutes. Share that there are no “right” or “wrong” answers. Ask if they have any questions before getting started and if the student is okay with being recorded].

Are you okay with me recording our conversation?

1. What did you like about being in the program? What was the most memorable part of the program?
2. Thinking back on the summer program, can you tell me about working with the scientists? What kinds of things did you do? What kinds of things did the scientist do? *If “no” ask if they have met scientists elsewhere. Tell me about it. What kinds of things did you do? What kinds of things did the scientist do?*
3. Do you feel like a **science person**? Why/why not? Do you feel like you are a scientist? Why/why not? What makes a person a scientist or a science person? *Probe: What do scientists look like or do? Probe: What is the difference between a science person and a scientist (if any)?*
4. How would you describe your feelings about science? *Probe: Interests; Value; Relevance*
5. Do you think other people [friends, family, teachers] think you're a **science person**? Do you think they think you are a scientist? Which other people? How do you know? *Probe: What do scientists look like or do? Probe: What is the difference (if any)? Would you like to be seen as a scientist/science person? How would this make you feel?*
6. Can you tell me some things you learned about the environment (both indoor and outdoor)? Can you tell me anything else that you know about the environment? What else would you like to know? *How about pollution and your environment?*
7. On a scale from 1-5, with 5 being the highest, how confident are you in using technology to communicate information about science (blogs, social media, websites)? What about using technology to do science (collect information, measure things)? Why did you pick # and not #-1, why not a #+1.
8. Can you think of some ways you could study the environment to help the people and places where you live, like your neighborhood, school, street, you live on, etc.? *Is studying your environment something you would feel confident doing?*
9. What are some things you learned about the environment where you live from the program sessions? What else would you like to know? What do you think other people who live near you would like to know? *Probe: Is there anything about the air, noise, soil, the human-built environment, pollution, etc. of your community?*

10. What type of career do you see yourself pursuing in the future? *Probe: Intentions of pursuing STEM or environmental science-related careers; Awareness of STEM or environmental science careers.*
11. What else would you have wanted to do in the program? *Probe: Is there anything else you wanted to do or learn about?*
12. Is there anything else you would like to tell me about the program, and what we have talked about today?

Codebook

Community Views: How the student sees how other people in their community views community

- Family: Influences from what their family says
- Viewpoints: their understanding of the community or understanding of how the community feels

Competence [knowledge and understanding of science content (may be less visible than performance)]

- Scientific competence: student's perception or expression of their science knowledge.
- Social media and online communication: student's expression of their knowledge and understanding of social media and other technology to communicate about science.
- Technology competence: student's perception or expression of their knowledge of technology
- Tools competence: student's perception or expression of their knowledge of scientific tools and instruments.

Environmental perspectives: How the student views and understands the Environment (big E environment here)

- Content knowledge - environment: their understanding of the environment
- Need to change: Students express the need to change the current environment, solution for changes, or the need to understand past environmental conditions
- Pollution: when the student talks about pollution in any sense.

Future careers: How the student views their future in education, jobs, and careers.

- Non-STEM: Career and/or future in a non-STEM field
- STEM (STEM to include medical jobs): Careers and/or future in an STEM field

Non-STEM interest: Student's description of their non-STEM interest, means of non-STEM learning, or other things that they like/like to do, excluding science.

- Learning, knowledge, and perceptions: Student's expression of the way they learn and how they perceive their abilities/attitudes/emotions in non-STEM areas.
- Non-STEM interest: Non-STEM interests the student has.

Performance: How the student views their ability to apply scientific knowledge and perform scientific tasks. (social performances of relevant scientific practices - e.g., ways of talking and using tools)

- Science tools/instruments: student's perception of their ability to use scientific tools and instruments.
- Scientific tasks: students' perception of their ability to perform scientific tasks or apply science knowledge exclusive of science tools or technology.
- Social media and online communication: Student's use of online tools like Zoom, Google meets, YouTube, TikToks and other platforms to communicate science.
- Use of technology for science: Student's view of using technology to perform scientific tasks.

Recognition (external and internal): How the student perceives others' view of them and their view of themselves (recognizing oneself and being recognized by others as a "science person")

- Attitude: their thoughts/feelings/emotions/attitude on what it would feel like to be recognized as a scientist (can be positive or negative)
- External recognition: the student expresses being recognized by others as a scientist/science person

- Internal recognition: the student internally acknowledges that they do or do not feel as though they are a science person.

Science identity: View of self as scientist and interest in science. When the student talks about themselves as a scientist or their interest in science.

- Expression of science identity: Attributes that indicate a student's degree of identity as a scientist.

Science Interest: Student's description of their science interest, means of science learning, or other things that they like related to science.

- Expression of science interest: Attributes that indicate a student's degree of interest in science.
- Learning, knowledge, and perceptions: Student's expression of the way they learn and how they perceive their abilities/attitudes/emotions regarding science/STEM.
- Student science interest: Science interests the student has.

View of scientists and/or science: How the student sees and understands scientists, and any past interactions or knowing of scientists.

- Appearance: student's view of what scientists look like.
- Interactions: student's description of the interactions with scientists.
- Scientists actions: student's understanding of what scientists do.
- View of science: Expression of how the student places views or value (negative, neutral, or positive) science and scientists

Pre-program and post-program long survey

Name: _____

Date: _____

Please answer the following questions by writing an X or check mark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

<i>I am able to...</i>	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
Create a science project to answer a question about the environment.					
Use data to explain things about my environment.					
Record measurements using different kinds of monitors and tools.					
Collect measurements that help answer questions about my environment.					
Understand how the data I collect contribute to science.					
Share my research findings with my friends.					
Share my research findings with people in the community.					

Please answer the following questions by writing an X or check mark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

<i>In the past, I have worked with others to...</i>	Never	Rarely	Sometimes	Often	Always
Create a science project to answer a question about the environment.					
Use data to explain things about my environment.					
Record measurements using different kinds of monitors and tools.					
Collect measurements that help answer questions about my environment.					
Understand how the data I collect contributes to science.					
Share my research findings with my friends.					
Share my research findings with people in the community.					

Please answer the following questions by writing an X or check mark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

<i>Other people, like my teachers, parents, and friends, notice that I can ...</i>	Never	Rarely	Sometimes	Often	Always
Create a science project to answer a question about the environment.					
Use data to explain things about my environment.					
Record measurements using different kinds of monitors and tools.					
Collect measurements that help answer questions about my environment.					
Understand how the data I collect contributes to science.					
Share my research findings with my friends.					
Share my research findings with people in the community.					

Please answer the following questions by writing an X or check mark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I am a science person.					
I am a scientist.					
Other people think I'm a science person.					
I feel like I fit in with others in the world of science.					
I can see myself doing the work of a scientist.					

Please answer the following questions by writing an X or check mark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

	Strong ly disagr ee	Disagr ee	Neither agree nor disagree	Agree	Stron gly agre e
I plan to go to college.					
I plan to study science in college.					
I plan to study about the environment in college.					
I will get a job in a science-related area.					
I will get a job in an environmental science-related area.					
I would enjoy a job in science.					
I would enjoy a job in environmental science.					

Name: _____

What grade are you starting this coming school year?: _____

Age: _____

I am (circle your choices): Female Male Non-binary Other: _____

I am (circle all of your choices): American Indian Alaska Native

Asian American African American Black Hispanic Latino/a

Native Hawaiian Other Pacific Islander White Other: _____

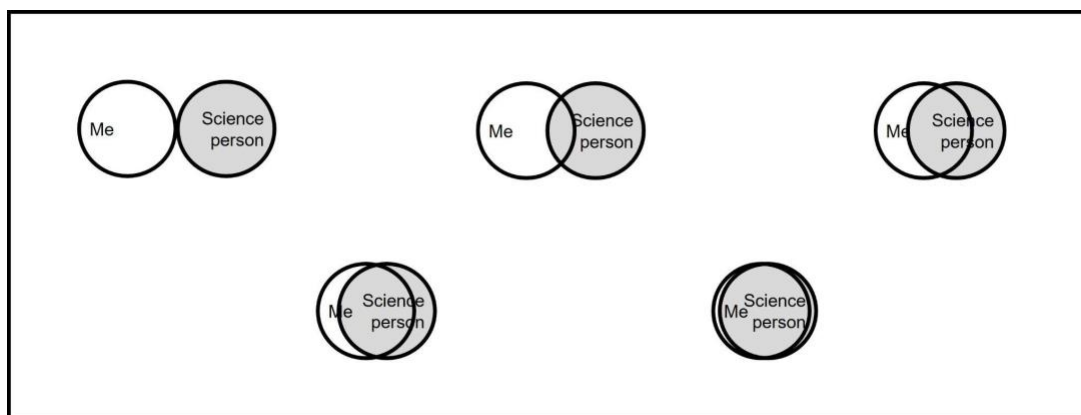
Repeated measures short survey

Name: _____ Date: _____

Please answer the following questions by writing an X or checkmark in the box. Remember, there is no right or wrong answer. We just want to understand what you think. If there is a question you do not want to answer, you do not need to. If you have any questions, please raise your hand and someone will help you out!

	Strongly disagree	Disagree	Neither agree nor disagree	Agree	Strongly agree
I see myself as a science person.					
Being a science person is important to me.					
I think of myself as a science person.					
I am a science person.					
I am a scientist.					
In general, being a science person is an important part of who I am.					
Knowing more scientists with a similar background as mine makes me feel more like a science person.					

Mark the picture below that best describes the overlap of the image you have of yourself and the image you have of a science person.



A Review of Teaching and Learning Sequences for the Scientific Practice of Modeling

Aaron J. Sickel
University of Hawai'i at Mānoa

ABSTRACT

Modeling is widely recognized as a core scientific practice and an essential component of science learning. Although numerous teaching and learning sequences (TLSs) have been proposed to support modeling in science education, the field lacks both a comprehensive review and a shared understanding of their design features. This review addresses that gap by systematically identifying and analyzing 18 distinct modeling-focused TLSs proposed in peer-reviewed articles and book chapters (representing a 1987-2025 timespan), with a focus on points of convergence and divergence across the TLSs. While several common components and learning goals emerged, I highlight five key areas of divergence that raise critical considerations for the future development of modeling-based TLSs. Building on these findings, I propose a set of design and research considerations for scholars seeking to advance work in this domain. I conclude with a description of three future research agendas that aim to strengthen the theoretical and practical foundations of modeling-focused TLSs in science education.

Keywords: modeling, science education, teaching and learning sequences

Introduction

Scientists use models to help explain the natural world (Gilbert & Boulter, 2000; Gilbert & Justi, 2016; Greca & Moreira, 2000; Hestenes, 1987; Schwarz et al., 2009). From the development of atomic models to climate change simulations, models help scientists develop abstracted representations which can be used to understand and predict the structure and interactions of various phenomena. Science education reform documents support the goal of producing a scientifically literate population; one that not only understands overarching scientific concepts, but also engages in the scientific enterprise (National Research Council [NRC], 1996; NGSS Lead States, 2013). Due to the importance of modeling in scientific practice, numerous science scholars argue that science instruction should place greater focus on students working with models to understand the natural world (Campbell et al., 2012; Gilbert, 1991; Grosslight et al., 1991; Greca & Moreira, 2000; Schwarz et al., 2009; Windschitl et al., 2008). Moreover, students who participate in learning experiences focused on modeling have the potential to gain several benefits, such as deeper conceptual learning (Justi et al., 2009; Wells et al., 1995) and increased ability to engage in scientific reasoning (Malone & Schuchardt, 2023).

Teaching and learning sequences (TLSs) refer to the progression of intentional phases in formalized instruction with the purpose of engaging learners in achieving particular learning outcomes. TLSs are models to guide a teacher's instruction in a productive way. Instruction is not organized according to the whims of a teacher, but as part of a coherent framework. Within the field of science education, TLSs have been designed to help grades K-16 learners engage in scientific practices and use them to construct scientific explanations in structured lessons and units. Examples include the

learning cycle approach - exploration, invention, discovery (Kaplus & Thier, 1967) - and the 5E model - engage, explore, explain, elaborate, evaluate (Bybee et al., 2006). Both of these sequences support learners with constructing science concepts while engaging in scientific practices that include asking scientific questions, data analysis, and designing scientific investigations. When teachers organize their construction according to these TLSs, there have been positive outcomes (Bybee et al., 2016). Other examples that can be applied to science education include problem-based learning (Larmer, 2015) and engineering design (Museum of Science, Boston, n.d.). In all of these TLSs, educators have given careful consideration to the nature, order and components of each phase.

Despite the advantages of and interest in TLSs, to date, the science education field lacks a review on TLSs specific to modeling. A cursory knowledge of the literature reveals that a consensus TLS for modeling does not exist. Science educators lack a comprehensive resource that brings together and makes sense of the TLSs proposed over the years. The purpose of this article is to address that gap. A review of the literature yielded 18 different TLSs focused on modeling that have been developed and articulated by science education scholars. In reviewing these TLSs, I aimed to answer the following questions:

1. What are the points of convergence and divergence across the TLSs?
2. What are considerations for science education researchers and scholars in the future based on those points of convergence and divergence?

This article is organized by four ensuring sections: (1) defining models and modeling, (2) articulating the methodology (3) presenting the findings, and (4) discussing final conclusions. These sections are addressed below.

Defining Models and Modeling

In the context of science education, models are “specialized representations that embody aspects of mechanism, causality, or function to illustrate, explain, and predict phenomena” (Schwarz et al., 2009, p. 634). Models can incorporate representations of objects (cell, airplane), concepts (force, energy), systems (digestive system, ecosystem trophic levels), events (projected object), or processes (natural selection, photosynthesis) (Gilbert et al., 2000). Their foci may vary, from representing a narrow set of relationships (an object’s speed is determined by examining the relationship between distance and time) to conceptualizing complex systems and processes (graphic organizer connecting multiple sources and components associated with the greenhouse effect). They can connect concepts of different physical scales together (drawing boiling water at the macro- and molecular level) and become generalized tools to help an individual understand a wide range of phenomena (using free-body diagrams to understand motion in two dimensions).

We rely on modeling to make sense of our experiences, constantly simulating everyday situations in our minds as we interpret events and carry out tasks (Chart, 2000). If these mental models are given attention and opportunities for formal examinations, they can be transformed into conceptual models that have predictive and interpretative power (Garnham, 1997; Ormrod, 2008). Once an individual communicates their mental model, it is said to be an *expressed model* (Boulter & Buckley, 2000). Through teaching interventions and learning opportunities, learners may be exposed to or begin constructing models that are more abstract and provide conceptual insights into phenomena and their interactions. I refer to such models as *conceptual models*. The learner’s initially *expressed model* may have some conceptual components, but often contain similar depictions of the event or object under study (Grosslight et al., 1991). Through participation in specially designed learning experiences, their models have the potential to become more abstract and conceptual over time.

Gilbert (1991) argued that scientific practice is essentially “a process of constructing predictive

conceptual models” (p. 73). The practice of developing models is referred to as *modeling*. It can be a means to help learners refine their mental models into conceptual models, to test explanations against evidence (Windschitl et al., 2008) and eventually encompass some compatibility with current scientific knowledge (Greca and Moreira, 2000).

Methodology

Below, I provide details of the identification criteria, selection process, and analysis procedure that guided this review.

Identification Criteria

The goal of my review was to find peer-reviewed articles or book chapters that propose a TLS for the scientific practice of modeling within the last 40 years (1985-2025). The following boundaries were placed at the outset:

1. **Teaching and learning boundary:** The purpose of the TLS needed to emphasize its potential use in grades K-16 learning contexts, rather than models or frameworks that only elucidate the nature of modeling. It is important to distinguish TLSs from learning progressions. A learning progression designates different levels of performance, and how educators might help learners advance to higher or more sophisticated levels (Schwarz et al., 2009). A learning progression is not a TLS because it does not communicate information about how learning experiences should be organized within a lesson or unit. Therefore, learning progressions on modeling were not included in this review.
2. **Modeling focus boundary:** The TLSs for this review needed to articulate at least two phases dedicated to some aspect of modeling, rather than just one phase for modeling within a broader framework. The purpose of this review is to understand TLSs for which modeling was a central feature.
3. **Clear description and rationale boundary:** This boundary required that articles or book chapters referencing a TLS not just share a table or figure of a TLS, but also dedicated space for explaining the rationale of the TLS and/or a clear articulation of each phase of the TLS.

Screening & Inclusion

I conducted a search across peer-reviewed books and journal articles in science education. The search proved to be challenging, however, because of the following reasons:

1. There is a vast amount of literature related to modeling and science education. An initial search in the ERIC database combining the words, *modeling*, *science*, and *education*, yielded a list of 13,303 hits.
2. There are not agreed-upon terms. People use words like *framework*, *method*, *sequence*, and *cycle* to represent a TLS, and the TLSs they propose are sometimes embedded within an empirical study (which means the titles and abstracts might be related to the findings of the study more so than the TLS).

Decisions had to be made to narrow the search for the purpose of practicality. I conducted the search for peer-reviewed products by searching for key words in titles and abstracts, with the

justification that key words would likely be in those places if the authors dedicated space to the discussion of a TLS. In the ERIC and Academic Search Complete databases, as well as an R1 university search engine, the key words of *modeling* and *model-based* were consistently used in conjunction with words likely to be used when describing a TLS (i.e. sequence, instruction, inquiry, theory, cycle, phase, construct) in a range of permutations. The greater specificity of these searches narrowed each list of hits to approximately 50-300 articles to review at a time. I scanned each title in these lists, and reviewed abstracts that had the potential to align with the selection boundaries. I also used the university library search engine to probe exclusively for books using the words, *model-based*, *modeling*, *science*, and *education*. Chapters with potential to meet the boundaries listed above were selected for further review.

41 articles and book chapters selected from the initial searches were thoroughly reviewed. Of those, 11 fully met the boundary criteria. However, with the focus of their writing on a TLS, those authors referenced other TLSs. I subsequently combed through each reference list of the original 11 and identified 7 more peer-reviewed articles or book chapters that described a TLS. A final count of 18 TLSs clearly met the boundaries, spanning the years 1987 to 2025. A full list of those TLSs can be found in the appendix. Each TLS had a unique set of phases with respect to total number, order, and description, with one slight exception. Two authors referred to the same TLS but provided some key differences in their descriptions (both Clement and Steinberg (2002) and Khan (2007) discuss a generate-evaluate-modify cycle). Therefore, both versions of the TLS were included in the review.

It is important to note that TLSs may have been missed in the process, due to great amount of scholarship related to modeling in science education (generally speaking) and the challenges of diverse phrasing mentioned above. However, every attempt was made to be thorough during the screening process. The final set of 18 TLSs reveals patterns, tensions, and insights that can inform future work in science education.

Analysis

To document information for each TLS, I read each article/chapter and took detailed notes. Using the strategy of constructing a descriptive framework (Yin, 1994), these notes were placed into two digital spreadsheets.

To gain a foundation of clarity about the TLSs and how they were structured, my first action was to determine the major components and goals of TLS phases. In the first spreadsheet, each TLS encompassed three columns. One column included a summary of each phase of the TLS, a second column noted learner actions, and a third column noted teacher actions. Three rounds of coding ensued. In the first round of coding, I aimed to stay close to the authors' descriptions, documenting the major actions that learners and teachers were expected to take. These actions were later called 'components.' Following recommendations from Saldaña (2009), in the second round I honed the preliminary component codes into final component codes, clarifying and adjusting the meaning of each code. In the third round, I collapsed component codes into three overarching categories (Saldaña, 2009), later termed 'goals.' Figure 1 highlights an example of coding for one TLS. I display the TLS phases in order from left to right. The component codes are placed underneath the phases in the 2nd row, and then the corresponding, more abstracted goal codes are in the third row. Figure 2 provides a broad snapshot of 17 TLSs, coded at the level of the three overarching goals. This figure helps the reader see the total number of phases for each TLS, and how the three goals progress through the sequence. One TLS was not included in Figure 2 because it is completely non-linear in nature (Gilbert & Justi, 2016). This TLS is discussed in the *points of divergence* section of the article.

In the second spreadsheet, I began documenting commonalities and differences across TLSs. These notes were updated based on each successive article and continual references back to the first

Note: Orange = Prepare, Green = Construct, and Blue = Use. Eight of the TLSs were described as non-linear, but also had a reference to a general order of events, either by numbering the phases or following arrows on a flowchart. For the purpose of Figure 2, I used the general order but note that non-linear paths are possible.

Findings

The findings are organized in two sections: (1) points of convergence, and (2) points of divergence across the TLSs.

Points of Convergence

I found that all TLSs encompassed three overarching goals for learners, with common patterns of sequencing of those goals. These two points of convergence are detailed in turn, followed by future considerations for science education researchers and scholars.

Point of Convergence #1: Common Goals

The three common goals were: (1) prepare for modeling, (2) construct models, and (3) use models. Each goal comprises at least two possible components, and every TLS utilized at least one component from each goal. While every component was not a part of every TLS, below I describe all components to better articulate the nature of the overarching goal.

Prepare for modeling. All of the TLSs include a component in which learners have the opportunity to prepare for the modeling process, with two options, as explained in turn.

Engage with a Catalyst. For this component, learners engage with an initial question, scenario, piece of information, or phenomenon that can serve as the catalyst for future modeling activity. This component is typically open-ended in nature (Garrido & Couso, 2025) and might include the teacher presenting a physics word problem (Sands, 2021), engaging students in generating a hypothesis after examining a data set (Khan, 2007), or asking students to consider an everyday, relatable phenomenon (Baek et al., 2011). During this component, learners may begin to engage in analogous thinking (Clement, 1989).

Engage in Pre-modeling Empirical Investigations. Learners may prepare for modeling by recording observations (Louca & Zacharia, 2015) or completing laboratory activities (Wells et al., 1995; Windschitl et al., 2008). This component privileges systematic data collection.

Construct models. Across all TLSs, learners are constructing models. To accomplish this goal, they may be prompted to consider connections among ideas and phenomena, consider causal relationships between variables, and/or be explicitly asked to express ideas through particular representations.

Consider Connections Among Ideas and Phenomena. TLSs describe this component as explicitly thinking about and comparing ideas with a question, scenario, prior piece of information, or investigation. This typically leads to an elicitation of their initial ideas to explain a phenomenon. As Garrido and Couso (2025) articulate it – “an explicit reference to the objects and relationships, principles, and so forth behind the students' predictions, hypothesis or initial explanations” (p. 404).

Consider causal relationships. When generating explanations of phenomena, one often needs to consider causal relationships, or how one variable under study directly influences another. Wells et al (1995) showcased students explicitly being asked to construct a mathematical relationship,

Force = mass * acceleration, which could explain a modified Atwood machine. In doing so, students begin considering the relationships among variables; given the same mass, if force is increased, acceleration is increased.

Express Ideas through Representations. Learners may be explicitly encouraged to express their models through different types of representations, for example diagrams, concept maps, and flow charts (Windschitl et al., 2008). Representations are not limited to visual aids, however, as this component can also include verbal presentations (Wells et al., 1995) and role plays (Chen, 2022). As learners express their ideas, they may also be asked to consider the nature and purpose of models to support their thinking (Schwarz et al., 2009).

Use models. The TLSs placed great emphasis on students using their models. Students may be asked to use models to interpret and predict phenomena, compare and evaluate models, and/or revise models in light of evidence or theoretical ideas.

Use Models to Interpret and Predict Phenomena. Conceptual models can help learners explain what happens in the physical world. The interviews with scientists from Clement's (1989) study demonstrated that they developed their models of physical phenomena with the intention to help predict what might happen in a given scenario. Schwarz et al (2009) provide an example of an elementary student using their model of water movement and phase changes to interpret what happens inside a solar still. The student represented the evaporation of water at the molecular level and demonstrated condensation along the solar rim. In this way, the student is using their model to both interpret and predict what is happening in a real-world phenomenon. Students might begin explicitly mapping parts of their model to the phenomenon (Louca & Zacharia, 2015).

Compare and Evaluate Models. In addition to explaining phenomena, learners will be able to strengthen their models if they compare and subsequently evaluate their models against one another. Chen's (2022) study pointed to the value of group presentations of 5th grade students' models of a human diaphragm. With the opportunity to compare and evaluate, learners are internalizing meaning and figuring out how different representations and designs might supplement or modify their developing knowledge. As learners evaluate their models, they may be asked to consider the criteria for revision (Baek et al., 2011), or how data helps them think about the theoretical parts of the model (Windschitl et al., 2008).

Revise Models in Light of Evidence or Theoretical Ideas. When students use their models to understand phenomena, they may find inconsistencies between what their models' explanatory or predictive value and what they observed or logical arguments posed to them. This may lead to a consideration of changing and/or revising the model. As an example, Khan (2007) describes a teaching episode within a generate-evaluate-modify sequence, during which students constructed models supporting the notion that high molecular mass of a substance results in higher boiling points. When they tested this model against teacher-provided data, they found that hydrogen bonding was a better determinant for boiling point, and thus revised their model to reflect this relationship. In an example of Stammen et al.'s (2018) TLS, high school students often revise their models of DNA replication based on an analysis of provided data.

Point of Convergence #2: Common Patterns of Sequence

In addition to sharing three overarching goals, there were common trends in the sequencing of these goals across all TLSs. Figure 2 highlights diversity in the number of phases. However, it highlights important commonalities in the sequencing of the goals:

- All 17 TLSs that includes a potential linear sequence of phases begin with the goal of preparing for modeling

- The goal of constructing models always follows initial preparation, whether it occurs within the same phase as preparation or has its own dedicated phase
- The goal of using models always follows the goal of constructing models, whether those two goals are in the same phase, or when the goal of using models occurs in a subsequent phase
- The final phase of all 17 TLSs includes the goal of using models

This common sequence – preparing before constructing and constructing before using – helps us see a strong point of convergence in how different science education scholars are conceptualizing modeling in grades K-16 learning contexts.

Future Considerations

Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- Relying on the strong foundation of the commonalities mentioned above, consider reviewing TLSs being used in work with teachers to see if gaps exist with certain components or goals and how they can be strengthened or become more explicit.
- When developing new TLSs that include other components, goals, or goal sequences, consider using the commonalities in this article as a normative reference of comparison. New variations can now be better understood and unpacked when compared to shared understandings of the past.

Points of Divergence

While there is commonality in the sequence of overarching goals, there are also several key differences that emerged from my examination. Each of the five points of divergence are explained below with illustrative examples across the TLSs. In this section, I provide future considerations for each point of divergence.

Point of Divergence #1: Model Evolution or Scientific Inquiry as the Organizing Feature Illustrates Different Emphases on a Few Key Components and the Extent of Instructor Guidance

Across several of the TLSs, two different organizing features emerged. This difference can be best articulated from a comparison of Clement's work on the one hand (and those who cite him as an influence), with several other scholars on the other hand (e.g. Chen, 2022; Halloun, 2007; Schwarz et al., 2009; Windschitl et al., 2008).

The model evolution focus. Clement (1989) presented a model construction cycle in an influential chapter from the *Handbook of Creativity* (Glover et al., 1989). The cycle is based on how scientists engage in hypothesis development and gives considerable attention to the cognitive complexities involved with constructing and refining mental models. First, initial model construction, which occurs when encountering a phenomenon or scenario, influences and is influenced by observations and the activation of analogies. Drawing on the analysis of think-aloud protocols in which scientists were asked to develop explanations for vexing physics scenarios, Clement (1989) asserted that analogy (in conjunction with prior knowledge) plays a crucial role in generating hypotheses and models. As an example, one think-aloud participant developed an analogous case of

pushing on a lever to explain the optimal place to push on a wheel when rolling it up the hill. A person's formation of an analogous case to the phenomenon under examination can serve as an initial model or prompt new questions or observations to consider. Additionally, the eventual explanatory model has analogous features that relate to the target phenomenon.

Clement's (1989) cycle allows for a model testing phase that must include the possibility of continual modification through "a series of successive refinements" (p. 347); hence the phrase, model evolution. Based on Clement's cycle (1989), the *model of modelling framework* was developed and discussed in two publications (Gilbert & Justi, 2016; Justi & Gilbert, 2002). Again, the critical role of model revision was highlighted in this framework, suggesting that learners need to be able to engage in significant revision processes after an initial construction phase.

Examples of modeling in action are detailed in the exploration of a high school student learning about energy transformation and electricity through tutoring (Clement & Steinberg, 2002) and in the description of a protocol for modeling in a life science class (Clement, 2008). In these examples, the focus is on model evolution, or scaffolded refinements in the initially expressed model. In the electricity example, there is a back and forth between demonstrations and trying out ideas with a circuit kit and probing questions from the tutor. Through the use of "small analogies" and "small discrepant events" (Clement & Steinberg, 2002, p. 440), the goal of the tutor is to challenge the student's erroneous ideas in the support of a revised and more accurate model. In one specific discussion, when the student asks, "Do all wires have a tendency to want to...[have equal inflows and outflows]?", the tutor replies, "No. Now, knowing that that doesn't have to be true, think about this again" [continues to ask a probing question about the concept of pressure in a circuit] (p. 411). The same continuous cycle of generating, evaluating, and modifying iterations of a model is described in the life science class as well. Clement (2008) presented an example of an eighth-grade class engaging in continual changes of their jointly constructed model of human lungs. These changes were made after discrepant questions from the teacher, e.g. "Does your model show air going out the bottom of the lung into the body?" and "Are there operations where they remove one lung" (p. 13)? These examples clearly allow for a significant amount of teacher planning, guidance, and discursive intervention. While constantly engaged in questioning and eliciting students' ideas, Clement's (2008) sequence is "teacher directed about 85% of the time" (p. 15). This degree of teacher guidance is supported by Khan's (2007) study, in which an undergraduate chemistry instructor engaged in many directive actions, e.g. demonstrating how to read data sets, selecting variables for students to test, and selecting extreme cases to push students' thinking.

In addition to analogies and revision, Justi and Gilbert's (2002) flow chart brought a new delineation to modeling as a sequence - separate steps for (1) producing a mental model and then (2) expressing that model in various modes of representation, which refer to media through which models are developed and communicated (Boulter & Buckley, 2000). These modes include concrete (3D material constructions), verbal (description, explanation, narrative), visual (diagram, simulation, animation), mathematical (formulae, equations), gestural (body movements), and any mixtures of the five (e.g., verbal and visual).

The scientific inquiry focus. Windschitl's et al.'s (2008) articulation of a TLS is rooted in an argument for authentic scientific inquiry. They provide a historical backdrop of *the scientific method*, a common sequence that developed largely from the work of John Dewey (1910) as a heuristic for science learning, which proliferated through science education curricula throughout the 20th century. They argued its linearity, assumption of 'one' method, and lack of focus on epistemic activity all make it an unsuitable organizer of instruction. As a replacement, they developed and proposed "model-based inquiry." They developed this approach based on the epistemic practices of scientists: "we must first identify core knowledge producing activities that scientists engage in, then consider current classroom practice in light of these" (Windschitl et al., 2008, p. 943). The four phases are rooted in

five features of scientific knowledge – namely, that it is testable, revisable, explanatory, conjectural, and generative. The proposed outcome is not only a model, per se, but the use of a model to construct a defensible, evidence-based explanation. The four phases are conversations that engage learners in (1) Organizing what we know and want to know; (2) Generating testable hypotheses; (3) Seeking evidence; and (4) Constructing an argument.

For those authors whose conceptual anchor is modeling as a scientific practice, more priority seems to be given to students engaging in an empirical investigation within the TLS. Windschitl et al. (2008) note that empirical investigations are not an absolute requirement in conversation #3, but many of their example prompts point to that aim (p. 958):

- “How can we operationalize our variables in ways that will allow us to record unambiguous measurements?”
- “What will it mean to collect data “systematically”?”
- “To test our hypothesis, should we observe the phenomenon as it is or actively manipulate some variables while controlling others?”

Other scholars whose approach is rooted in authentic scientific practices seem to privilege students engaging in the design, data collection and analysis of empirical investigations. They have set aside specific phases for empirical investigation, either as a pre-modeling activity (Schwartz & Gwekwerere, 2007; Wells et al., 1995) or as a learning experience to take place after the initial construction of a model (e.g. Baek et al., 2011; Halloun, 2007; Louca & Zacharia, 2015; Schwarz et al., 2009). Schwarz and Gwekwerere (2007) assert that the *investigate* phase of their EIMA framework should give “high priority for data collection and analysis of those data into patterns” (p. 160). Halloun’s (2007) approach includes two phases for empirical investigation. The author articulates and provides examples of how modeling is “teacher-mediated” (p. 3) but also points to times when they must step back: “all through the actual investigation process, the teacher retracts from the arbitration role to supervise the process from a distance as a moderator” (p. 26). Wells et al. (1995) noted the importance of allowing students the opportunity to design their own investigations: “the teams are permitted to pursue their own experimental procedures without intrusion from the instructor” (p. 616).

In addition to empirical investigation, another emerging and highlighted feature from the scientific inquiry TLSs is to engage learners in metamodeling prompts as they construct and use models. The goal, here, is for students to reflect on the norms of modeling as they engage in the practice itself (Schwarz et al., 2009). For example, in association with the component of *expressing ideas through representations* (constructing models), Schwarz et al.’s (2009) TLS asks learners to consider the nature and purpose of models as part of that component. The scholars who are explicit about engaging students in discussion about the qualities for evaluation and revision of models (Baek et al., 2011; Chen, 2022; Windschitl et al., 2008) tend to nest themselves in the science inquiry foundation. Students are prompted to consider questions like “How can I negotiate my model?” (Chen, 2022, p. 392) and “How do the data help us infer about theoretical (unobservable) events in our model?” (Windschitl et al., 2008, p. 959), thereby adding an extra dynamic to the goal of using models.

Comparison of foci. The model evolution approach from Clement and Steinberg (2002), and those who cite Clement’s model construction cycle (1989) as major influences (e.g. Gilbert & Justi, 2016; Justi & Gilbert, 2002; Khan, 2007) tend to focus on mental modeling. This focus originates in the cognitive sciences. Clement (1989) references the constructs of conceptual change and mental processing, and both Justi and Gilbert (2002) and Khan (2007) reference cognitive psychologists, such as Johnson-Laird (1983). In turn, their focus on mental processing seems to have given rise to the inclusion of analogical thinking and the different media by which we *express* representations (e.g.

referring to *modes of representation* more so than representations as final products). Rather than emphasizing these types of references when building the justification for a TLS, scholars focused on the scientific inquiry perspective are more likely to cite literature that highlights the authentic practice of modeling within disciplines of science (Windschitl et al., 2008) and calls for reform in science education to reflect that authenticity (Louca & Zacharia, 2015; Schwarz et al., 2009). This focus is more likely to be accompanied with phases that privilege empirical investigations and opportunities to engage in metamodeling discussions about modeling as they construct and use models. The delineation is by no means perfect, as one can find references across the TLS articles that mention mental modeling and scientific practices or inquiry. However, the differing emphases across these two camps brings about helpful considerations for those will engage learners in modeling TLSs in the future.

Each of the four actions in this comparison has value. From the model evolution perspective, both analogical thinking and modes of representation are important. Regarding analogical thinking, learners consider their prior knowledge and seek associations between the phenomenon under study and prior knowledge as well as previous phenomena they have encountered (Clement, 1989). In some cases, these analogies can end up being unhelpful due to erroneous connections (Dolphin and Benoit, 2016), and so placing emphasis on analogical actions while constructing and using models can be helpful. Regarding modes of representation, the emphasis on the medium of representation instead of representations as products can be powerful. Many of the representations listed in TLSs are visual products, like diagrams (Windschitl et al., 2008) and dioramas (Garrido & Couso, 2025). Bending toward the perspective of representational *modes* opens up many options for learners to consider. The students in Chen's (2022) study relied on gestures through role play as they enacted a model of the human diaphragm. Using chairs as diaphragms and students playing the role of air molecules, they were able to represent the relationship between space and pressure through movement. Fackler and Ruth (2024) found that elementary pre-service teachers simultaneously "used their gestures, drawings, and verbal expressions as a central component of their thinking about molecules getting, having, and releasing energy" (p. 11). Their study supports the argument that the deployment of multiple modes of representation can contribute to meaningful learning when students engage in modeling.

From the scientific inquiry perspective, engaging learners in empirical investigations and metamodeling discussions are also worthwhile endeavors. Designing and conducting investigations has a long tradition in science education reforms (NRC, 1996; NGSS Lead States, 2013), to engage students in a considerable facet of scientific work. Students have the opportunity to participate and understand the important role of empiricism in generating scientific knowledge, which supports their developing understanding of what science entails (i.e., their science literacy). Indeed, understanding the role of inquiry in science is thought to be a central aim in producing a scientifically literate population (Lederman, 2007). The integration of metamodeling discussion into the goals of constructing and using models parallels previous calls for understanding scientific inquiry. These metamodeling discussions will help them understand the nature of the scientific practice.

Future Considerations. Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- Consider both foci (model evolution and scientific inquiry) when developing or deploying the TLS with teachers and students.
- For TLSs that may be rooted in one focus over another, consider updates to the TLS that more explicitly integrate the other focus. An example of this integration is provided in Table 1 by comparing Clement and Steinberg's (2002) model evolution focus with Windschitl et al.'s (2008) scientific inquiry focus. The point of Table 1 is not to criticize the two TLSs but to illustrate what these considerations might look like in practice.

Table 1.*Example of Including the Model Evolution Focus and Scientific Inquiry Focus into TLSs*

TLS Focus	TLS phases	Possible actions to include from the other perspective	Components aligned to actions from the other perspective
Model evolution (Clement & Steinberg, 2002)	Generate Model 1		
	Evaluate Model 1 • Learners experience observation discrepancies and consider analogous case • Learners modify model to...	Engage learners in an empirical investigation	Empirical investigation
	Modify Model 1 into Model 2	Engage learners in metamodeling questions: • How does the model help me connect two ideas? • What parts of my model are not helping me explain the phenomenon?	Metamodeling about constructing and using models
Scientific Inquiry (Windschtil et al., 2008)	<i>Conversation 1 – Organizing what we know and what we want to know.</i>	Engage learners to consider which modes of representation (visual, verbal, mathematical, gestural, and concrete) will be most useful.	Modes of representation
	<i>Conversation 2 – Generating testable hypotheses.</i>	Consider potential analogies to help learners make the initial leap from phenomenon to the initial model.	Analogies
	<i>Conversation 3 – Seeking evidence.</i>		
	<i>Conversation 4 – Constructing an argument.</i>	Support learners in the use of analogies to help explain the phenomenon	Analogies

Point of Divergence #2: The TLSs Differ with Respect to Being Linear or Non-linear

The second issue of divergence is linear vs. non-linear TLSs (see the appendix for the designation of each one).

Description. Of the 18 TLSs, ten are presented as linear, with one phase proceeding to the next in one direction. This could occur as a vertical list of phases, such as the case of Sands (2021) or Baek et al. (2011). Or, this could be described as a cycle yet with no indication of moving in any direction but forward, such as the case of the GEM cycle described by Khan (2007).

Eight of the TLSs are non-linear or have non-linear options. Seven have an overarching linear path one can follow, but the representations of the TLS and/or authors denote that one can move back and forth across the phases or there are points of divergence with options to move toward one

phase or another. For example, Chen (2022) lists four phases vertically, but with one double arrow running across all phases, noting:

It is important to understand that this framework does not reflect a linear process, even if represented in a linear manner. The process of modeling is a dynamic, iterative, and cumulative practice in which teachers often move back and forth between different phases depending on how the various epistemic uncertainties occur and are managed across phases. (p. 388).

Such declarations are common among the non-linear TLSs. Windschitl et al. (2008) share it is “rarely the case” (p. 960) that their overarching sequence would proceed exactly in the order of their four conversations, referring to how scientists’ work would unfold: “scientists often learn things that put them back into conversations about the models and assumptions they are working from” (p. 960). Both Buckley et al. (2004) and Justi and Gilbert (2002) present their TLSs in the form of a flow chart, following the actions of a modeler. For Buckley et al. (2004), once a learner forms a mental model and interacts with phenomena or other models in the curriculum, they might go in one direction and begin evaluating the model or go in the other direction and revisit or add new information to their existing model. Garrido and Couso (2015) built in a mini-cycle of learners expressing, using, and revising models, before moving on to the next phase of building a consensus model.

One of the non-linear TLSs is unique from the other seven. Gilbert and Justi’s (2016) version 2 of their *model of modelling framework* is represented as a tetrahedron with four vertices that are all equidistant from each other (each vertex represents a phase). Modelers in this TLS will engage in creating, expressing, testing and evaluating models in any particular order or sequence, with no preference for the entry point: “the use of this geometric shape emphasises three essential characteristics of the process: to be cyclic, non-linear, and not predetermined” (p. 36). In a departure from their first *model of modelling framework* (presented in Justi & Gilbert (2002) and Gilbert & Justi (2016)), version 2 “was produced by having in mind that modelling is a scientific epistemic practice” (p. 36), a nod to similar arguments from Windschitl et al. (2008).

An underlying reason for the two different approaches is not explicitly clear. As mentioned above, it would make sense for those scholars rooted in the scientific inquiry perspective to privilege non-linear TLSs, but this is not always the case. The inquiry perspective has led to linear sequences (e.g. Louca & Zacharia, 2015), and Buckley et al. (2004)’s approach cites previous mental modeling research and presents a non-linear path. The one intriguing trend is this: of the five TLSs that provide an example of modeling in elementary contexts (Baek et al., 2011; Chen, 2022; Louca & Zacharia, 2015; Schwarz & Gwekwerere, 2007; Schwarz et al., 2009), four present linear sequences and one (Chen, 2022) was presented in a mostly linear way.

It is possible that those science education researchers working in elementary contexts are inherently thinking about the clear articulations and scaffolding needed to support early learners. A linear sequence has the potential to reduce the cognitive load for both learners and teachers. Baek et al. (2011)’s TLS included eight phases, with actions broken up into clearly articulated actions. There is a clear phase for constructing the initial model, then engaging in an empirical investigation, then engaging in scientific ideas and computer simulations, and so on. When comparing the flow charts of Buckley et al. (2004) and Justi and Gilbert (2002) with the TLS of Baek et al. (2011), it becomes evident that the former TLSs give insight into the thinking processes a modeler undertakes, while the latter TLS provides more cues for a teacher about how the time will be organized. For example, Baek et al.’s (2011) phase dedicated to peer evaluation seems to point to how the students should interact in a typical classroom setting. This may explain why several of the TLSs applied to elementary contexts include a greater number of phases, each with a specific set of actions and at a fairly small grain size (e.g. Baek et al., 2011; Garrido & Couso, 2025; Schwarz et al., 2009).

The partnership between elementary contexts and linear TLSs raises the broader question of whether non-linear TLSs influence teachers’ planning. Cases of significant flexibility and non-linearity are difficult to find. For example, an examination of learning activities in Buckley et al.’s (2004) study

resembles a typical curriculum in which one activity proceeds to the next without a clear mapping back to their non-linear TLS. There are a few cases of a teacher combining two phases, but still in a mostly linear sequence. The description of learning experiences in Chen's (2022) study suggests the teacher completed phase 1 entirely before moving on (problematizing a phenomenon), then engaged in a back and forth of phase 2 and 3 (material practice and argumentative practice), then completed phase 4 (conceptualizing theory and application of final models). Mendonça and Justi's (2011) study of middle school chemistry students in Brazil tested the use of the *model of modeling framework* (Justi & Gilbert, 2002). The list of learning experiences presented in table format follows the flowchart in a mostly linear fashion, proceeding along four ordered sub-processes, but occasionally including two sub-processes together (e.g. sub-processes 1 and 2 occurring together, producing a mental model and expressing it in modes of representation). Despite these findings, one can argue that an overly linear sequence can make it difficult to maintain connections to authentic scientific practice (Windschitl et al., 2008).

Future Considerations. Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- For those using linear TLSs, consider representing instances of moving back and forth across actions or components when they're implemented in K-16 learning contexts. Where might there be opportunities to represent that back and forth within the TLS?
- For those using non-linear TLSs, consider the extent of non-linear actions among modelers and whether those actions can be bundled together in a clear phase that guides a teacher's planning.

In Figure 3, I present an example of the difference between an idealized model and an implemented model from Chen's (2022) article. The model presented in the introduction of the article includes four separate phases, but with options to move back and forth among them. However, as mentioned above, during implementation, only phases 2 and 3 were combined. While this was only one iteration of implementation, if such a reality was observed repeatedly, it could lead to a modification of the TLS. In the bottom row of Figure 3, the TLS has three overarching phases while acknowledging an inherent back and forth between material and argumentative practice in phase 2. Perhaps that update would be misguided after more attempts at implementation, but the Figure illustrates an example of balancing the benefits of a mostly linear sequence and maintaining authenticity in scientific work.

Figure 3

Comparison of Chen's (2022) Introduced and Implemented TLS

Chen's (2022) TLS	Phases progressing left to right			
TLS presented in introduction	1. Problematizing a phenomenon	2. Material practice	3. Argumentative practice	4. Conceptualizing theory and application of final models
TLS based on description of unit implementation	1. Problematizing a phenomenon	2. [Combined phases] Material practice (develop and explain models)  Argumentative practice (evaluate and revise models)		3. Conceptualizing theory and application of final models

Point of Divergence #3: The Total Amount of Time Needed to Complete a TLS is Rarely Clear, and Differences in Duration Appear to Exist

It is the nature of TLSs to have some flexibility in implementation, particularly with regard to the total length of time it takes to implement it. Nonetheless, it is important to share the finding that a substantial range of possible timeframes appears to exist. I use the word ‘appear’ because there is not a common unit that authors use when providing examples of the TLSs during implementation.

Description. At one end of the continuum, Schwarz et al.’s (2009) 7-phase TLS was applied to 6-week units in 5th and 6th grade classrooms. While it is not clear how days, lessons, or hours were spent on the unit, interview references to “lesson #15” seem to imply a lengthy total duration (p. 645). At the other end of the continuum, Khan’s analysis of generate-evaluate-modify cycles in an undergraduate chemistry course revealed that the instructor facilitated 52 total cycles across a semester, with an average of 2 cycles per class period. Stammen et al.’s cycle did not include a timeframe, but after reading the example of high school students developing possible models of DNA replication with pop beads and chenille stems, it seems the activity (or a modified version of it) could be completed in a class period or two since there is no need for extensive experimental design. The number of phases does not necessarily translate to the total duration, as Wells et al.’s (1995) TLS only consisted of two phases but can take 2-3 weeks to complete (there is considerable work to be done in their ‘model development’ phase). Some TLSs can be completed in about a week; the *model of modelling framework* (Justi & Gilbert, 2002) has been implemented during five in-class activities (Justi et al., 2009). The key issue here is that time is very important to teachers, as they consider how to parse and organize it when making decisions (Hargreaves, 1994). Yet, currently, unless the author of the TLS is working directly with the teacher, it is not clear what teachers should expect with regard to total time, relative time on each phase, and the adaptability of the TLS to be implemented in shorter or lengthier timeframes.

Future Considerations. Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- Since references to weeks or lessons do not provide a common unit of understanding, consider reporting projected or implemented timeframes in the unit of *instructional hours*.
- Consider reporting projected or implemented timeframes for each phase in the TLS

Point of Divergence #4: The TLSs Differ with Respect to the Extent of Information Regarding Teachers’ Actions, with Most TLSs Written from the Learner’s Perspective

The TLSs varied in how they wrote about and described the TLS.

Description. For the majority of TLSs, the description of phases is written primarily from the learner or modeler’s perspective, a concern noted by Garrido and Couso (2025) prior to developing their TLS. Most often, a description of what the learners are doing accompanies the title or summary of each phase. Examples of learner actions include considering an introductory phenomenon, exploring relationships among variables, expressing mental models, and sharing and revising those models after opportunities to test them. The role and actions of the teacher are not as clear. In 16 of the 18 articles, the teacher’s actions were explicitly described for some (but not all) phases or for none of the phases. A breakdown is provided below.

- 9 TLSs include only the learners’ perspective for each phase
- 7 TLSs includes the learner’s perspective and partially the teacher’s perspective

- 1 TLS includes only the teacher's perspective for each phase
- 1 TLS includes both the learners' and teacher's perspectives for each phase

There is value in describing both the learners' and teachers' role when communicating about science instruction, as the reader can better imagine and simulate the phase when considering both perspectives. Be it for science education researchers, curriculum developers, or science teachers, describing both learner and teacher actions helps to communicate a fuller scope of each phase.

Table 2 presents an example of comparing the described actions across two TLSs. Baek et al. (2011) describe the learners' actions for their *peer evaluation phase*, and Garrido and Couso (2025) describe both learners' and the teacher's actions in their *express consensus model* phase. These two phases are at similar places of their respective TLSs (toward the end). The purpose of making this comparison is not to criticize Baek et al.'s (2011) TLS, as they provide clear and helpful learner descriptions of each phase. In addition, there is a mention of a teacher giving a "mini-lecture" (p. 203) when describing the implementation of a particular unit, but actions like this are not articulated in the table that presents each phase. Rather than serving as a source of criticism, the comparison helps the science education community understand what can be gained from being just as descriptive about the teacher's role as we are the learners' role. At a general level, all articles acknowledged the importance in the teacher's role to bring about and support meaningful learning in the TLSs. Meaning-making in science classrooms relies on both dialogic and authoritative discourse, and the learning experiences – whether more controlled or more open-ended – are introduced and monitored by the teacher (Halloun, 2007; Mortimer & Scott, 2003; Sickel et al., 2013). In Garrido and Couso's (2025) description of the *express consensus model* phase, we learn important responsibilities of the teacher - guiding a class discussion, promoting the sharing of errors and challenging ideas, and ensuring the funneling process leads to a consensus model that is scientifically accurate.

Table 2

Learners' and Teachers' Actions for Baek et al's (2011) Peer Evaluation Phase and Garrido and Couso's (2025) Express Consensus Model Phase

Authors	Learners' actions	Teachers' actions
Baek et al., 2011, p. 201	• "each student presents his/her model and the others give their evaluative feedback on it" • [students] "discuss criteria for evaluating models"	
Garrido & Couso, 2025, p. 406	• "students' models are shared and made public among the entire group"	• Guide "a collaborative negotiation and synthesis of students' explanations into a completer and more sophisticated final consensus model" • "cultivate a positive and secure classroom environment wherein errors are embraced" • make sure the final model "is aligned, similar, and coherent to the instructional target"

Future Considerations. Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- Consider describing the phases from the perspective of learners' actions/roles *and* the teacher's actions/roles, to communicate the nature of the phase more completely to readers.

- If educative curriculum tools accompany the TLS, consider uploading them online as supplementary items to the published article.

Point of Divergence #5: There are more Examples of the TLSs within the Physical Sciences when Compared to Life Science and Earth Science

Articles and book chapters under this review were selected if they dedicated space to proposing and describing a TLS. In 13 of the 18 articles or book chapters, the author(s) also supplied an articulated example of the TLS in action and how it might unfold in a classroom. In the following section, I discuss these examples in accordance with the science disciplines represented.

Description. The following science disciplines were represented across 13 TLSs as shown in table 3 below. Louca and Zacharia (2015) provided two examples in their article, one in physical science and one in life science (hence, the total number of examples is 14):

- 10 articles discussed physical science examples
- 3 article discussed life science examples
- 1 article discussed an Earth science example

Table 3 presents a breakdown of the example topics organized by grade band, author, and science discipline.

Table 3
Topics Addressed in Examples of TLSs

Science discipline	Author(s)	Grade Band	Topic(s)
Physical sciences	Clement (1989)	University doctoral students	Mechanics
	Hestenes (1987)	University undergraduate	Mechanics
	Wells et al. (1995)	University undergraduate and secondary	Mechanics
	Sands (2021)	University undergraduate	Mechanics
	Khan (2007)	University undergraduate	Vapor pressure and boiling point
	Clement & Steinberg (2002)	Secondary	Electric circuits
	Schwarz et al. (2009)	Middle school and elementary	Properties of matter (evaporation and condensation), and the nature of light
	Baek et al. (2011)	Elementary	Properties of matter (evaporation and condensation)
	Garrido & Couso (2025)	Elementary (participants were pre-service teachers but the topic is aligned with elementary)	Floating, buoyancy, and forces
	Louca & Zacharia (2015)	Elementary	Accelerated motion
Life science	Stammen et al. (2018)	Secondary	DNA replication
	Chen (2022)	Elementary	Human respiratory system
	Louca & Zacharia (2015)	Elementary	Plant parts and their role
Earth science	Windschitl et al. (2008)	Middle school	Motion of Earth, Moon and Sun

In the vast majority of articles, the examples that accompanied the TLSs included topics that have a long tradition in school science; they could be investigated with common school supplies and tested in controlled settings with immediate feedback. Examples in the physical sciences include pre-service teachers using cups and water to investigate the forces involved with floating (Garrido and Couso, 2025) and a secondary student testing out ideas about the nature of electric currents using a basic circuit kit (Clement & Steinberg, 2002). For life science, elementary students used their bodies and classroom chairs to model how a lung works (Chen, 2022), and secondary students used chenille stems and beads to test out their ideas about DNA (Stammen et al., 2018).

Illustrations of the TLSs in life science and Earth science have great potential to reveal nuances that may need to be considered for those disciplines. Biology was once a discipline characterized by stores of information such as classifications of organisms and anatomical descriptions. We are now in the age of studying life at the molecular level (Flannery, 2001), which means biology relies heavily on a foundation of physical science (Fanelli & Glänzel, 2013). Due to the *physics first* movement, there are arguments for secondary students to learn physics and chemistry before they complete a biology course in 11th grade (Bybee & Gardner, 2006). Earth science is also viewed as having an overarching and integrative role in the science curriculum. OpenSciEd chose to construct their high school curriculum into the separate disciplines of biology, chemistry, and physics, but paired Earth science with each of them (Activate Learning, n.d.). Complex phenomena, such as climate change (often associated with Earth science), incorporate understandings of energy (physical science) and understandings of organic processes like photosynthesis (life science). With this perspective, conceptual models in life and Earth science have the likelihood of becoming extremely complex and abstract. Zangori et al.'s (2017) study of secondary students creating models in a unit on climate change seems to support this notion; they found that students needed to first grapple with the inner-workings of the carbon cycle before they could make meaningful connections to the more complex phenomenon of climate change.

More investigations of Earth and space science may reveal a component in constructing models that is not explicit across the 18 TLSs in this review, that of *physical scale and perspective*. As evidenced from research on levels of representation (Chandrasegaran et al., 2008) and research on students' understanding of scale (Jones & Taylor, 2009), understanding scientific concepts requires students to think about the relative scale of phenomena, from the large-scale macro world to the small-scale micro world. When constructing models, it might be helpful for learners to think about how scale of phenomena may play a role in their model. For instance, spatial reasoning and scale are both important to and can inhibit learning about Earth-Moon-Sun phenomena (Wilhelm et al., 2018). Investigations may require consideration of the sizes of the Earth, Sun, and Moon, and/or the distances between them, and how variables of magnitude can be represented within a model. Spatial orientation is particularly important when science educators rely heavily on two-dimensional models of complex phenomena. Park et al.'s (2023) investigation of 10th grade students' drawings of water and sustainability revealed that focusing on only one spatial orientation, e.g. top-down, can highlight some important components (i.e. runoff) while ignoring other ones (i.e. groundwater). Beyond an exploration of causal factors and connections among ideas and variables, *physical scale and perspective* may need more attention within the goal of constructing models.

Future Considerations. Researchers and scholars engaging with TLSs for modeling are encouraged to do the following:

- Consider including descriptive examples of the TLS within life science or Earth science learning contexts, to provide the field more examples across the disciplines and potentially uncover new factors to examine based on the nature of the discipline.

- Consider including descriptive examples of the TLS that are complex and integrative across disciplines, e.g. climate change, to understand how TLSs work with topics that cannot be investigated quickly with typical school science materials.
- Consider placing attention on the role of physical scale and perspective in modeling, and how that component may connect to an overarching modeling goal (e.g. constructing models).

Conclusion

I have reviewed the literature to understand the various ways science education researchers and scholars have developed and presented TLSs for the scientific practice of modeling. In doing so, points of convergence and divergence have been examined, and future considerations have been identified. I believe those considerations provide practical advice that can help move the field of modeling education forward in productive ways. Zooming out from those considerations, I also envision that science education researchers and scholars might consider three broad research agendas in this space – what I will call the *synthesis*, *visionary*, and *scalability* agendas. I advocate that working across these agendas will be helpful, because they will push and pull on one another as the modeling education field pursues its potential while also thinking about issues of practicality and diverse learning contexts.

The **synthesis agenda** would be concerned with developing a meta-model based on the extant literature. The meta-model might be organized by the common goals and sequencing from this review (see points of convergence above). It would inventory the range of components and goals, including those identified in future TLSs, ensuring the modeling evolution and scientific inquiry perspectives are balanced and taking stances on other key points of divergence (e.g. the issue of linearity and the need for projected timeframes). The meta-model would be developed from the argument that it is built off a foundation of over 30 years of prior TLSs, and therefore benefits from a strong foundation across multiple academic lineages.

The **visionary agenda** would be concerned with developing TLSs in alignment with particular visions. In many ways, prior TLSs have been visionary. Clement's work took a careful look at the cognitive intricacies of modeling and opened up new avenues for designing learning sequences in ways that helps learners refine their mental models. Windschitl et al's (2008) model-based inquiry proposed a very different way of thinking about scientific practice in the classroom (when compared to "the scientific method"), allowing for a learning environment characterized by active learning and epistemic agency. Researchers and scholars should continue to pursue visions that are in support of improving science education, and how TLSs for modeling can be designed to support those aims. For example, what TLSs for modeling might best support learning complex socio-scientific and interdisciplinary topics?

Finally, the **scalability agenda** would be concerned with developing and testing TLSs in the pursuit of discovering what is most scalable across a wide range of learning contexts. Currently, there is no evidence that sophisticated modeling-based teaching is occurring at a large scale. Individual research teams are trying out their TLSs with groups of teachers (e.g. Schwarz & Gwekwerere, 2007) and often find that the adoption of those TLSs is quite challenging and ambitious for them (e.g. McLaughlin & Pringle, 2011; Wells et al., 1995; Windschitl et al., 2008). The scalability agenda would need to begin with an approach that is rarely a starting point in the science education community; the premise that school is a very different place from a science laboratory and that teachers have general learning goals for their students beyond learning science (Friedrichsen & Dana, 2005). It would take an asset-based perspective, starting with the proposition that teachers have knowledge in areas that we do not sufficiently consider. Teachers have the responsibility to cultivate and manage socioemotional relationships and take into account what motivates their students to learn (Van Sickle

& Spector, 1996). They must differentiate instruction for learners with a wide range of accommodations and modifications based on individualized learning plans. Coming from this perspective, a TLS might be designed in collaboration with teachers (McLaughlin & Pringle, 2011) and build off of teachers' general pedagogical knowledge and responsibilities. The TLS might focus on a few phases with non-negotiable components of modeling practice but that include potential actions for differentiation and scaffolding. The TLS might need to be applicable to relatively short units of learning, because school curricula play a significant role in determining how much time can be dedicated to any one topic (Fitzgerald et al., 2019; Halloun, 2007). Researchers and scholars working under this agenda might begin with some of the components provided in the points of convergence section of this paper, but ultimately navigate the tension of engaging learners in meaningful modeling while also being consistently workable for teachers.

The final consideration for the field of science education relates to how we share TLSs in peer-reviewed outlets. In several cases, the articulation of modeling TLSs was a part of a typically structured empirical paper, which means the majority of the article space was dedicated to results and findings (e.g. Louca & Zacharia, 2015; Chen, 2022). This means there is limited space to fully articulate the rationale behind the development of the TLS and the practical knowledge that needs to accompany it. For those articles that utilize a greater amount of space for their rationalization along with an articulated example (e.g. Clement & Steinberg, 2002; Garrido & Couso, 2025; Wells et al., 1995; Windschitl et al., 2008), it is incredibly helpful for the field. In those instances, editors and reviewers were willing to consider a manuscript that was atypical in structure. More of these types of articles will be needed in the future. The literature on modeling TLSs has not reached consensus, and may never do so. However, there is excellent potential in pursuing the development, articulation, modification, and future testing of modeling TLSs in K-16 learning contexts to support the implementation of this important scientific practice.

The authors received no financial support for the research, authorship, and/or publication of this manuscript

Aaron J. Sickel (sickel@hawaii.edu) is an assistant professor of STEM Education in the School of Teacher Education at the University of Hawai'i at Mānoa. He teaches courses in mathematics methods, science methods, and integrated STEM. Research interests include investigating the nature of STEM teacher learning, the design of content-specific methods courses, and the implementation of constructivist-oriented instructional models.

References

- Activate Learning. (n.d.). The future of high school STEM education: Rethinking the high school science course sequence (part 2). <https://activatelearning.com/high-school-science-course-sequence-openscienced/>
- Baek, H., Schwarz, C., Chen, J., Hokayem, H., & Zhan, L. (2011). Engaging elementary students in scientific modeling: The MoDeLS fifth-grade approach and findings. In M. S. Khine & I. S. Saleh (Eds.), *Models and modeling: Cognitive tools for scientific enquiry* (pp. 195-218). Springer. https://doi.org/10.1007/978-94-007-0449-7_9
- Boulter, C. J. & Buckley, B. C. (2000). Constructing a typology for models in science education. In: J. K. Gilbert & C. J. Boulter (Eds.), *Developing models in science education*. (pp. 41-58). Kluwer.
- Buckley, B. C., Gobert, J. D., Kindfield, A. C. H. (2004). Model-based teaching and learning with BioLogica™: What do they learn? How do they learn? How do we know? *Journal of Science Education and Technology*, 13, 23–41 (2004). <https://doi.org/10.1023/B:JOST.0000019636.06814.e3>

- Bybee, R. W. & Gardner, A. L. (2006). High school biology & the physics first movement. *The American Biology Teacher*, 68(3), 134-138. [https://doi.org/10.1662/0002-7685\(2006\)68\[134:HSBTPF\]2.0.CO;2](https://doi.org/10.1662/0002-7685(2006)68[134:HSBTPF]2.0.CO;2)
- Bybee, R. W., Taylor, J. A., Gardner, A., Van Scotter, P., Powell, J. C., Westbrook, A. et al. (2006). *The BSCS 5E instructional model: Origins, effectiveness, and applications*. Unpublished white paper. Colorado Springs, CO: BSCS.
- Campbell, T., Oh, P. S., & Neilson, D. (2012). Discursive modes and their pedagogical functions in model-based inquiry (MBI) classrooms. *International Journal of Science Education*, 34(15), 2393–2419. <https://doi.org/10.1080/09500693.2012.704552>
- Chandrasegaran, A. L., Treagust, D. F., & Mocerino, M. (2008). An evaluation of a teaching intervention to promote students' ability to use multiple levels of representation when describing and explaining chemical reactions. *Research in Science Education*, 38, 237-248. <https://doi.org/10.1007/s11165-007-9046-9>
- Chart, D. (2000). *A theory of understanding: Philosophical and psychological perspectives*. Ashgate.
- Chen, Y.-C. (2022). Epistemic uncertainty and the support of productive struggle during scientific modeling for knowledge co-development. *Science Education*, 59(3), 383-422. <https://doi.org/10.1002/tea.21732>
- Clement, J. (1989). Learning via model construction and criticism: Protocol evidence on sources of creativity in science. In: J. A. Glover, R. R. Ronning, & C. R. Reynolds (Eds.), *Handbook of creativity*. (pp. 341-381). Plenum Press.
- Clement, J. J. (2008). Student/teacher co-construction of visualizable models in large group discussion. In: J. J. Clement & M. A. Rea-Ramirez (Eds.), *Model based learning and instruction in science (Models and modeling in science education, Volume 2)*. (pp. 11-22). Springer. https://doi.org/10.1007/978-1-4020-6494-4_1
- Clement, J. J. & Steinberg, M. S. (2002). Step-wise evolution of mental models of electric circuits: A “learning-aloud” case study. *Journal of the Learning Sciences*, 11(4), 389–452. https://doi.org/10.1207/S15327809JLS1104_1
- Dewey, J. (1910). *How we think*. DC Heath and Company.
- Dolphin, G. & Benoit, W. (2016). Students' mental model development during historically contextualized inquiry: how the “Tectonic Plate” metaphor impeded the process. *International Journal of Science Education*, 38(2), 276-297. <https://doi.org/10.1080/09500693.2016.1140247>
- Fackler, A. K. & Ruth, R. M. (2024). Empowering elementary preservice science teachers: Harnessing diverse language resources in the practice of modeling. *Science Education*, published online 02 January 2025. <https://doi.org/10.1002/sci.21934>
- Fanelli, D. & Glänzel, W. (2013). Bibliometric evidence for a hierarchy of the sciences. *PLoS ONE* 8(6): e66938. <https://doi.org/10.1371/journal.pone.0066938>
- Fitzgerald, M., Danaia, L. & McKinnon, D. H. (2019). Barriers inhibiting inquiry-based science teaching and potential solutions: Perceptions of positively inclined early adopters. *Research in Science Education*, 49, 543-566. <https://doi.org/10.1007/s11165-017-9623-5>
- Flannery, M. C. (2001). Where is biology? *American Biology Teacher*, 63(6), 442-447. <https://doi.org/10.2307/4451152>
- Friedrichsen, P. M. & Dana, T. M. (2005). Substantive-level theory of highly regarded secondary biology teachers' science teaching orientations. *Journal of Research in Science Teaching*, 42(2), 218-244. <https://doi.org/10.1002/tea.20046>
- Garnham, A. (1997). Representing information in mental models. In: M. A. Conway (Ed.), *Cognitive Models of Memory*. (pp. 149-172). MIT Press.
- Garrido, A. & Couso, D. (2025). The IPM cycle: An instructional tool for promoting students' engagement in modeling practices and construction of models. *Journal of Research in Science Teaching*, 62(2), 391-425. <https://doi.org/10.1002/tea.21979>

- Gilbert, J. K. & Boulter, C. J. (2000). *Developing models in science education*. Kluwer.
- Gilbert, J. K., Boulter, C.J., Elmer, R. (2000). Positioning models in science education and in design and technology education. In: Gilbert, J.K., Boulter, C.J. (Eds) *Developing models in science education*. Springer. https://doi.org/10.1007/978-94-010-0876-1_1
- Gilbert J. K., & Justi, R. (2016). *Modelling-based teaching in science education (Models and modeling in science education, Volume 9)*. Springer.
- Gilbert, S. W. (1991). Model building and a definition of science. *Journal of Research in Science Teaching*, 28, 73-79. <https://doi.org/10.1002/tea.3660280107>
- Glover, J. A., Ronning, R. R., & Reynolds, C. R. (1989). *Handbook of creativity*. Plenum Press.
- Greca, I. M. & Moreira, M. A. (2000). Mental models, conceptual models, and modelling. *International Journal of Science Education*, 22(1), 1–11. <https://doi.org/10.1080/095006900289976>
- Grosslight, L., Unger, C., Jay, E., & Smith, C. L. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching*, 28, 799-822. <https://doi.org/10.1002/tea.3660280907>
- Halloun, I. A. (2007). Mediated modeling in science education. *Science & Education*, 16(6-7), 653–697. <https://doi.org/10.1007/s11191-006-9004-3>
- Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in a postmodern age*. Teachers College Press.
- Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics*, 55(5), 440-454. <https://doi.org/10.1119/1.15129>
- Johnson-Laird, P. N. (1983). *Mental models*. Cambridge University Press.
- Jones, M. G. & Taylor, A. R. (2009). Developing a sense of scale: Looking backward. *Journal of Research in Science Teaching*, 46, 460-475. <https://doi.org/10.1002/tea.20288>
- Justi, R. S. & Gilbert, J. K. (2002). Modelling, teachers' views on the nature of modelling, and implications for the education of modellers. *International Journal of Science Education*, 24(4), 369–387. <https://doi.org/10.1080/09500690110110142>
- Justi, R., Gilbert, J. K., & Ferreira, P. F. M. (2009). The application of a 'Model of Modelling' to illustrate the importance of metavisualisation in respect of the three types of representation. In J. K. Gilbert & D. Treagust (Eds.), *Multiple representations in chemical education. Models and modelling in Science education, vol 4*, (pp. 285-308). Springer. https://doi.org/10.1007/978-1-4020-8872-8_13
- Karplus, R. & Thier, H. D. (1967). *A New Look at Elementary School Science*. Rand McNally & Company.
- Khan, S. (2007). Model-based inquiries in chemistry. *Science Education*, 91(6), 877-905. <https://doi.org/10.1002/sce.20226>
- Larmer, J. (2015, July 13). *Project-based learning vs. problem-based learning vs. X-BL*. Edutopia. <https://www.edutopia.org/blog/pbl-vs-pbl-vs-xbl-john-larmer>
- Lederman, N. G. (2007). Nature of science: Past, present, and future. In: S.K. Abell & N.G. Lederman (Eds.), *Handbook of Research on Science Education*. (pp. 831-880). Lawrence Erlbaum Associates.
- Louca, L. T. & Zacharia, Z. C. (2015). Examining learning through modeling in K-6 science education. *Journal of Science Education and Technology*, 24, 192–215. <https://doi.org/10.1007/s10956-014-9533-5>
- Malone, K. L., Schuchardt, A. M., & Sabree, Z. (2019). Models and modelling in evolution. In U. Harms, & M. Reiss (Eds.), *Evolution education re-considered* (pp. 207-226). Springer. https://doi.org/10.1007/978-3-030-14698-6_12
- McLaughlin, C. & Pringle, R. (2011). Conflict and consonance: Implementing an evidence-based middle school science curriculum. In M. S. Plakhotnik, S. M. Nielsen, & D. M. Pane (Eds.),

- Proceedings of the tenth annual College of Education & GSN Research Conference* (pp. 138-146). Florida International University.
- Mendonça, P. C. C. & Justi, R. (2011). Contributions of the model of modelling diagram to the learning of ionic bonding: Analysis of a case study. *Research in Science Education*, 41, 479–503. <https://doi.org/10.1007/s11165-010-9176-3>
- Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Open University Press.
- Museum of Science, Boston. (n.d.). Engineering design process. <https://yes.mos.org/impact/engineering-design-process/>
- National Research Council. (1996). *National science education standards*. Washington, DC: National Academy Press.
- NGSS Lead States. (2013). *Next generation science standards: For states, by States*. The National Academies Press.
- Ormrod, J. O. (2008). *Human learning* (5th ed.). Pearson.
- Park, B.-Y., Campbell, T., Kelly, M., Gray, R., Arnold, C., Chadwick, C., Cisneros, L. M., Dickson, D., Moss, D. M., Rodriguez, L., Volin, J. C., & Willig, M. R. (2023). Improving NGSS focused model-based learning curriculum through the examination of students' experiences and iterated models. *Research in Science & Technological Education*, 41(3), 983–1007. <https://doi.org/10.1080/02635143.2021.1978962>
- Saldaña, J. M. (2013). *The coding manual for qualitative researchers* (2nd ed.). SAGE Publications.
- Sands, D. (2021). Modeling as sensemaking: towards a theory of modelling in physics education. *European Journal of Physics*, 42(6), 064001. DOI: 10.1088/1361-6404/abcc80
- Schwarz, C. V. & Gwekwerere, Y. N. (2007). Using a guiding inquiry and modeling instructional framework (EIMA) to support preservice K-8 science teaching. *Science Education*, 91(1), 158-186. <https://doi.org/10.1002/sce.20177>
- Schwarz, C. V., Reiser, B. J., Davis, E. A., Kenyon, L., Acher, A., Fortus, D. et al. (2009). Developing a learning progression for scientific modeling: Making scientific modeling accessible and meaningful for learners. *Journal of Research in Science Teaching*, 46, 632-654. <https://doi.org/10.1002/tea.20311>
- Sickel, A. J., Witzig, S. B., Vanmali, B. H., & Abell, S. K. (2013). The nature of discourse throughout 5E lessons in a large enrolment college biology course. *Research in Science Education*, 43(2), 637-665. <https://doi.org/10.1007/s11165-012-9281-6>
- Stammen, A. N., Malone, K. L., & Irving, K. E. (2018). Effects of modeling instruction professional development on biology teachers' scientific reasoning skills. *Education Sciences*, 8(3), 119. <https://doi.org/10.3390/educsci8030119>
- Van Sickle, M. & Spector, B. (1996). Caring relationships in science classrooms: A symbolic interaction study. *Journal of Research in Science Teacher*, 33(4), 433-453. [https://doi.org/10.1002/\(SICI\)1098-2736\(199604\)33:4<433::AID-TEA5>3.0.CO;2-T](https://doi.org/10.1002/(SICI)1098-2736(199604)33:4<433::AID-TEA5>3.0.CO;2-T)
- Wells, M., Hestenes, D., & Swackhamer, G. (1995). A modeling method for high school physics instruction. *American Journal of Physics*, 63(7), 606-619. <https://doi.org/10.1119/1.17849>
- Wilhelm, J., Cole, M., Cohen, C., & Lindell, R. (2018). How middle level science teachers visualize and translate motion, scale, and geometric space of the Earth-Moon-Sun system with their students. *Physical Review Physics Education Research*, 14, 010150. <https://doi.org/10.1103/PhysRevPhysEducRes.14.010150>
- 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), 1-27. <https://doi.org/10.1002/sce.20259>
- Yin, R. K. (1994). *Case study research: Design and methods* (2nd ed). Sage.

Zangori, L., Peel, A., Kinslow, A., Friedrichsen, P., & Sadler, T. D. (2017). Student development of model-based reasoning about carbon cycling and climate change in a socio-scientific issues unit. *Journal of Research in Science Teaching*, 10, 1249-1273. <https://doi.org/10.1002/tea.21404>

Appendix: Teaching and Learning Sequences for Modeling

The 18 TLSs are presented in chronological order by year of publication.

Note: Eight of the TLSs are described as non-linear. Seven of those non-linear TLSs have a reference to a general order of events, either by numbering the phases or following arrows on a flowchart. For the purpose of this table, I used the general order provided but point out the authors make it clear that non-linear paths are likely. Additionally, the list of learners' actions is a summary and does not fully encapsulate the phases. It is highly recommended that readers go to the source publication to fully understand the TLS.

Author, Year, and Name of Model or Sequence	Grade-level and Science Discipline	Linear or Non-Linear; Cycle or Sequence	Presumed Order of Learners' Actions Based on Authors' Representation and Discussion
Hestenes (1987) General model development	Grade-level: not specified, example is for university Science discipline: all; example is for physics	Linear Sequence	● <i>Description Stage</i> – Students describe the phenomenon and identify variables, and may construct initial conceptual models like free-body diagrams ● <i>Formulation Stage</i> – Students consider how to develop a functional relationship between the variables ● <i>Ramification Stage</i> – Students work out special properties and implications of the mathematical model ● <i>Validation Stage</i> – Students test the model empirically
Clement (1989) Hypothesis development (model construction) cycle of conjecture, evaluation, and modification or rejection	Grade-level: not specified, example is university doctoral candidates and professors in technical fields Science discipline: all; example is for physics	Non-linear Cycle	● (A) Make Initial Observations & (B) Activate Possible Analogies and Related Model Elements ● (C) Construct Initial Model ● (D) Rational Evaluation ○ (E) Reject or Modify Model ● (F) Construct and Perform Empirical Tests ○ (E) Reject or Modify Model
Wells et al. (1995) Modeling method	Grade-level: secondary (high school) and university Science discipline: physics	Linear Cycle	● <i>Model Development</i> – Students explore phenomena and develop a model to explain it ● <i>Model Deployment</i> – Students test model against experiments or physics problems
Clement and Steinberg (2002) Role of analogies	Grade-level: not specified, example is for	Linear Cycle	● Generate Model 1 ● Evaluate Model 1 ○ Learners experience observation

and observations in GEM cycle of model generation, evaluation, and modification	secondary (high school) Science discipline: not specified, example is for physics		discrepancies and consider analogous case ○ Learners modify model to... ● Modify Model 1 into Model 2 ● Evaluate Model 2 ○ Learners experience observation discrepancies and consider analogous case ○ Learners modify model for the next step ● Modify Model 2 into Model 3 ● Evaluate Model 3 ○ Learners experience observation discrepancies and consider analogous case ○ Learners modify model for the next step [sequence continues as needed]
Justi & Gilbert (2002) and Gilbert & Justi (2016) Model of modeling framework	Grade-level: not specified Science discipline: not specified	Non-linear Cycle	● Decide on purpose ● Select source for model and/or have experience ● Produce mental model ● Express in mode(s) of Representation ● Conduct thought experiments ○ Reject or Modify Mental Model ● Design and Perform Empirical Tests ○ Reject or Modify Mental Model ● Fulfill Purpose ● Consider Scope and Limitation of Model
Buckley et al. (2004) Model-based learning framework	Grade-level: not specified, example is for secondary (high school) Science discipline: not specified, example is for biology	Non-linear Cycle	● Model formation into learners' mental models based on prior knowledge or new information ● Model use ○ Interacting with hypermodels and phenomena ● Model evaluation ○ Model reinforcement or revision of the model, or ○ Model rejection and formation of a new model
Khan (2007) GEM cycle	Grade-level: university Science discipline: chemistry	Linear Cycle	● <i>Generate hypothesis</i> – Students consider relationships between variables to develop explanatory models of molecular structure ● <i>Evaluate hypothesis</i> – Students design new tests, compare ideas, and seek new information to evaluate mental models and consider causal factors ● <i>Modify mental models</i> – Students revise and

			enrich mental models to have predictive value [cycles of GEM contribute to refinement of learners' mental models]
Halloun (2007) Modeling learning cycle (MLC)	Grade-level: secondary and university Science discipline: examples are for physics	Non-linear Cycle	● Exploration ○ <i>Monstration</i> – empirical situations are exposed to support cognitive disequilibrium ○ <i>Nominal models</i> – construction of a subsidiary model ● Model Adduction ○ <i>Plausible model</i> – time is devoted to the proposition and comparison of plausible models ○ <i>Investigative design</i> – proposes an investigation to test the model ● Model formulation ○ <i>Investigation and initial model formulation</i> – learners collaboratively conduct the investigation and refine the plausible model ○ <i>Rational model extrapolation</i> – learners rationally analyze the model and may reference previous plausible models ● Model deployment ○ <i>Elementary deployment</i> – learners deploy the constructed model in parts ○ <i>Paradigmatic deployment</i> – learners deploy the model with its integrity ● <i>Paradigmatic synthesis</i> – learners determine viability of the model and either return to previous phases for revision or begin a new cycle
Schwarz & Gwekwerere (2007) EIMA instructional framework	Grade-level: Elementary (used by elementary pre-service teachers) Science discipline: all	Linear Sequence	● <i>Engage</i> – students engage in the topic and draw on prior knowledge ● <i>Investigate</i> – students investigate the topic or phenomenon through data collection and analysis ● <i>Model</i> – students create models (focusing on causal aspects of patterns in phenomena) or explanations ● <i>Apply</i> – students apply models/explanations to different situations
Windschitl et al. (2008) Model-based inquiry	Grade-level: secondary (middle and high school) Science discipline: all	Non-linear Cycle	● <i>Conversation 1 – Organizing what we know and what we want to know.</i> Students consider topic under study and think of ways to represent current understandings ● <i>Conversation 2 – Generating testable hypotheses.</i> Students use model to articulate sets of

			relationships or events to test the model ● <i>Conversation 3 – Seeking evidence.</i> Students consider what data would help test their models ● <i>Conversation 4 – Constructing an argument.</i> Students consider whether original model is consistent with the data, how the data helps them make theoretical inferences, what the model does not predict, whether the model should change, and how to apply their models to new phenomena
Schwarz et al (2009) Instructional modeling sequence for elementary curriculum materials	Grade-level: elementary and middle school Science discipline: not specified, examples are for physics and chemistry	Linear Sequence	● <i>Anchoring phenomena</i> – consider driving questions and phenomena for a concept. ● Construct a model ● Empirically test the model ● Evaluate the model ● Test the model against other ideas ● Revise the model ● Use the model to predict and explain
Baek et al. (2011) Model-centered instructional sequence	Grade-level: elementary Science discipline: all, example is for physical science	Linear Sequence	● <i>Anchoring phenomena and central questions</i> – phenomena from everyday life are introduced ● Constructing an initial model ● <i>Empirical investigations</i> – work in groups to complete a set of investigations. Learners make and share predictions and revise models based on evidence. ● <i>Scientific ideas and computer simulations</i> – scientific ideas and theories are introduced via texts, the teacher, and simulations. ● Evaluating and revising the model ● <i>Peer evaluation</i> – present models in small groups, give and receive feedback, and discuss criteria for evaluation ● Constructing a consensus model ● Using the model to predict or explain related phenomena
Louca & Zacharia (2015) Paraphrased title: Four discrete phases of modeling-based learning	Grade-level: not specified, example is for elementary Science discipline: not specified, example is for physical science and life science	Linear Sequence	● <i>Collection of observations and experiences phase</i> – learners examine a natural world phenomenon, to identify its parts or predict/explain its behavior ● <i>Construction of the model phase</i> – learners construct an initial model. They consider how to represent concepts symbolically and how well they represent the components or features. The model is modified and enhanced.

			● <i>Evaluation of the model phase</i> – learners apply their models to new contexts, testing one component at a time, and receiving feedback. ● <i>Revision of the model phase</i> – learners may reject the model and start over, or may revise or elaborate the model.
Gilbert & Justi (2016) The Model of Modelling v2	Grade-level: not specified Science discipline: not specified	Non-linear Cycle	Four equidistant stages, with no one entry point or sequence from any one stage to another ● Expression with sub-processes (e.g. select modes of representation) ● Creation with sub-processes (e.g. make analogy, get new empirical data) ● Test with sub-processes (e.g. design test, interpret data, generate explanation) ● Evaluation with sub-processes (e.g. define scope of model)
Stammen et al. (2019) Modeling cycle Note: this model is cross-referenced to Malone et al. (2019), but it is explained in Stammen et al. (2019). Malone is an author of both publications.	Grade-level: not specified, example is for secondary Science discipline: not specified, example is for biology	Linear Cycle	● <i>Phenomena</i> – learners encounter phenomena ● <i>Model with Representations</i> – learners develop a model ○ Develop model based on data ○ Representations may be algebraic, verbal, diagrammatic, verbal, graphical, or any combination of the four ● <i>Correspondence with Phenomena</i> – learners use model to make predictions based on correspondence with phenomena, and continues the cycle by making revisions as necessary
Sands (2021) ACME Protocol	Grade-level: university Science discipline: physics	Linear Sequence	● <i>Assess the problem</i> – develop a mental model by identifying properties of objects ● <i>Construct the model</i> – translate mental model into a mathematical structure ● <i>Evaluate the model</i> – run the model, produce an output, and compare to known data if possible
Chen (2022) Variations of epistemic uncertainty in the four phases of modeling	Grade-level: not specified, example is for elementary Science discipline: not specified, example is for life science	Non-linear Sequence	● <i>Phase 1, Problematising a phenomenon</i> – learners work to make sense of a phenomenon ● <i>Phase 2, Material practice</i> – learners develop mental and physical models to explain a phenomenon ● <i>Phase 3, Argumentative practice</i> – learners debate and evaluate their models based on evidence ● <i>Phase 4, Conceptualizing theory and application of final models</i> – learners conceptualize a final

			model and connect it to scientific theory
Garrido & Couso (2025) Instruction Performance Modeling (IMP) Cycle	Grade-level: primary and secondary, example is with preservice teachers Science discipline: not specified, example is for physical science	Non-linear Cycle	● Recognize the need of a model ● Use and express initial model ● Evaluate model ● Revise model ● Express consensus model ● Use consensus model

Understanding How an Environmental Educator Identifies and Thinks about Environmental Issues in Their Education Programming

Hamza Malik 

Lloyd Center for the Environment & Stonehill College

Stephen B. Witzig 

University of Massachusetts Dartmouth

ABSTRACT

Environmental educators work directly with students and their teachers to implement environmental education programs in both in-school and out-of-school settings. However, the literature is limited in focusing on these environmental educators' thinking and teaching perspectives on environmental issues. This narrative inquiry focuses on an educator (Greta – pseudonym) from a local environmental education center to understand the ways in which she thinks about and identifies the relevance of environmental issues, utilizing in-school and out-of-school settings. Our findings indicate that Greta thinks about environmental issues using her past and current experiences. Greta uses a more local and personalized approach with students, utilizing social and cultural aspects to identify the relevance of environmental issues. This study contributes to research exploring environmental educators' thinking about environmental issues and their different teaching perspectives when interacting with students and when teaching environmental education content in out-of-school and in-school contexts. This study has implications for both practice and future research in environmental issues.

Keywords: environmental issues, relevance of environmental issues, environmental educators, thinking, experiences, education programs

Key Abbreviations: Environmental Education (EE), Environmental Education Center (EEC)

Introduction

Communities around the globe are being directly and indirectly affected by environmental issues (e.g., deforestation, rising sea levels, biodiversity loss, plastic pollution). Scientists and policymakers have warned us about these environmental issues for years (Kelani, 2015; Klein, 2015). These problems are dangerously impacting and threatening the existence of the ecosystem(s) around us. Bigelow and Swinchart (2014) argued that we live in a world where we cannot separate ourselves from the environment. As a society, we are all connected and share a common planet. Harper and Snowden's (2017) book highlights human connections to environmental issues as a primary topic for people interested in teaching and learning about environmental issues. Their chapters aim to give readers a snapshot of how these human connections are made between different fields. To achieve equilibrium in our society, where natural resources and human impacts are balanced, we need to build a sustainable future for ourselves, especially by focusing on educating people about environmental issues.

Environmental educators are at the forefront of teaching students about the environment and environmental issues (Malik et al., 2024a). The work we present in this paper highlights the importance of understanding how environmental educators think about and identify environmental issues in their teaching. Environmental educators are perceived as science content experts by both students and teachers (Stronach et al., 2026). These educators are actively involved with students in various in- and out-of-school programs throughout the year. These different teaching and learning contexts (in-school vs. out-of-school) demand careful consideration by the environmental educator regarding the teaching strategies needed to help students (and teachers) learn environmental science concepts (Hesson & Roth, 2024). For this research study, we were specifically interested in learning how these educators present environmental issues in their programs and their underlying thinking related to environmental issues. We hope our findings will help school administrators, teachers, environmental education team members, and science teacher educators/researchers better understand environmental educators' perspectives on environmental issues and the support they need to be productive and effective in their teaching experiences.

Seminal works by Corcoran (1999), Palmer (1993), and Tanner (1980) provide an essential foundation for environmental educators' overall experiences and understanding. For example, Corcoran (1999) noted various influences that significantly affect and inform the lived experiences of environmental educators in the United States. They focused on unpacking environmental educators' narratives and their experiences. However, these studies do not provide evidence of how environmental educators think about and present environmental issues in their programs, across both in- and out-of-school settings. Exploring this area has been called for in recent research (Rennie, 2014; Stronach et al., 2026).

It is important to note that the *International Handbook of Research on Environmental Education* (Stevenson et al., 2013) includes 51 chapters on environmental education research. Some of these chapters have environmental educators' perspectives embedded in their content. However, there is no single chapter focused solely on environmental educators. Having said that, we would like to note that all the chapters in this handbook are crucial and offer valuable insights into environmental education research. What the field needs, though, are more in-depth accounts of the pedagogical decisions made by environmental educators when teaching science to students in both in-school and out-of-school settings.

Environmental educators are directly responsible for enhancing students' understanding and knowledge about environmental issues (Malik et al., 2025). There is a need to engage the public in experiences that make them more literate about scientific knowledge and understanding (American Association for the Advancement of Science, 2022). We believe that one way to enhance the public's scientific knowledge and understanding is through education, which takes place in both in-school and out-of-school settings. In this study, we focus on identifying environmental educators as critical community members tasked to teach our students, teachers, and the general public about environmental issues. Their work is of tremendous importance for the health and well-being of our communities. These educators utilize both in-school and out-of-school settings for their programs. This research study highlights environmental educators' perspectives to unpack how they think about environmental issues and identify their relevance when teaching in various contexts.

Environmental Education (EE) is a challenging field within science education due to its interdisciplinary nature. The goal or objective is not easily visible as it is intertwined mainly with social and cultural aspects (Dillon & Herman, 2023). Environmental issues, by nature, are related to a society's social and cultural perspective. For example, the water crisis in Flint, Michigan (The Associated Press, 2021) is not an issue affecting one individual. Instead, it is affecting the whole community, with implications for more significant systemic problems (environmental, political, and infrastructure) that must be addressed. So, the approach to teaching students about such environmental issues must consider cultural and social factors. While there are journals that focus

specifically on environmental education (*Journal of Environmental Education*, *Journal of Outdoor and Environmental Education*, *Environmental Education*, etc.) interdisciplinary journals like the *Electronic Journal of Research in Science and Mathematics Education* have begun to include a foci on issues related to ‘informal science and environmental education’ to reach a wider audience (Hesson & Roth, 2024; Stevenson et al., 2014). Ng and Leng (2022) noted that people who are actively engaged with nature tend to exhibit pro-environmental behaviors. Environmental educators actively engage students with nature and their surroundings (Malik et al., 2024a).

Building on Ng and colleagues (2022), we emphasize that environmental educators operate in teaching spaces that regularly connect students with nature, influencing them to be more pro-environment in their thinking and experiences. Zeiske et al. (2021) stated, “As children are the future generations that would need to adopt large-scale behavioral changes to mitigate climate change, understanding which factors encourage children to engage in pro-environmental behavior is pivotal to reduce environmental problems” (p. 1119). One of the core responsibilities of environmental educators is to teach and design various learning experiences for students that influence their pro-environmental behaviors and their understanding of the underlying science behind the issues. Athman and Monroe (2001) argued that environmental educators “can help foster learners’ innate curiosity and enthusiasm, providing them with continuing opportunities to explore their environment and engaging them in the direct discovery of the world around them” (p. 37). In this study, we present empirical evidence to inform decisions about advancing EE as a field and to better understand environmental educators’ thinking and perspectives.

Few studies explore how an educator from an Environmental Education Center (EEC) perceives and identifies the relevance of environmental issues. Most studies discuss EE and EEC as a field (Braund & Resis, 2006; Casas et al., 2021; Christenson, 2004; Frensley et al., 2022; Haluza-DeLay, 2013; Henderson & Zarger, 2017; Yanniris, 2015). These studies do not address the perspective of an environmental educator. Environmental educators are still developing their place within the science education field; hence, there is limited knowledge about how they operate and identify relevant environmental issues. The field needs in-depth accounts from environmental educators who navigate the unique teaching and learning space that is sometimes formal (in-school context), sometimes informal (out-of-school context), and occasionally interformal (contains elements of both the in-school and out-of-school contexts) when teaching school children and the public about significant and complex environmental issues (Eshach, 2017; Rennie, 2014; Stronach et al., 2026). For this research study, we utilized narrative inquiry to understand the life stories of an EEC educator who plays a dual role in providing field trips, nature walks, summer programs, and after-school programs to students in an out-of-school setting (informal science education context) and also partners with local schools as a science content specialist that includes routine visits to the schools to conduct activities and instruction on EE (formal science education context). This qualitative research study aimed to understand how an educator from an EEC thinks about and identifies the relevance of environmental issues, and to explore the following research question: in what ways does an educator from an environmental education center think about and identify the relevance of environmental issues?

Conceptual Framework and Review of Relevant Literature

Theoretical lens

As science education researchers aiming to understand the perspective of an environmental educator, we relied on a theory rooted in social constructivism (Hess, 1995). Social constructivists believe that social interaction and how people interact with their surroundings affect their learning experience. Environmental educators are vital community members, and diverse social, cultural, and

political experiences shape their understanding and perspectives. Richardson (2003) defines social constructivism as an approach that “centers on how power, the economy, political and social factors affect how groups of people form understandings and formal knowledge about their world” (p. 1624). This argument connects to the social and cultural roots of environmental education as a field. Environmental issues are generally politically charged and socially embedded due to their global impact.

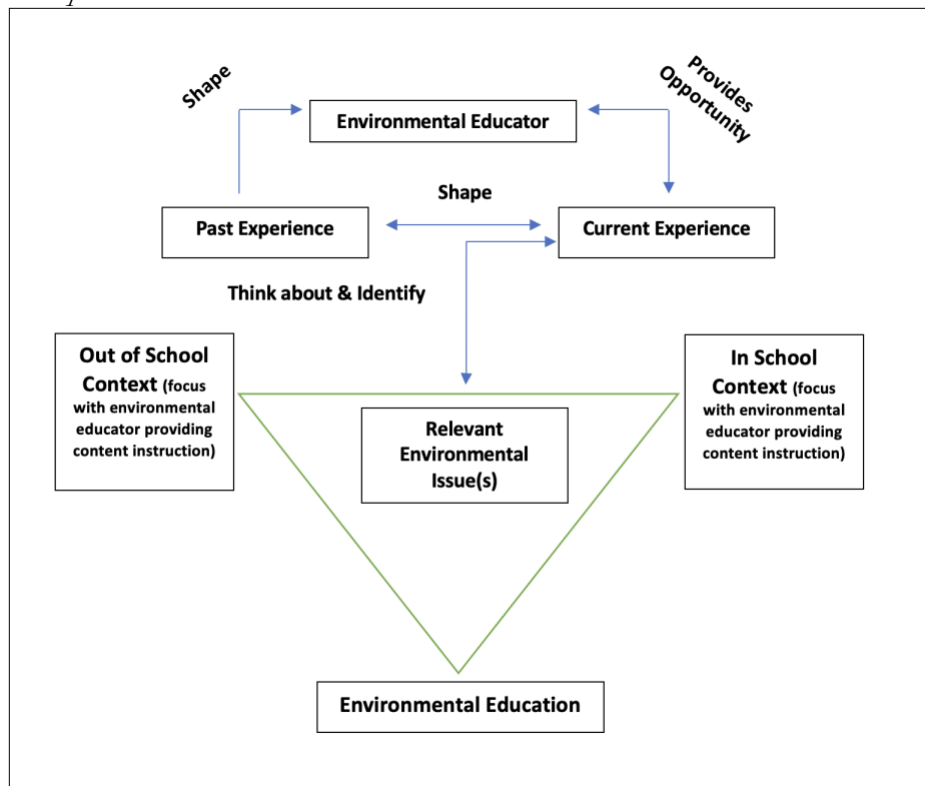
Educators are at the forefront of driving change in student learning, or at least providing opportunities to motivate students to think outside the box to address problems such as environmental issues (Malik et al., 2024a). Aronowitz and Giroux (2003) stated that human beings socially construct knowledge as they interact with their surroundings, which influences their understanding and that of others with whom they interact regularly. Therefore, it is safe to imply that there are various ways an educator develops their own opinions and influences students, particularly when linking instruction with environmental issues. Thus, identifying and unpacking the specific methods environmental educators use in their teaching and thinking about environmental issues is essential to determine what our students are learning, how they are learning, and which social and cultural perspectives influence environmental educators' teaching.

Substantive Constructs

The substantive constructs that comprise the conceptual framework and review of relevant literature for this research study include 1) Experiences, 2) In-school/Out-of-school contexts, 3) Environmental Education, and 4) Relevant Environmental Issues (see Figure 1). These constructs, viewed through the social constructivist lens, helped frame this study.

Figure 1

Conceptual Framework



Note. The conceptual framework illustrates how past and current experiences shape and provide opportunities for educators to develop their thinking about environmental issues. Their experiences can include anything they have learned or experienced throughout their life. They use these experiences to think about and identify their understanding of relevant environmental issues. Once they have developed their thinking about environmental issues, they can use an out-of-school context, an in-school context, or a combination of both, to identify the relevance of these issues to students' environmental education.

Experiences

An educator from an EEC considers environmental issues through their life experiences, encompassing both past and present aspects (see Figure 1). Au (1998) stated, "Because reality is seen to be created through processes of social exchange, historically situated, social constructivists are interested in the collective generation of meaning among people" (p. 299). These social, historical, and cultural aspects highlight an environmental educator's past and present experiences. Combined or individually, these life experiences (past and present) inform the educator's thinking about environmental issues. Barak (2017) argued that social constructivists believe cognitive change is enhanced by social dynamics and subconscious discourse. Taken together, our framework highlights how educators' social interactions, including past and current experiences, inform their thinking about problems like environmental issues. These educators focus on how individuals interact and emphasize taking a bigger picture of society to develop their understanding of the given topic, which in this research study is environmental issues.

In-school and Out-of-school Contexts

An environmental educator from an EEC identifies environmental issues using past and current experiences (see Figure 1). These educators often use both in-school and out-of-school contexts (Eshach, 2017; Malik et al., 2024a; Stronach et al., 2026) to provide students with opportunities to learn about environmental issues while drawing on social and cultural perspectives in their instruction. Falk and Dierking (2010) stated that people who utilize both in-school and out-of-school learning experiences enjoy science-related topics. In a school setting, these educators are seen as science content experts and provide students with opportunities to enhance their science learning. Dillon and Herman (2023) argued that science learning should involve processes beyond traditional classroom and lab settings. In the out-of-school context, these educators can interact with students to provide a practical, hands-on approach to the same science concepts students encounter in their schools. Rennie (2014) stated that "... out-of-school contexts usually involve a degree of volition, so individuals' unique backgrounds and experiences and the choices they make about the nature of their activities can bring about very different learning outcomes" (p.120). Furthermore, Schmäing and Grotjohann (2022) describe out-of-school educators as individuals who teach students about environmental issues at out-of-school locations. Together, this dynamic connection between in-school and out-of-school contexts makes it vital for educators to identify the relevance of environmental issues for students. However, we need more voices in the literature from environmental educators who navigate the complex landscape of teaching science both in and out of schools.

Environmental Education

EE emerged as a field in the 1960s when Rachel Carson published an inspiring book titled *Silent Spring* that highlighted the significance of environmental issues (Sisson, 2018). Shortly after that, the UN Tbilisi Declaration of 1977 became an essential and pivotal convention by documenting an

international trajectory of EE (Hungerford & Volk, 1990). These events shaped the field of environmental education, which continues to develop today. Organizations such as the *North American Association for Environmental Education* (NAAEE) have been working for the last five decades to advance the EE field while facilitating various educators from formal and informal contexts. There are environmental education core concepts embedded in the *Next Generation Science Standards* (NGSS Lead States, 2013) shaping science instruction in K-12 schools, there are textbooks devoted to teaching and learning environmental education at the college level (Ayers, 2017; Chiras, 2009; Robertson, 2021), the NAAEE provides an international emphasis on the teaching and learning of environmental education for both K-16 schools as well as informal science/environmental education contexts (Simmons et al., 2025). Furthermore, international organizations such as the World Wide Fund for Nature (WWF) help shape environmental education in K-16 contexts worldwide (Fien, 2002). These resources can significantly shape an educator's perspective when implementing EE programs and finalizing teaching strategies. This unique nexus of in- and out-of-school settings makes EE a dynamic, socially constructed field of science education (see Figure 1).

Relevant Environmental Issues

One core objective of EE is to teach students about environmental issues using both in-school and out-of-school settings (Dillon & Herman, 2023; Rennie, 2014). Environmental educators teach students about these issues, but how they identify relevant environmental issues remains a question that is not fully understood or explored in the science education and EE literature. We argue that understanding how environmental educators think about identifying relevant environmental issues can be approached by focusing on their narratives. Suppose we want to understand how environmental educators include relevant environmental issues in their teaching practice. In that case, we need to observe them teaching and then talk with them about their pedagogical decisions when working with students. We need to understand environmental educators' knowledge of environmental issues and their causes (Kollmuss & Agyeman, 2002), the scale of the environmental issue ranging from global issues such as anthropogenic climate change (Robinson & Bowen, 2000) to point source pollution such as the degradation of water quality at a local river (Geary & Lucas, 2019; We & Chen, 2013), and we need to understand how the environmental educators transform their understanding of these environmental issues to make them relevant to their students (Malik et al., 2024b). An environmental educator's social and cultural aspects are critical for understanding how they identify relevance, and how they connect it to our conceptual framework embedded in past and current experiences (see Figure 1).

Methods

Methodology: Narrative Inquiry

We used narrative inquiry with a social constructivist lens to investigate the following research question: In what ways does an educator from an environmental education center think about and identify the relevance of environmental issues? Savin-Baden and Niekirk (2007) define narrative inquiry as “an approach that focuses on the use of stories as data” (p. 459). They further argue that narrative inquiry is an approach that goes beyond mere storytelling; instead, it addresses the fundamental concepts that underlie participants' life experiences, shaping their thinking on a given topic. So, our role was to focus on educators' narrative stories to understand their thinking and identify relevant environmental issues.

Participant and Context

One environmental educator, Greta (pseudonym), was selected for this in-depth inquiry from a pool of seven education team members due to her fourteen years of experience working as an educator at the Lloyd Center for the Environment (LCE). Greta's experience includes working with students in both in-school and out-of-school settings. The methodological decision to focus this investigation on a single participant was intentional, as it aligns with our inquiry approach, which uses narrative inquiry (Savin-Baden & Niekerk, 2007). This approach affords us the opportunity to provide an in-depth analysis of our participant's experiences over the years she has spent teaching EE across various contexts during her employment at the LCE. LCE is a Southeastern Massachusetts non-profit organization whose mission includes environmental research, education, and outreach. The center is on 82 acres of land that consists of a maritime forest, salt marshes, an estuary, an outdoor learning pavilion, and two indoor classrooms with various live exhibits (e.g., fish, touch tanks), where they host programming for students of all ages. In addition, the education staff serves as science content experts through sustained formal in-school programming in multiple districts throughout the school year (Malik et al., 2024b). A view of one of the teaching and learning spaces on the LCE campus is shown in Figure 2.

Figure 2

Students are biking at the Lloyd Center for the Environment's Property.



Additionally, LCE offers summer programs annually for students ages six to twelve. From 2017 to 2021, LCE offered 35 overarching education programs catering to approximately 47,250 students of all ages from several major cities and towns in Massachusetts. More relevant information about Greta and LCE is embedded in the findings section for the readers.

Positionality, Bias, Limitations

The first author of this study has five years of international experience with the World Wide Fund for Nature in both Pakistan and the Fiji Islands. During this research, he conducted a capstone research project as part of his doctoral program in science education research. He then participated in a research and teaching fellowship at the LCE, ultimately being hired as its research director after successfully defending his dissertation. During that time, this research project was revisited and adapted for publication. The second author is an associate professor of science education who has established a 12-year working relationship with the LCE, with much of his research agenda focused on teaching and learning in environmental education within the LCE's programs. Together, we are uniquely positioned to conduct this in-depth narrative analysis of Greta's story as an environmental educator. As qualitative researchers, we embrace our bias towards this work and believe it provides us with unique insight into telling Greta's story in ways that would otherwise not be possible (Creswell, 2013; Creswell & Poth, 2016; Maxwell, 2008). While this might be a limitation to this research, and while one could argue that just including one participant would also be a limitation, we view these as strengths, as we were able to explore Greta's story in-depth over time and space, as is typical in narrative inquiry (Savin-Baden & Niekerk, 2007).

Data Collection

The data for this research was collected in Spring 2020 and include document analysis, field observations, and interviews. Initially, document analysis was conducted, including an in-depth review of the LCE's website and program fact sheets, to develop an understanding of the center and its activities and programs. Additionally, two semi-structured narrative interviews, following the guidelines of Savin-Baden and Niekerk (2007) and Creswell (2013), were conducted as primary data sources. Each interview lasted between 60 and 90 minutes. We also observed six education programs (three in school and three out of school) in which Greta served as an instructor. We observed four programs (two in school and two out of school) between the first and second interview, whereas two observation sessions (one in school and one out of school) happened after the final interview. We intentionally had a two-week window between interviews to analyze the first interview data, observations, and documents, and use them to inform the second interview. In total, we produced 51 open-ended questions in three sections: Participant Background, Lloyd Center for Environment, and Context. We used our conceptual framework as a guide, but remained flexible enough to follow the story Greta shared, prompting her to reflect on her teaching and learning about environmental issues. Throughout the interviews, documents were collected (and subsequently analyzed) from Greta, including images, student work, observations, and her instructional materials.

Data Analysis

Interview data were transcribed verbatim. The analysis method was adapted from Creswell (2013), which included data sorting in chronological order for stories typical of narrative inquiry (Savin-Baden & Niekerk, 2007) and developing themes. The analysis helped us understand themes and their connections to Greta's past and current experiences. Using the constant comparative method, we employed an interactive inductive-deductive data analysis model (Strauss & Corbin, 1998). Our deductive analysis started with the constructs in our conceptual framework, and continued as new constructs and themes emerged from our inductive analysis. Our document analysis informed interviews and observations, which, in turn, informed additional document analysis, ultimately leading to our final interview with Greta. The participant's narrative was then

linked to the underlying experiences, and assertions were developed in response to our research question.

Findings

The findings are organized into four sections that include emergent themes related to the constructs in our study: Experiences - Past and Current, Nexus of In-School and Out-of-School Context, Environmental Education from a Local Perspective, and Relevant Environmental Issues as a Local Problem. Each section includes an embedded assertion, supported by excerpts from Greta's narrative.

Experiences - Past and Current

Our first assertion posits that Greta's thinking about environmental issues is developed by her past experiences and is simultaneously matured and informed by her current experiences. These experiences inform Greta's environmental behaviors regarding her thinking about and identification of the relevance of environmental issues. Greta decided to join the LCE because she was motivated by her past experiences working with children and her love for animals. She explained, "*I have always enjoyed working with kids, and I have always enjoyed working with animals. And so, I figured, you know, I would like to learn how to teach.*" Greta explained that her early life experiences helped her choose the profession of environmental education. She stated:

The education piece to me kind of came more as I started doing it more with the [Lloyd] Center. There was a little bit before the center when I was volunteering for the Cape project stranding network. So, other than that, I did not have any teaching background to refer to. But I did a lot of social things growing up. I was in clubs and things like that.

The experiences that drove her to become an educator included volunteering in EE, working with social clubs, and teaching young students. Greta developed most of the formal knowledge and thinking about environmental issues after starting to work at the LCE. This includes part of her current/present experience. She explained,

So, the education part of it for me, I guess, has been slowly cultivating over the years, and so has the environmental part. The [Lloyd] Center really took all of that and kind of put it into focus for me of what I wanted to do with it.

Her past experiences led her to decide to become an educator. However, her current expertise at the LCE has informed and developed her thinking about environmental issues. We asked Greta to explain how she thinks about her past and current experience, informing her thinking as an environmental educator; she replied, "*I would say it's definitely a mixture of both for me. The more I teach, the older I get, the more I understand about what is going on. That's what I try to bring to my teaching as well.*" These experiences (past and current) combined contribute to her thinking about environmental issues and inform how she identifies the relevance while teaching.

Nexus of In-School and Out-of-School Context

Our second assertion states that Greta works at the nexus of in-school and out-of-school contexts, promoting critical student discussions and hands-on activities to identify relevant environmental issues. We first support this assertion with evidence from Greta's in-school teaching experiences, then follow this up with Greta's out-of-school teaching experiences at the LCE.

In-School Settings

In the in-school context, our goal was to unpack stories about Greta's strategies for engaging students during school visits. She explained, *"I try to be a little more engaging in school because they don't have all of those natural resources around them, yet we bring animals into the school and things like that, but you want to capture their attention."* For LCE, the content development of the education programs aligns with the 2016 Massachusetts Curriculum Framework for Science and Technology/Engineering, which is based on and adapted from the Next Generation Science Standards (NGSS Lead States, 2013). Greta's teaching experiences are embedded in the local schools, as the LCE has contracts to provide environmental education throughout the school year. Therefore, aligned with assertion 1, Greta's experiences with instructional materials developed by the LCE and her experiences (past and current) in nature and at the LCE inform her teaching in school contexts. For example, LCE's Erosion program is conducted in an in-school setting. It is designed to be a three-part, classroom-based, hands-on experience investigating how erosion and weathering occur. Within this program, LCE engages students through an activity that includes a student worksheet on ocean waves. Students are divided into small groups (2-4 students per group) and instructed to simulate a wave on a sloped plastic tray filled with sand (a model beach).

Our observations highlighted that Greta's role as an environmental educator involves bringing materials to set up a model beach in school settings to capture students' attention and teach them about the environmental issues related to the topic. Our document analysis found that students are then asked to observe what happens to the sand after the waves subside (see Figure 3). This is one of the critical examples in which educators from LCE focus on tapping into students' social and cultural understandings to address issues such as erosion scenarios. Greta focuses on critical student discussion when conducting in-school programming. For example, the Wave activity is designed to spark conversations with students, encouraging them to self-reflect on their roles within an ecosystem and draw on their knowledge of local shorelines.

Figure 3

The erosion program activity sample worksheet from an Avengers (Pseudonym) student group.

Before test:
Shade in the boxes where the soil appears.

Water												

How many boxes have soil? 6

After test:
Shade in the boxes where the soil appears.

Water												

How many boxes have soil now? 12 1/2

What happened? Erosion & Wethering worked together.

A discussion follows up on the activity on possible solutions to the erosion-related issues. Students experiment with different materials/applications to mitigate the impact of weathering and erosion caused by natural events. Students are engaged in discussions about these topics with the educators to identify not only the science underlying these issues but also to understand how the interaction of these social and cultural settings can help in finding better design solutions to issues like

natural disasters, such as earthquakes, blizzards, coastal flooding, etc., connected to weathering and erosion.

From 2017 to 2021, LCE completed the Erosion program with 1,022 students from across Southeastern Massachusetts. Other LCE programs provide students with different hands-on experiences in an in-school setting, such as bringing various species (e.g., owls, turtles, salamanders, crabs) into the classroom to teach various science topics. Introducing these species in the school sparks discussions among students and high participation (Malik et al., 2024a). For Greta, these stories and interactions with students make teaching the students about environmental issues easier. She believes that teaching environmental content without connecting students socially and culturally will not be impactful. She added,

I think the classroom experience is a very good introduction for students. Sometimes, it's all we get to do. Sometimes we don't get to take kids into the field, so we try to bring as much of the enthusiasm into the classroom and make it fun as well as educational for them.

This includes adding interactive presentations for the kids that involve examples from the local/surrounding areas. For example, one of the LCE programs is called Delphi the Whale, a hands-on classroom experience designed to introduce students to the diversity of marine mammals and their unique adaptations to life in the ocean. Students work alongside LCE educators as they assemble a life-sized wooden model of a pilot whale. Instructors like Greta provide whale artifacts, such as baleen and teeth, for students to examine up close. Using a measuring tape, students experience the length of a blue whale. By doing so, Greta makes it easier for students to understand and learn about local species in their classrooms, with a hands-on perspective.

Out-of-School Settings

To explain her understanding and teaching perspective regarding the out-of-school context, Greta mentioned, *"When you get them out in the field. You kind of let nature do that work for you. You're just there to kind of guide them and direct them."* She explained that utilizing such context allows the students to explore independently and find connections with the environment around them. She added, *"I think the trails and grounds are very important. It gives people a unique view into the environment and a safe place."* This also serves as a learning opportunity for students to work towards safeguarding the natural resources around them. Talking about education activities at the center, she added, *"So, having our teaching pavilion is a really important foundation in the forest, especially for kids who want to touch and look and get distracted and pick things up and do this."* We observed that the teaching pavilion (see Figure 4) overlooks the brackish water body, nested among the maritime forest and surrounded by old stone walls, highlighting the remains of an old farm trail.

This 600-square-foot open-air structure provides a dedicated space for students to engage in hands-on learning. The structure and wooden frame were designed to blend into the center's natural environment. Guiding students to make a connection with the ecosystem that is close to their house or school area. She explained,

We get kids who have never before even looked at the ocean. I mean, they live in [neighboring small city], but they have never actually gone to the beach. And they just look at you, and they go, 'This is the best field trip of my life'.

Figure 4

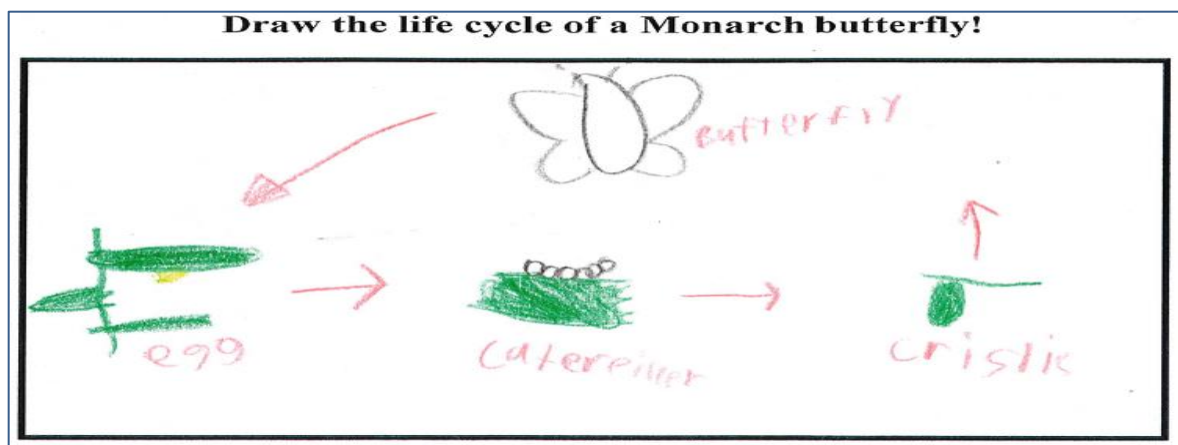
The outdoor learning pavilion at the Lloyd Center for the Environment.



Field trips like these have a lifelong impact on students' learning about environmental issues, providing them with a new perspective on their environment. It helps them understand that their local environment is connected to them and can directly impact their lives and those of all the species around them. For example, one of the LCE's programs that focuses on the out-of-school setting is Monarch Migration. This program is designed to cater to different grade levels. The program introduces students to a butterfly's life cycle, adaptations, identification, and habitats. Students in this program learn about the human impacts on butterflies and that one generation of the monarch butterfly travels a particularly great distance during its life cycle. Our document analysis highlighted that students are engaged in understanding that insects and flowers rely on each other to survive, how an increase or decrease in resources can significantly affect a species, i.e., the monarch population, and how impacts on a monarch's environment can impact their growth. An example of a student activity sheet utilized within this program is shared below,

Figure 5

Monarch Migration activity sample worksheet from a student named Camilla (Pseudonym).



Greta stated that she engages students through follow-up discussions, prompting them to reflect critically on their role as humans who influence the Monarch's life cycle. From 2017 to 2021, LCE completed the Monarch Migration program with 549 students from across Southeastern Massachusetts.

Environmental education operates at the nexus of both in-school and out-of-school settings, as shown in Figure 1. Greta uses these settings to engage students in discussions and highlight relevant environmental issues. For Greta, these discussions and activities are informed by her past and current experiences.

Environmental Education in a Local Perspective

Greta defines EE as a field that teaches students about environmental issues. Our third assertion states that Greta frames EE with a local perspective when connecting with her students. Greta explained,

We teach kids about plants - that they are living things too. We need plants, but a lot of people don't look at it that way. If something's killing off a certain plant, they're like, 'ah, it's a plant,' but it's a very big part of nature, we make that connection.

The idea here is to provide students with a localized approach connected to their surroundings. Suppose kids are struggling to understanding how a local plant affects the global environment's health. In that case, students can see how what might seem like a small or insignificant plant is vital in balancing the larger environment. Similarly, suppose students can see how a small oil leak from their family's car makes its way to the storm drain in their neighborhoods. In that case, they can connect this to how an oil tanker spill in the Atlantic Ocean will affect shorelines and marine pollution problems in the neighboring shipping/fishing harbor. We observed (see Figure 6) students making those connections during their field trip to a local beach.

Figure 6

Students visit a local beach with the Lloyd Center for the Environment's education team.



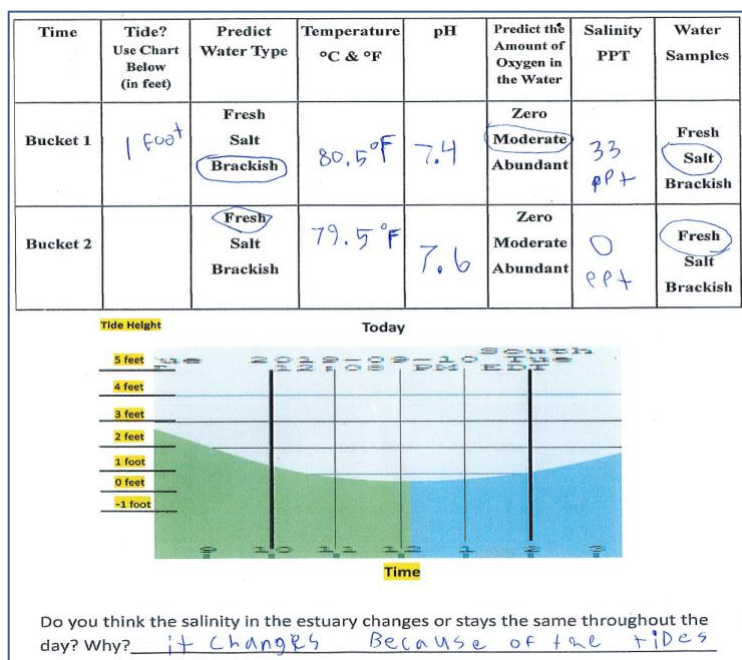
If they can make those local-to-global connections, then EE objectives would be considered a success. Greta explained that the LCE aims to teach students about environmental issues. She added,

Our goal is to educate, inform, and cause a desire to want to help when it comes to nature and the environment. We hope that by not only educating but getting people to understand, they like and appreciate the world around them.

As part of her EE approach, Greta tends to make these problems more personalized and localized for the students. Greta uses examples that align with the students' social and cultural contexts. She explained, *"I am not a global picture person. I look at small local things because that's what I can do. So, when I talked to the kids, I tend to focus on a smaller scale, the more local issues."* A signature EE program at the LCE is Turn The Tide, a three-part, hands-on field- and classroom-based experience designed for upper elementary students. Students participate in a hands-on field experience and two classroom visits. During the program, the field trip section has students explore four water-related stations on the LCE's property. Students engage in weather observations and water-quality testing, explore the biodiversity of the salt marsh, examine the many species of seaweed as indicators of ecosystem health, and participate in dip-net fishing in the estuary. During classroom visits, students learn about the types and amounts of water on Earth, as well as the adaptations of animals living in aquatic environments. Below, our document analysis shows the sample student work in this program (see Figure 7). Students are prompted to think about how a salt marsh or an estuary is connected and the scientific concept of high/low tides in the ocean.

Figure 7

Turn the Tide activity sample worksheet from a student, Camilla (Pseudonym).



The focus of this program is to teach fundamental science concepts so that these students can make visible connections that environmental issues are local problems. Greta explains to students that climate change is a local issue due to rising sea levels and how it affects our local communities. The program highlights local watershed issues to make this relevant to students' lives and experiences. For example, high nitrogen pollution is a significant problem in the estuary along LCE's border. The students learn that the primary source of this pollution is failing septic systems in the watershed, where some of these students may live. From 2017 to 2021, LCE completed the Turn the Tide program with 2,155 students from across Southeastern Massachusetts. Greta thinks that educating students is more impactful when these issues are something the students can personally experience in their surroundings or see.

Relevant Environmental Issues as a Local Problem/Examples

Environmental issues receive significant attention when they affect a local population. Here, we present how Greta explains the relevance of environmental issues. She explained that environmental issues for her entail "*Something that impacts the well-being of a living part of an ecosystem, whether it be plant or animal life. And that can include humans, as well as into the animal life part.*" Developing on her description of environmental issues that involve humans, animals, and society, she shared,

I like to focus on the small things, like, you can recycle, you can, you know, like, the car example, check and see if the car is dripping oil. I like to try to focus on smaller things that kids can look at and do themselves.

Our fourth assertion states that Greta uses social and cultural examples with the students to make environmental issues more relevant. For instance, Greta encourages students to recycle at home or helps them understand how oil dripping from their parents'/guardians' cars can impact the local ecosystem and the environment. Such examples challenge students to reflect on their social and

cultural identities. Engaging students in problems that incorporate social and cultural perspectives on their day-to-day lives helps Greta teach. She further explained,

Explaining the relevance or importance to them. Again, brings me back to the locality of it, what's going to impact them directly. The younger kids are kind of self-centered because, to them, the world is this big [using hand gestures to show small size]. You know their world is small. And so, relating it to things that they interact with or they are going to encounter is a really important step to getting them to care about it and understand it.

Greta explained that understanding students' perspectives is essential when teaching about environmental issues. Each student has their own experience and social/cultural perspective. If we can connect with students on their level, EE's objectives can be achieved more easily. Getting to the level of students' thinking and understanding of the world around them is critical to Greta's ways to identify environmental issues. She believes in providing students with examples that can be easily monitored on an individual basis.

Discussion and Implications

This research study examined Greta's life stories to understand the ways she thinks about and identifies the relevant environmental issues in the region. Sosu et al. (2008) explained in their study that educators would teach a topic effectively if they found motivation or personal meaning in the content. Greta's stories highlight a positive connection between her past and present experiences and her thinking about environmental issues. Greta's life stories of working with animals and doing outdoor activities as a kid informed her decision to become an environmental educator. Ng and Leung (2022) argued that people who are connected to nature tend to exhibit more pro-environmental behaviors. Greta's interest in the outdoors and her passion for teaching drove her to become an educator and join the LCE. These early childhood experiences helped her think about environmental issues from a more personal perspective, such as viewing them as local problems. Something that affects her day-to-day life.

Greta's experience in the out-of-school context involves activities such as field observations, field trips to local beaches and wetlands, and other activities. Eshach (2007) explained the importance of out-of-school learning as "scientific field trips to science center can generate a sense of wonder, interest, enthusiasm, motivation, and the future" (p. 178). Greta explained that in an outdoor context, nature takes over, and students get more engaged in the teaching. Such experiences involving out-of-school settings allow students to understand and experience how a scientist operates in a field (Schmäing & Grotjohann, 2022). However, Greta makes a significant effort to bring the outdoors into the classroom for students when leading school environmental education programs. Greta transitions seamlessly between these two spaces.

Greta shared stories about the various methods she uses in in-school programs, such as discussion-based lectures, hands-on activities, and classroom lectures, to keep students engaged and interested in environmental issues. Christenson (2004) argued that for environmental education content to be more impactful, it needs to be more explicit in connecting the educational objective to the goal of the activities. Greta's teaching methods promote collective contribution from participants to create, discuss, and improve shared science ideas and knowledge (Akgun & Sharma, 2023). Greta's local perspective approach, while setting the objectives of the education programs and activities, helps build these social and cultural connections with students' local environments, aligning directly with a social constructivist worldview (Richardson, 2003). She provides examples that connect to students' day-to-day lives. This connects with Hungerford's (2009) assertion that the educator's role in promoting critical thinking can inform positive science learning. Greta's expertise as an EE, along

with her purposeful efforts to bring the outdoors into the classroom with live animals and artifacts, aligns with recent research on this instructional space being interformal (Stronach et al., 2026).

To highlight the relevance of environmental issues, Greta focuses on educating students about local issues, mainly by using examples they can connect with, such as their homes, schools, or local areas. For example, in the Monarch Migration program, Greta's goal is to provide students with hands-on experiences while connecting them to their local environment, i.e., the local beachside area. She helps students connect these environmental issues to socio-cultural aspects. For Greta, this is EE's primary objective. Zeiske and colleagues (2021) highlight environmentally responsible student engagement as an essential factor given the seriousness and urgency of environmental problems for future generations. The positive relationship between students and Greta helps her identify the relevance of environmental issues across in- and out-of-school contexts, using personalized and localized examples, such as impacts on their local watershed areas and failing septic systems — i.e., point-source pollution factors — which align with our conceptual framework. We conclude that to achieve positive science learning, students must connect such environmental issues with their day-to-day lives. Greta engages students to look in their neighborhoods or at their homes and analyze how these social and cultural aspects affect the environment around them. For example, while teaching students about different low and high tides in the ocean, Greta connects climate change to rising sea levels. Wang et al. (2019) highlight climate change as one of our most significant environmental issues. Our findings support the idea that such connections for students are helpful because they make them realize that humans and nature are interconnected and that climate change is a reality everywhere around the globe.

Recommendations: Implications for Practice and Future Research

The findings from this research study add to the body of research about environmental educators who operate from an EEC. These educators are the people who interact with the students and teach them about core environmental issues (Malik et al., 2024c). These educators also frequently interact with schoolteachers, parents, and administrators (Malik et al., 2024b). Our findings suggest that understanding how these educators think and identify environmental issues is valuable for classroom teachers who often partner with them on classroom activities. For example, classroom teachers can use the same examples and perspectives to align their teaching. Future studies should focus on collaboration among formal classroom teachers and educators typically categorized as informal science/environmental educators, perhaps building on the work describing this space as interformal (Stronach et al., 2026). More research is needed to understand how collective resource sharing can occur among these distinct yet sometimes overlapping groups of educators. Perhaps science teacher preparation programs could also address ways classroom teachers can collaborate with and learn from local science/environmental educators. Some of this is occurring (Malik et al., 2024b; Hesson & Roth, 2024), but more research is needed in this area to address how the experiences of environmental educators can inform teaching and learning practice. Understanding environmental educators and their instructional decision-making also provides a unique example of how these teachers acquire new or update existing understandings of environmental issues.

Funding

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Hamza Malik (hmalik@lloydcenter.org) is the Director of Research Programs and Strategic Partnerships at the Lloyd Center for the Environment- a Southeastern Massachusetts non-profit organization whose mission is to instill lifelong respect and affection for nature in all people through

research, education, and outreach. Dr. Malik is currently also serving as an Adjunct Professor of Environmental Sciences and Studies at Stonehill College.

Stephen B. Witzig (switzig@umass.edu) is an associate professor of science education at the University of Massachusetts Dartmouth and the graduate program director for the Ph.D. program in STEM Education. Stephen's research focuses on the development of teachers' specialized knowledge for teaching science, scientific practices, socioscientific issues, and bridging research relationships among scientists, classroom teachers, environmental educators, and science teacher educators.

References

- Akgun, M., & Sharma, P. (2023). Exploring epistemic agency in students' problem-solving activities. *Interdisciplinary Journal of Environmental and Science Education*, 19(1), e2303. <https://doi.org/10.29333/ijese/12970>
- American Association for the Advancement of Science. (2022, August 23). *About Project 2061*. <https://www.aaas.org/programs/project-2061/about>
- Aronowitz, S., & Giroux, H. A. (2003). *Education under siege: The conservative, liberal and radical debate over schooling*. Taylor & Francis.
- Athman, J. A. & Monroe, M. C. (2001). Elements of effective environmental education programs. In A. Fedler (Ed.), *Defining best practices in boating, fishing, and stewardship education* (pp. 37-48). Recreational Boating and Fishing Foundation.
- Au, K. H. (1998). Social constructivism and the school literacy learning of students of diverse backgrounds. *Journal of Literacy Research*, 30(2), 297-319. <https://doi.org/10.1080/10862969809548000>
- Ayers, J. C. (2017). *Sustainability: An environmental science perspective*. CRC Press.
- Barak, M. (2017). Science teacher education in the twenty-first century: A pedagogical framework for technology-integrated social constructivism. *Research in Science Education*, 47(2), 283-303. <https://doi.org/10.1007/s11165-015-9501-y>
- Bigelow, B., & Swinehart, T. (Eds.). (2014). *A people's curriculum for the earth*. Rethinking Schools.
- Braund, M., & Reiss, M. (2006). Towards a more authentic science curriculum: The contribution of out-of-school learning. *International Journal of Science Education*, 28(12), 1373-1388. <https://doi.org/10.1080/09500690500498419>
- Casas Jr, E. V., Pormon, M. M., Manus, J. J., & Lejano, R. P. (2021). Relationality and resilience: Environmental education in a time of pandemic and climate crisis. *The Journal of Environmental Education*, 52(5), 314-324. <https://doi.org/10.1080/00958964.2021.1981205>
- Chiras, D. D. (2009). *Environmental science*. Jones & Bartlett Publishers.
- Christenson, M. A. (2004). Teaching multiple perspectives on environmental issues in elementary classrooms: A story of teacher inquiry. *The Journal of Environmental Education*, 35(4), 3-16.
- Corcoran, P. B. (1999). Formative influences in the lives of environmental educators in the United States. *Environmental Education Research*, 5(2), 207-220.
- Creswell, J. W. (2013). *Qualitative inquiry & research design: Choosing among five approaches* (3rd ed.). Sage Publications.
- Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage Publications.
- Dillon, J., & Herman, B. (2023). Environmental education. In N. Lederman, D. Zeidler, & J. Lederman (Eds.), *Handbook of research on science education, volume III* (pp. 717-748). Routledge.

- Eshach, H. (2007). Bridging in-school and out-of-school learning: Formal, non-formal, and informal education. *Journal of Science Education and Technology*, 16(2), 171-190.
<https://doi.org/10.1007/s10956-006-9027-1>
- Fien, J. (2002). Exploring principles of good practice: Learning from a meta-analysis of case studies on education within conservation across the WWF network. *Applied Environmental Education and Communication: An International Journal*, 1(3), 153-162.
- Frensey, B. T., Stern, M. J., Powell, R. B., & Sorice, M. G. (2022). Investigating the relationships among students basic psychological needs, engagement, and environmental literacy at a residential environmental education center. *The Journal of Environmental Education*, 53(4) 186-198. <https://doi.org/10.1080/00958964.2022.2081654>
- Haluza-Delay, R. (2012). Educating for environmental justice. In R. B. Stevenson, M. Brody, J. Dillon, & A. E. J. Walls (Eds.), *International handbook of research in environmental education* (pp. 390-399). Routledge.
- Harper, C. L., & Snowden, M. (2017). *Environment and society: Human perspectives on environmental issues*. Routledge.
- Hess, David J. (1995). *Science and technology in a multicultural world: The cultural politics of facts and artifacts*. Columbia University Press
- Hesson, N., & Roth, O. (2024). Effects of Informal versus School-Based Field Experience on Elementary Preservice Teachers' Self-Efficacy for Teaching Science. *The Electronic Journal for Research in Science & Mathematics Education*, 28(1), 76-92.
- Henderson, J. A., & Zarger, R. K. (2017). Toward political ecologies of environmental education. *The Journal of Environmental Education*, 48(4), 285-289.
<https://doi.org/10.1080/00958964.2017.1336978>
- Hungerford, H. R. (2009). Environmental education (EE) for the 21st century: Where have we been? Where are we now? Where are we headed?. *The Journal of Environmental Education*, 41(1), 1-6.
- Hungerford, H. R., & Volk, T. L. (1990). Changing learner behavior through environmental education. *The Journal of Environmental Education*, 21(3), 8-21.
- Kelani, R. R. (2015). Integration of environmental education in science curricula in secondary schools in Benin, West Africa: Teachers' perceptions and challenges. *The Electronic Journal for Research in Science & Mathematics Education*, 19(3).
- Klein, N. (2015). *This changes everything: Capitalism vs. the climate*. Simon & Schuster.
- Kollmuss, A., & Agyeman, J. (2002). Mind the gap: Why do people act environmentally and what are the barriers to pro-environmental behavior. *Environmental Education Research*, 8(3), 239-260.
- Malik H., Stronach, R. L., & Witzig, S. B. (2024a). Unpacking Roles of Nature Photography as a Tool to Spark Discussions & Actions on Environmental Issues. *Journal of Outdoor and Environmental Education*. 1-23. <https://doi.org/10.1007/s42322-024-00187-z>
- Malik, H., Stronach, R. L., & Witzig, S. B. (2024b). Highlighting 20+ years of Partnerships Between a Nature Center and Local Schools. *Connected Science Learning*. 6(5), 200-209.
<https://doi.org/10.1080/24758779.2024.2405694>
- Malik, H., Stronach, R. L., & Witzig, S. B. (2024c). Engaging Students in an Action-Oriented Coastal Ecology Course. In B. Simmons, C. Bennett III, Z. Grant, K. Kester, K. Khalil, L. Kirkland, G. Medina, M. Monroe, B. Morton, F. Niepold, A. Umali, S. W. Ridge, J. Stallard, L. Mann, & S. Weir (Eds), *Educating for climate action and justice: Guidelines for excellence* (Vol. 7, pp. 70-71). North American Association for Environmental Education.
- Malik, H., Witzig, S. B., & Stronach, R. L. (2025). Diverse Roles of Environmental Educators: Science Content Experts, Professional Development Providers, Environmental Advocates, and Mentors. *Paper Presented at Annual International Conference of the National Association for Research in Science Teaching*, National Harbor, MD, United States.

- Maxwell, J. A. (2008). Designing a qualitative study. *The SAGE handbook of applied social research methods*, 2(7).
- Ng, H. K. S., & Leung, A. N. M. (2022). Nature Connectedness and Nature Exposure Interactively Influence Social Dominance Orientation and Policy Support for Marginalized Groups during the COVID-19 Pandemic. *Environment and Behavior*, 54(2), 412–449. <https://doi.org/10.1177/00139165211031198>
- NGSS Lead States. (2013). *Next generation science standards: For states, by states*. National Academies Press.
- Palmer, J. A. (1993). Development of concern for the environment and formative experiences of educators. *The Journal of Environmental Education*, 24(3), 26-30. <https://doi.org/10.1080/00958964.1993.9943500>
- Rennie, L. (2014). Learning science outside of school. In N. G. Lederman & S. Abell (Eds.), *Handbook of research in science education*, Vol. 2 (pp. 120-144). Routledge.
- Richardson, V. (2003). Constructivist pedagogy. *Teachers College Record*, 105(9), 1623-1640.
- Robertson, M. (2021). *Sustainability principles and practice*. Routledge.
- Robinson, M., & Bowen, W. M. (2000). Global Environmental priorities of engineering students in Krakow Poland. *The Electronic Journal for Research in Science & Mathematics Education*.
- Savin-Baden, M., & Niekerk, L. V. (2007). Narrative inquiry: Theory and practice. *Journal of Geography in Higher Education*, 31(3), 459-472. <https://doi.org/10.1080/03098260601071324>
- Schmäing, T., & Grotjohann, N. (2022). The potential and limitations of the Wadden sea in university, school, and out-of-school contexts from the perspective of biology education. *Interdisciplinary Journal of Environmental and Science Education*, 18(4), e2286. <https://doi.org/10.21601/ijese/12072>
- Simmons, B., Bhagwanji, Y., Born, P., Drossman, H., Franzen, R. L., Mony, P., ... & Weir, S. (2025). Environmental Educator Knowledge and Skills: Guidelines for Excellence.
- Sisson, S. R. (2018). *Spring after spring: how Rachel Carson inspired the environmental movement*. Roaring Brook Press.
- Sosu, E. M., McWilliam, A., & Gray, D. S. (2008). The complexities of teachers' commitment to environmental education: A mixed methods approach. *Journal of Mixed Methods Research*, 2(2), 169-189. <https://doi.org/10.1177/1558689807313>
- Stevenson, R. B., Brody, M., Dillon, J., & Wals, A. E. J. (Eds.). (2013). *International handbook of research on environmental education*. Routledge.
- Stevenson, K. T., Carrier, S. J., & Peterson, M. N. (2014). Evaluating strategies for inclusion of environmental literacy in the elementary school classroom. *The Electronic Journal for Research in Science & Mathematics Education*, 18(8).
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research techniques* (pp. 1-312). Sage Publications.
- Stronach, R. L., Malik, H., & Witzig, S. B. (2026). The space between: Teacher perceptions of an interformal elementary science education program. *Journal of Research in Science Teaching*. 63(3), 195-214. <https://doi.org/10.1002/tea.70036>
- Tanner, T. (1980). Significant life experiences: A new research area in environmental education. *The Journal of Environmental Education*, 11(4), 20-24. <https://doi.org/10.1080/00958964.1980.9941386>
- The Associated Press (2021, January 21). *Key moments in Flint, Michigan's lead-tainted water crisis*. AP News. <https://apnews.com/article/us-news-health-michigan-rick-snyder-flint-7295d05da09d7d5b1184b0e349545897>
- Wang, J., Geng, L., Schultz, P. W., & Zhou, K. (2019). Mindfulness Increases the Belief in Climate Change: The Mediating Role of Connectedness With Nature. *Environment and Behavior*, 51(1), 3–23. <https://doi.org/10.1177/0013916517738036>

- Yanniris, C. (2015). 20+ years of environmental education centers in Greece: Teachers' perceptions and future challenges. *Applied Environmental Education & Communication*, 14(3), 149-166. <https://doi.org/10.1080/1533015X.2015.1067578>
- Zeiske, N., Venhoeven, L., Steg, L., & van der Werff, E. (2021). The Normative Route to a Sustainable Future: Examining Children's Environmental Values, Identity and Personal Norms to Conserve Energy. *Environment and Behavior*, 53(10), 1118–1139. <https://doi.org/10.1177/0013916520950266>

PCK Grain Size in the 21st Century

William Veal 
The University of Charleston

Jan Sermeus 
KU Leuven & Royal Observatory of Belgium

ABSTRACT

Pedagogical content knowledge (PCK) is specific to a given discipline, such as science. For example, the knowledge that a physics teacher has is different than the knowledge that a chemistry teacher has. This concept was refined with the introduction of different grain sizes of PCK that include domain-specific and topic-specific PCK. However, as was discussed during the 2024 Lorentz workshop on 21st-century science education, science teaching, learning, and understanding are becoming focused on broader issues that require knowledge and understanding outside of traditional disciplines. Furthermore, the focus on 21st-century skills (e.g., digital literacy, critical thinking, creativity, and collaboration and communication) necessitates students adapting to how they learn and embracing inter- and transdisciplinary approaches to solving societal issues. The new skills and nature of scientific knowledge will impact the PCK of 21st-century science teachers in that science teachers must be prepared to adapt their discipline-specific content knowledge and teaching methods to include inter- and transdisciplinary approaches. The purpose of this paper is to present toy cases as a tool to illustrate the impact of inter- and transdisciplinary science education on teachers' PCK, to address 21st-century skills, and to refine the grain size approach to PCK.

Keywords: pedagogical content knowledge, grain size, science teacher preparation

Introduction

Given the complexity and wickedness of the global and local scientific challenges that society faces in the 21st-century (e.g., climate change, energy, water, AI, and global health), there are calls for more collaboration between natural science disciplines (e.g., Weart, 2013) and for collaboration of the natural sciences with the humanities (e.g., Marcoci et al., 2023). Moreover, the complex societal challenges and the inter- and transdisciplinary approach needed to address them require that scientists develop and utilize collaboration and communication skills, critical thinking, creativity and flexibility, digital and data literacy, and problem-solving skills (i.e., 21st-century skills). The complex societal challenges, the inter- and transdisciplinary approach to (natural) sciences, and the required development of 21st-century skills are not just limited to science but also impact science education. The purpose of this paper is to utilize *toy cases*, an adaptation from toy-models and gedankenexperiments from physics (Redish, 2021), to highlight the knowledge bases required by teachers when teaching inter- and transdisciplinary topics, including socio-scientific issues, and to develop students' 21st-century skills.

Indeed, natural sciences are increasingly going beyond disciplinary boundaries in two ways. On the one hand, there is increasing interdisciplinary collaboration between the natural sciences. For example, there is research in physical chemistry/chemical physics, biochemical engineering, geophysics, computational mathematics, chip-based biosensors, and environmental science. On the other hand, the natural sciences also draw on other fields, such as public health and the humanities, and work in transdisciplinary fields of study. Examples include researchers involved in psychoacoustics, medical nuclear physics, urban studies (combining geography, architecture, history, and sports science), and disaster preparedness and response. Klein (2010) employed a typology from the Organization for Economic Cooperation and Development to define multidisciplinary, interdisciplinary, and transdisciplinary. Multidisciplinary teaching offers distinct methods and perspectives for examining a theme, allowing disciplines to remain separate. Interdisciplinary teaching integrates knowledge, theories, and methods from multiple traditional disciplines (e.g., science) to analyze an issue and synthesize them into a coordinated whole. Transdisciplinary teaching, in contrast, transcends traditional disciplinary boundaries and involves other subject areas (e.g., arts, business, mathematics, and public health) to create innovative solutions to complex, real-world problems.

The increased inter- and transdisciplinary research at universities already translates into inter- and transdisciplinary educational programs offered in university curricula. There is also increased attention in science education for developing 21st-century skills, both for future scientists and for citizens. Indeed, the so-called 21st-century skills are needed for adults in developed and developing nations to enter the workforce and society (Martin, 2018). If students need to develop these skills, it can be expected that their teachers also need to have and be able to transfer these skills.

Teachers can transfer science content and skills to students using diverse and appropriate methods if they have the necessary pedagogical content knowledge (Shulman, 1987). Schulman (1986) introduced the concept of pedagogical content knowledge (PCK) as being “an amalgam of content and pedagogy that is uniquely the province of teachers.” Science can be used as a rich context for developing 21st-century skills in education (Hilton, 2010). The increasingly inter- and transdisciplinary character of science offers additional opportunities to develop these skills, while also introducing students to science in the 21st-century.

In this paper, the impact of inter- and transdisciplinary science education and how it relates to 21st-century skills on the PCK of science teachers will be described. First, 21st-century skills, interdisciplinary and transdisciplinary science, and PCK are defined, including the original description of PCK grain sizes. Based on these definitions and the importance of socio-scientific issues, toy cases are presented that illustrate the impact of inter- and transdisciplinarity on science teachers’ PCK. Next, the relationship between PCK and 21st-century skills is highlighted, resulting in an updated PCK model with additional grain sizes. Finally, the significance of the new model is discussed in the context of science teacher education.

Problem Statement

At the 2024 workshop on 21st-century Science Education (Chan et al., 2024), science education researchers learned about and discussed how the 21st-century brings about several challenges in science teaching and its impact on science teacher training. In short, science knowledge is growing exponentially (Bornmann et al., 2021), scientific information is spreading very quickly through digital information sources (but, unfortunately, so is misinformation), scientists are increasingly specialized, and science is increasingly inter- and transdisciplinary (e.g., Barthel & Seidl, 2017; Van Noorden, 2015).

Inter- and transdisciplinary research brings about specific challenges. The scientists working together ideally have some understanding of the other field and know which questions they can ask their colleagues. Additionally, they must find common ground across disciplines in the theoretical

foundations of their work, research methodologies (i.e., What are common research practices?), epistemological aspects (i.e., What constitutes knowledge or evidence in a field?), and communication (How are journal articles written, which journal to choose, and where to disperse information?). Moreover, as collaboration is a relatively new phenomenon, it is unclear how it manifests itself. For example, Gleason (2020) argues that this could be anything from ignoring each other to collaborating as equal partners to one serving the other. All approaches have advantages and pitfalls.

Given that science is increasingly inter- and transdisciplinary, it can be expected that science education will follow suit. Indeed, there exist inter- and transdisciplinary master's programs (e.g., Science and Math for Education, Environmental Studies, Public Health, Biophysics, Biochemistry, Biotechnology, and Sustainability). The challenges presented earlier also find their way into science education and the classrooms of the 21st century, including the K-12 level. Indeed, there is an increasing call to teach 'STEM' in an integrated way and move beyond the siloed approach that was common in the early years of STEM education (Widya et al., 2019). The question thus arises of how the inter- and transdisciplinary nature of 21st-century science (teaching) impacts the PCK of science teachers.

Background

Teaching science content in the 21st-century requires a diverse set of knowledge and skills that can be associated with PCK. Shulman's (1986) initial concept of PCK focused on students, content, teaching strategies, and context as a knowledge base for teaching that was unique to the profession. At first, PCK was "the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations and demonstrations – in a word the ways of representing and formulating the subject that make it comprehensible to others" (Shulman, 1986, p. 9). It soon became one part of a teachers' knowledge base comprised of seven categories of knowledge: content knowledge; pedagogical knowledge; curriculum knowledge; knowledge of learners; knowledge of context; knowledge of educational ends, purposes and values; and pedagogical content knowledge (Shulman, 1987). The PCK of teachers in the 21st-century must change to adapt to the integrated approach needed to solve societal issues.

21st-century Skills

There have been many frameworks to conceptualize "21st-century skills." Different literature reviews of papers published in the last 25 years have resulted in lists of emergent categories of skills (e.g., Chen, 2021; González-Perez & Ramirez-Montoya, 2022; Kotsiou et al., 2022). Chen's review of the literature identified seven domains of 21st-century or future skills, including higher-order thinking skills; dialogue skills; digital and STEM literacy skills; values; self-management; lifelong learning; and enterprise, leadership, and flexibility skills (2021). González-Perez and Ramirez-Montoya (2022) created a list of Education 4.0 core competencies. The transversal competencies include critical thinking, cooperation, collaboration, communication, and creativity. The disciplinary competencies include developing technical knowledge and using emerging technologies. Karaca-Atik et al. (2023) developed a list of 21st-century skills for social science students that also included communication, problem solving, critical thinking, initiative and self-direction, and social and cross-cultural skills. For the purposes of this paper, 21st-century skills encompass a composite list of the literature reviews cited, and they include digital literacy, critical thinking and problem solving, creativity and flexibility, and collaboration and communication.

Skills are extremely important for changing the economic, scientific, and technological landscape. One of the most crucial 21st-century skills is digital literacy. With the rapid advancement

of technology, educators and learners need to be proficient in using digital tools and platforms. This includes basic computer skills and the ability to navigate and utilize various educational technologies effectively to distinguish fact from fiction. Digital literacy encompasses information, technology, and media literacies (Bliss, 2019), all of which require epistemic vigilance (Osborne, 2024). Epistemic vigilance is the ability of students to evaluate a digital media source, and “the expertise level of the source and to evaluate the claim itself, using scientific standards and appropriate argumentation heuristics” (Bielik & Krell, 2025, p. 869). Digital literacy enables students to access a wealth of information and resources online, while teachers can enhance their teaching methods and engage students in more interactive and innovative ways.

Another vital 21st-century skill is critical thinking and problem-solving. In today's complex and fast-paced world, the ability to analyze information, think critically, and solve problems creatively is more important than ever (Jin et al., 2023). Critical thinking was defined as the “purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as an explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based” (Facione, 1990, p. 3). Science teachers need to foster an environment that encourages students to question assumptions, evaluate evidence, and develop well-reasoned arguments like the tenets of the nature of science (Galili, 2019). This skill set is essential for students to navigate the vast amount of information online and make informed decisions about the science content. Moreover, critical thinking and problem-solving skills are highly valued in the workforce, making them indispensable for students' future careers (Humphreys & Kelly, 2014).

Creativity and flexibility are also necessary in the 21st-century for both teachers and students. For example, when teachers make decisions, they are aided by the ability to creatively develop new pedagogical tools to enhance the transferability of the content to the students. To achieve this, teachers must learn to be flexible in the classroom so that students can learn the content in appropriate formats (Al-Abdali & Al-Balushi, 2016). Knowing the students' context, culture, and backgrounds and the different structures and representations of the content topics affords the teacher the flexibility to create a rich environment for learning science content. It is the interplay between PCK and creativity in teachers that enables the development of 21st-century skills (Veal & Kind, 2025).

Collaboration and communication are key 21st-century skills that teachers and students must develop. The modern world is becoming increasingly interconnected, and the ability to work effectively with others is becoming increasingly crucial. Teachers should promote collaborative learning experiences where students can collaborate on projects, share ideas, and learn from one another (Rodriguez et al., 2019). Effective written and verbal communication skills are essential for expressing ideas clearly and persuasively. By developing these skills, students can better prepare for the collaborative nature of the modern workplace, while teachers can create a more dynamic and inclusive classroom environment.

Interdisciplinary Science Education

In addition to the development of 21st-century skills by students, which are necessary for living in a complex society, the world's problems can no longer be evaluated and solved through a single lens. Schools need to prepare these students for the real world through progressive pedagogical approaches that highlight the interconnectedness of different subjects and domains of knowledge. The 21st-century classroom thus necessitates an interdisciplinary and transdisciplinary approach to science content (e.g., Hubber et al., 2022). This, in turn, requires that science teachers adjust how and why they teach science content so that their students are prepared for the changing societal demands.

One method is to develop a perspective on science teaching that integrates other science disciplines and other subject areas.

A prominent pedagogical approach that has gained favor in the last 20 years is the integration of science, technology, engineering, and mathematics into STEM. The research literature contains examples of the implementation of integrated science lessons using different pedagogical approaches (e.g., problem-based learning and inquiry) (e.g., Deák et al., 2021). Integration has been synonymous with an interdisciplinary approach. Most research studies have focused on the integration of two disciplines (e.g., mathematics and technology) (e.g., Furner, 2024) or a cross-disciplinary approach (e.g., Chang & Chen, 2020). This body of research also includes many studies on STEM in which integrating four disciplines has often been recommended to support students in developing 21st-century skills. However, many STEM studies have focused on a multidisciplinary, rather than an interdisciplinary, approach to teaching. For example, most studies focus on science, mathematics, and technology as separate content areas that can contribute independently to a central theme (e.g., English, 2023). In contrast, an interdisciplinary approach requires all content areas to be seamlessly used to solve the theme/problem (e.g., Sikand et al., 2021).

In other words, science teachers still need to be experts in their own content discipline but also will need to be able to integrate science topics from other disciplines and more topics from other subject areas. This implies that science teachers will have to expand their content and PCK to include the application of interdisciplinary and transdisciplinary knowledge.

Pedagogical Content Knowledge

Since its inception, PCK has been considered and researched as its own knowledge base (e.g., Magnusson et al., 1999; Veal & MaKinster, 1999). In the brief history of PCK, many researchers have defined PCK as having multiple knowledge bases (e.g., Grossman, 1990; Marks, 1990; Cochran et al., 1993; Magnusson et al., 1999; and Veal & MaKinster, 1999). Indeed, research has shown that there are multiple knowledge bases that a teacher must have. However, there is no consensus on the relationship among the components (Friedrichsen et al., 2011).

There have been a variety of models describing the attributes or knowledge bases related to and describing PCK. For example, Grossman (1990) stated that a teacher with PCK had knowledge of representations, students' understanding, curriculum, and assessment. These were like Shulman's but focused specifically on effective teaching practices. Other researchers (e.g., Magnusson et al., 1999; Marks, 1990) similarly described PCK using the same knowledge bases and expanded some to include the subject matter and general pedagogy. Veal and MaKinster (1990) elaborated on the initial idea that PCK was its own knowledge base and included many other knowledge bases. All the knowledge bases described by other researchers were included in the PCK taxonomy developed by Veal and MaKinster (1999). Table 1 shows the initial development of the definition of PCK to include multiple knowledge bases.

Table 1*Pedagogical Content Knowledge and its Knowledge Bases*

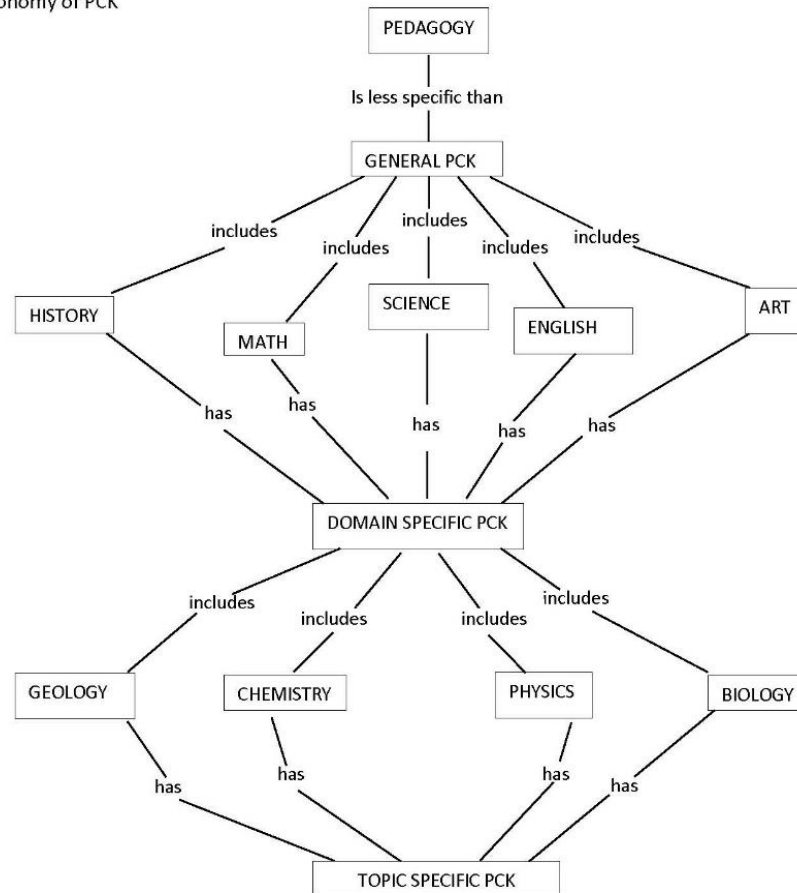
	Grossman (1990)	Marks (1990)	Magnusson et al. (1999)	Veal & MaKinster (1999)
Knowledge of				
Representations	X	X	X	X
Students' Understanding	X	X	X	X
Subject Matter		X		X
Curriculum	X	X	X	X
Assessment	X		X	X
Environment/Context				X
Socio-cultural issues				X
General Pedagogy				X
Classroom Management				X

Veal and MaKinster (1999) introduced an elaboration to understanding how teachers learn to teach science by determining that there are different grain sizes to PCK. Figure 1 shows the original conceptualization of the grain sizes related to PCK. All teachers possess general pedagogical skills, which include the ability to design a lesson plan, deliver a lecture, form students into groups, and administer assessments. General PCK is common across all disciplines and includes such skills as providing an advanced organizer for students before instruction and facilitating transitions between concepts and activities. However, teachers also have more specific forms of PCK. There is PCK specific to teaching a discipline of 'science' (e.g., introducing the Nature of Science to their students), as well as PCK specific to a domain of science (e.g., physics teachers introducing their students to model-based thinking) and PCK specific to the topics within a domain of science (e.g., a physics teacher working with their students on a lesson on mechanics to recognize and avoid the weight-mass confusion). The cascading and refined knowledge base of PCK begins with general PCK, which is shared by all teachers, and progresses to Domain-specific PCK (DS-PCK) and ultimately to Topic-specific PCK (TS-PCK). Figure 1 represents the flow of grain size originally established by Veal and MaKinster (1999). This hierarchy defines the grain size of PCK, which has been evaluated and applied to science teaching research studies since its inception (e.g., Rollnick & Mavhunga, 2015). It remains an integral part of the Consensus Model (Gess-Newsome, 2015) and the Reformed Consensus Model (Carlson et al., 2019) of PCK.

Figure 1

Grain Size for PCK (presented with permission from the author)

General Taxonomy of PCK



Aims and Methods

This work aims to discuss the impact of an approach to science education that embraces inter- and transdisciplinarity and that develops 21st-century skills in science teachers' PCK by presenting several toy cases. This methodology is inspired by toy models or gedankenexperiments (Luczak, 2017), a known methodological tool in theoretical physics. These are deliberately simplistic and idealized models and experiments used to explore and explain the core mechanisms of a more complex physical system (Redish, 2021). By intentionally stripping away complicating factors, physicists can gain insight into a fundamental principle without getting bogged down in real-world complexity. The model is not meant to be a precise description of reality, but a "caricature" that is easy to analyze and helps build intuition. In an analogous way, "toy cases" are simplified hypothetical situations used to think through complex educational situations.

Cases have been used in various educational settings for decades, including education, social science, business, law, and medicine. Shulman (1986) stated, "While cases themselves are reports of events, the knowledge they represent is what makes them cases" (p. 22). They focus on collecting

qualitative and quantitative data to explore complex issues in real-life contexts. By examining specific cases, researchers can identify patterns, relationships, and outcomes that contribute to a deeper understanding of broader phenomena. Toy cases in this paper serve as idealized, simple, and focused representations of how to teach specific inter- and transdisciplinary concepts, including socio-scientific issues, that reflect the need to use 21st-century skills.

For this research, hypotheticals were developed to reflect a particular form of narrative that examines a particular instance, individual, group, or event. These toy cases are used to elucidate and illuminate professional knowledge bases and attributes of science teachers when teaching interdisciplinary and transdisciplinary science issues. It should be noted that the toy cases are inspired by and describe a realistic classroom situation that incorporates facts, examples, and ideas (Kagan, 1993). The toy cases in this paper are intended to portray scientific issues that span the interdisciplinary and transdisciplinary domains of education. While the science content of the toy cases is critical, it is the way that the inter- and transdisciplinary nature of the content impacts the method of instruction that should be the centerpiece of discussions.

A compact and qualitative notation has been developed to describe the different kinds of PCK a science teacher needs in the 21st-century. The notation described in Table 2 is based on the notation typical in conditional probability. So, in general, $PCK(A|B)$ should be read as the PCK necessary to teach A given B. Throughout the text, the notation uses cap B, cap C, cap P, and cap G to refer to biology, chemistry, physics, and geosciences, respectively. Non-natural science disciplines are written in full or using common abbreviations. If PCK is specific to the intersection between disciplines, this is noted with $\otimes$. Finally, PCK that is specific to a person (or group) is marked as PCK_{Tp} to be read as: the PCK of Teachers of Physics. This notation will only be used if it is needed to differentiate between teachers. The notation does not distinguish between personal and collective PCK, as introduced in the refined consensus model of PCK (Carlson et al., 2019). Rather, in this work, a teacher's personal PCK is simplified to be equivalent to the collective PCK. As an example, in this notation, the last sentence is written as $PCK_{Tp}(topic) = PCK(topic|P)$, assuming that a physics teacher approaches topics from a Physics perspective.

Table 2

Notation of Toy Cases as the Content Applies to PCK Grain Size

PCK level	Notation	To be read as
Transdisciplinary PCK	$PCK(topic P \otimes history)$	PCK about teaching a topic at the interface between physics and history.
Interdisciplinary PCK	$PCK(topic P \otimes C)$	PCK about teaching a topic at the interface between physics and chemistry
Disciplinary-specific PCK	$PCK(P)$	PCK about teaching physics (in general)
	$PCK(P topic)$	PCK about teaching physics using a specific topic
Topic-specific PCK	$PCK(topic P)$	PCK on teaching a topic (within physics)

Case Studies

Toy Case 1 Multidisciplinary DS-PCK: The Optics of the Eye in Biology and Physics

Case

A physics teacher is introducing geometrical optics to his students. He works with his students through the optical paths of a (ideal) convex lens and constructs, on the blackboard, an image similar to that in Figure 2a. Subsequently, he shows an application of this model, the lens of the human eye, by showing Figure 2b. The teacher points out that this is a simplification. The eye is much more complex, but the teacher refers the students to their biology teacher for those details. “For example,” he adds, “notice how the image is upside down! The brain can correct this.” While the physics teacher also knows that the retina contains rods and cones, that humans are sensitive to red, blue, and green light, and that the optical nerve passes through the retina, creating a blind spot, he does not go into those details, trusting his biology colleague to explain this in a more thorough way.

Next door, his biology colleague is introducing the human eye to her students. She starts with a detailed diagram of the eye (Figure 2b), showing the different anatomical parts and explaining their functions. She points out that humans perceive the world upside down; luckily, the brain can correct this. This is due to the way lenses work, she explains by superimposing a simplified path of the light on top of the anatomical diagram (Figure 2b). “This is just one part of optics, though,” she points out to her students. For more details on the physics content of geometric optics, she refers the students to her physics colleague. Like her colleague, the biology teacher knows basic physics. She knows that concave lenses diverge light, how to create basic geometric optical diagrams, and the lens formula. But she, too, knows that her physics colleague will demonstrate the workings of lenses, has lab-work planned for the students, and will derive the lens-formula.

Case Discussion

The integration of the different disciplines is limited to a multidisciplinary approach. However, in this first case, there is PCK of the teachers that goes beyond their own discipline. When teaching about the optical path of light in the eye, both physics and biology teachers recognize that students are typically unaware that the image projected in the eye is upside down. Moreover, both teachers know the basic subject matter of the other discipline. They can use this knowledge to guide students through a representation of elements of both disciplines. However, they also know that their knowledge is limited. As such, based on their knowledge of what is planned in the curriculum of their colleague, they can refer to each other.

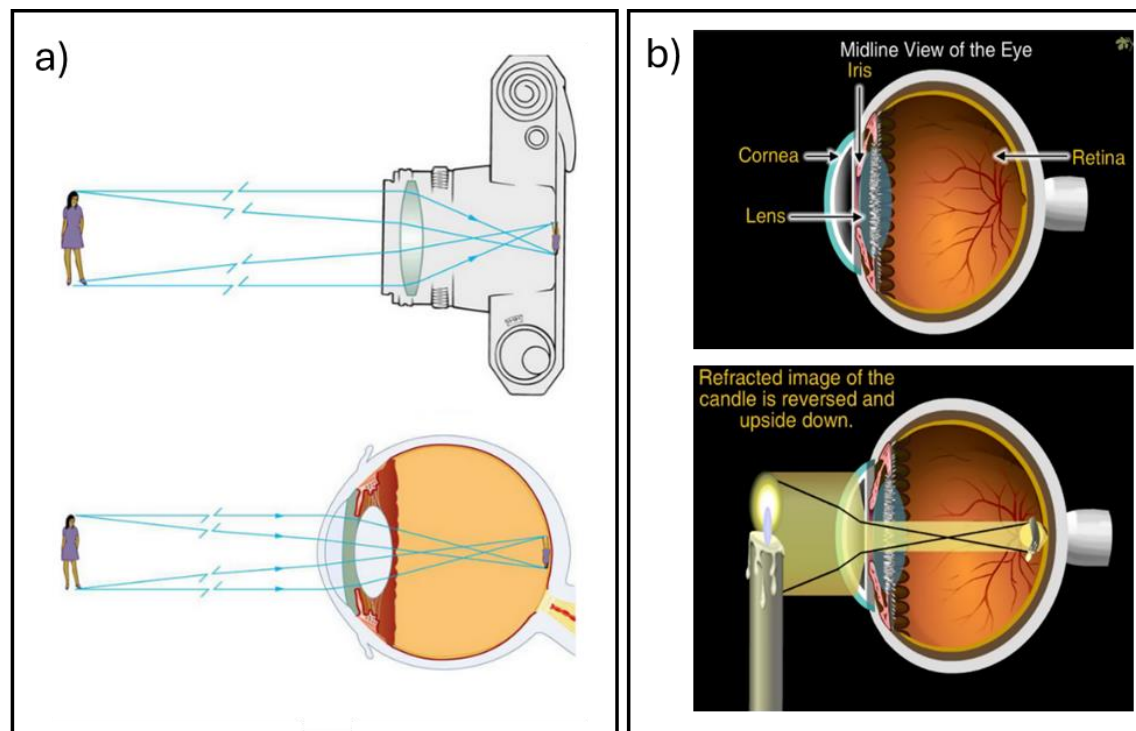
$$PCK_{T_P}(\text{eye}) = PCK(\text{optics}|P) + PCK(\text{eye}|P \otimes B)$$

$$PCK_{T_B}(\text{eye}) = PCK(\text{vision}|B) + PCK(\text{eye}|B \otimes P)$$

The $PCK_{T_P}(\text{eye})$ represents the physics teacher’s multidisciplinary PCK related to the eye. It includes TS-PCK of optics (as it is understood in physics) and knowledge of where physics meets biology for the particular content of the optics of the eye. Likewise, the biology teachers’ multidisciplinary PCK consists of their TS-PCK concerning vision and the knowledge of where biology meets physics for the particular content of the optics of the eye. Note that $PCK(\text{eye}|B \otimes P) \approx PCK(\text{eye}|P \otimes B)$, as the topic-specific PCK of the physics teacher is quite similar to the topic-specific PCK of the biology teacher. Of course, they approach the topic in a different order and include this segment about the optics of the eye in different lessons (the physics teacher focuses on optics, the biology teacher focuses on the biological working of the eye).

Figure 2

Images used by a) a Physics Teacher (adapted from Urone & Hinrichs, 2025) and b) a Biology Teacher (adapted from AniMed, 2017).



Toy Case 2 Interdisciplinary PCK: Health Monitor STEM Project

Case

A school organizes a STEM project. The mathematics teacher (Alice), the physics teacher (Bob), and the Information and Communications Technology (ICT) coordinator (Charly) work together to develop the lessons. Each teacher is responsible for their part. Alice wants to ensure that the mathematics in the STEM project goes beyond calculations, Bob wants to ensure that students develop some conceptual understanding of the science involved, and Charly wants students to learn the basics of coding. The project they develop challenges the students to “measure their physical condition,” making them think about energy, building a chronometer using an Arduino to measure how fast they run, and using (discrete) derivatives to calculate velocity and acceleration from positional sensor data.

Students from several classes work on this project in groups under the supervision of Alice, Bob, and Charly. If Alice or Bob gets questions regarding the Arduino, they call on Charly. Charly knows what mistakes beginners typically make in setting up and coding an Arduino (e.g., wiring mistakes, coding errors, syntax errors, powering everything using 5V, etc.). Charly can also provide coding examples that can serve as a starting point for the students. Similarly, Alice and Charly expect Bob’s interventions to force students to think about energy, velocity, and acceleration concepts. He knows common misconceptions and can quickly demonstrate how distance traveled, velocity, and acceleration are related. And Bob and Charly know that Alice will introduce the idea of derivatives as

the local slope of a curve and set students up to think about what happens if they make the time interval smaller and smaller.

After a while, Alice and Bob also gain the confidence to help students in their coding, especially when the students are struggling with translating the idea of the derivative into code. Likewise, Alice and Charly will, for example, recognize when students failed to take the second derivative to calculate the acceleration. Bob and Charly also grow in their support of the students when they are mathematically processing the data.

Towards the end of the project, a student asks Alice a question about energy usage in the body: “How is the energy in my sugary drink converted into the motion of my body?” Alice recognizes this as a science question and calls Bob. Bob recognizes that this is indeed a biology question, but also that this is not his expertise. As such, he first had to ask a biology colleague (not involved in the project) and could only get back to the student a week later. After the project is completed, the teachers discuss the possibility of including the biology teacher next year.

Case Discussion

In this toy case, the teacher initially approached the STEM project in a siloed way. Each teacher was responsible for their own subject. Of course, each teacher is expected to have PCK regarding their own subject. It enables them to distinguish aspects of the project that are within their own subject. Hence, initially, the combined PCK of the teacher team TT for this project $STEM$ can be written without any PCK at the intersection as

$$PCK_{TT,i}(STEM) = PCK_{T_M}(STEM|M) + PCK_{T_P}(STEM|P) + PCK_{T_{ICT}}(STEM|ICT)$$

The initial PCK of the teacher team related to STEM is equal to the combined PCK of the math, physics, and ICT teachers who approach the STEM project from their own disciplines. However, as is typical in a STEM project (and in the reality of S, T, E, and M beyond the school walls), the different subjects are not independent. Rather, they are interwoven. Velocity and acceleration, concepts from physics, rely heavily on mathematics. The time interval in which data are gathered is the denominator in the discrete derivatives but also depends on the technological capabilities of the sensor and the Arduino. Hence, as the teachers grew throughout the project, they recognized that they also needed to know where their subject touches upon neighboring subjects, and that they needed to know elements of those neighboring subjects. They also needed to know which aspects of the project are beyond their own expertise and they need to know when and whose expertise to call on. As such, the integration goes from multidisciplinary to interdisciplinary. Moreover, the PCK of the teacher team expands, in this notation, to

$$PCK_{TT,f}(STEM) = PCK_{T_M}(STEM) + PCK_{T_P}(STEM) + PCK_{T_{ICT}}(STEM) + 0 * PCK_{T_{P,M,ICT}}(B|STEM)$$

where

$$PCK_{T_P}(STEM) = PCK(STEM|P) + PCK(STEM|P \otimes M) + PCK(STEM|P \otimes ICT)$$

$PCK_{T_M}(STEM)$, $PCK_{T_P}(STEM)$, and $PCK_{T_{ICT}}(STEM)$ are analogous. Note that the final PCK of the physics teacher now includes not only their PCK on the STEM project as they would approach it from their own field (physics), expressed as $PCK(STEM|P)$, but also their knowledge of the boundaries of physics and where it meets mathematics and ICT, expressed as $PCK(STEM|P \otimes M) + PCK(STEM|P \otimes ICT)$. Also note that in the combined interdisciplinary PCK of the teacher team the $0 * PCK_{T_{P,M,ICT}}(B|STEM)$ refers to the inability of the team to recognize the biological elements in the project. The PCK needed by the team to teach this STEM project is dependent upon the collective and integrative knowledge of physics, math, and ICT teachers. This case reflects an interdisciplinary situation, even though it lacks a key component of a biology teacher.

Toy Case 3a Interdisciplinary PCK: Climate Change

Case

Climate change is a topic that is inherently not bound to a single discipline. Ms. Rodriguez is an Earth sciences teacher who calls on her biology, chemistry, physics, and engineering colleagues to spend time on climate change. She sets up a meeting to discuss who can take up which aspects of climate change in their lessons. The notes of this meeting are given in Figure 3a.

Case Discussion

By coming together and delineating who is best suited to discuss which aspect of climate change, the teacher team has taken a multidisciplinary, interdisciplinary approach to teaching climate change. After the meeting, all the teachers involved are aware of which aspects are covered in the other disciplines and of potential student difficulties (that may also arise in their own lessons). Combined, the teacher team (TT) has clarified and expanded their PCK for teaching the topic of climate change. It now, as a team, encompasses all individual PCKs and a shared understanding of the topic.

$$PCK_{TT}(CC) = PCK_{BT}(CC|B) + PCK_{CT}(CC|C) + PCK_{GT}(CC|G) + PCK_{ET}(CC|E) \\ + PCK_{PT}(CC|P) + PCK(CC|B \otimes C \otimes E \otimes G \otimes P)$$

Where E refers to engineering and

$$PCK_{PT}(CC) = PCK(CC|P) + PCK(CC|P \otimes B) + PCK(CC|P \otimes C) + PCK(CC|P \otimes E) \\ + PCK(CC|P \otimes G)$$

Note that the interdisciplinary PCK of the teacher team not only depends on the interdisciplinary PCK of each teacher (which, in turn, consists of TS-PCK from the discipline and knowledge of the intersections with other disciplines), but also on the whole team knowing where all disciplines come together. The latter is expressed as $PCK(CC|B \otimes C \otimes E \otimes G \otimes P)$.

Toy Case 3b Transdisciplinary PCK: Climate Change

Case

Ms. Rodriguez also teaches science to non-science majors. For this group, she calls on her colleagues from all other disciplines to explore how they can integrate climate change into their courses. She sets up a meeting to discuss who can take up which aspects of climate change in their lessons. The notes of this meeting are given in Figure 3b. Students are challenged to create a proposal to build a climate-proof house within a specific budget and present it to the Architectural Review Board of the local city council.

Figure 3a*Discussions and Assignments of Content for Teachers***Biology**

- Data and Examples: Discuss the impact of climate change on biodiversity, ecosystems, and species migration patterns. Use data on changing habitats and species extinction rates.
- Graphics: Include charts showing the decline in polar bear populations and maps of shifting biomes.
- Expected student difficulties: Climate change is a natural change.

Chemistry

- Data and Examples: Explain the chemical processes behind greenhouse gas emissions and their effects on the atmosphere. Use data on CO₂ levels (Keeling curve) and their correlation with global temperatures.
- Graphics: Provide molecular diagrams of greenhouse gases and graphs showing historical CO₂ concentrations.
- Expected student difficulties: The ozone hole is responsible for climate change, and only CO₂ is a greenhouse gas.

Physics

- Data and Examples: Describe the physical principles of energy transfer and the greenhouse effect. Use data on solar radiation and heat absorption.
- Graphics: Include diagrams of the greenhouse effect and charts showing the energy balance in the Earth's atmosphere.
- Expected student difficulties: Changing distance to the sun and changing solar activity.

Geology

- Data and Examples: Explore the geological evidence of past climate changes and their causes. Use data from ice cores and sediment layers.
- Graphics: Provide images of ice core samples and graphs showing historical temperature variations.
- Expected student difficulties: The Earth's climate has always changed.

Engineering

- Data and Examples: Discuss engineering solutions to mitigate climate change, such as renewable energy technologies and carbon capture. Use data on the efficiency and adoption rates of these technologies.
- Graphics: Include schematics of solar panels and wind turbines, and charts showing the growth of renewable energy capacity.
- Expected student difficulties: Technology (e.g., carbon capture) alone is enough to mitigate climate change, and nuclear power emits CO₂.

Figure 3b*Meeting Notes***Science**

- **Data and Examples:** Use data on rising global temperatures, melting ice caps, and increasing CO₂ levels. Include graphics such as the "hockey stick" graph showing temperature rise over the past century.
- **Experiments:** Conduct simple experiments to demonstrate the greenhouse effect and the impact of CO₂ on temperature.
- **Resilient Design:** Explore how infrastructure can be designed to withstand the impacts of climate change. Use examples of green buildings and sustainable urban planning.

Sociology and Public Health

- **Social Impact:** Discuss how climate change affects different communities, particularly vulnerable populations. Explore the health effects of climate change, such as the spread of diseases, heat-related illnesses, and respiratory issues due to air pollution. Use case studies of climate refugees and the social dynamics of communities affected by extreme weather events.
- **Class Discussions:** Facilitate discussions on environmental justice and the role of social movements in addressing climate change.
- **Media Literacy:** Teach students to critically evaluate climate change information in the media and identify misinformation.

Mathematics and Economics

- **Data Analysis:** Teach students how to analyze climate data using statistical methods. Include exercises on calculating trends and making predictions based on historical data. Use data on the cost of natural disasters and the economic benefits of sustainable practices.
- **Graphing:** Use graphs to illustrate changes in temperature, sea levels, and CO₂ concentrations over time. Teach students how to perform cost-benefit analyses for energy savings.

English and Literature

- **Literary Analysis:** Analyze literary works that address environmental themes, such as "Silent Spring" by Rachel Carson, "The Road" by Cormac McCarthy, or "Janouk zoekt een duurzame wereld" by Janouk Kelderman (Janouk is Looking for a Sustainable World).
- **Creative Writing:** Encourage students to write the project proposal, essays explaining the science behind the structure, and short stories about their perspectives on climate change.

History

- **Historical Context:** Provide a historical overview of climate change, including the Industrial Revolution's impact on the environment and key milestones in climate science.
- **Timeline:** Create a timeline of significant events in the history of climate change awareness and policy.

Case Discussion

Climate change is a wicked problem; and the solutions are all interconnected. There is also a huge societal component to climate change as it is human-induced. The efforts of Ms. Rodriguez to bring together a teacher team and to approach climate change from many angles and disciplines allow students to get both a broader and a deeper understanding of climate change. The project allows teachers to use their depth of content knowledge to guide students in the project and through other activities. The interconnectedness of the project variables reflects the myriad methods that the teachers must use to present new content and relate it to the real world. The guiding questions of the content should focus on the socio-cultural issues related to climate change using environmental justice. The project to build a sustainable house allows the students to evaluate online resources, collect data, organize data, critically think about solutions, and create their vision of a sustainable home. The project's culmination is an authentic assessment in which students must problem-solve and communicate their results to an outside body of experts.

$$PCK_{TT}(CC) = PCK_{NST}(CC|NS) + PCK_{historyT}(CC|history) + PCK_{econT}(CC|economy) + \dots + PCK(CC|NS \otimes history \otimes economy \otimes \dots)$$

The transdisciplinary PCK of the teacher team is equivalent to the PCK about climate change of natural science teachers and the collective PCK about climate change of the other team members in other subject areas as all subject areas are interfaced with each other. Where *NS* is short for Natural Science.

Toy Case 4a Interdisciplinary PCK: Evolution

Case

In a secondary science classroom, Mr. Smith, a dedicated science teacher, embarks on an innovative journey to teach the topic of evolution. Recognizing the complexity and interconnectedness of scientific disciplines, he adopts an interdisciplinary approach that integrates data, examples, and graphics from biology, geology, chemistry, and physics. This method enhances students' understanding of evolution and demonstrates the relevance of various scientific fields in explaining natural phenomena. The instruction uses multiple hands-on activities and labs to help students overcome traditional misconceptions related to evolution as a theory and its explanation of environmental biodiversity.

Mr. Smith introduces the fundamental concepts of evolution, emphasizing the role of natural selection, genetic variation, and adaptation. He starts by proposing some overarching and guiding questions to orient the students. For example, he can ask, "How do environmental and geological changes, and natural selection drive changes in a population?" He then uses the example of antibiotic resistance in bacteria to illustrate how genetic mutations can lead to advantageous traits that increase an organism's chances of survival and reproduction. This example is grounded in biology but also incorporates chemistry by explaining the molecular interactions that drive genetic mutations.

To provide a broader context, Mr. Smith integrates geology by discussing how the Earth's geological history, including the formation of continents, volcanic activity, and mass extinctions, has influenced the evolution of life. He uses data from geological surveys and paleontological studies to show how different species have evolved over millions of years. Students learn how to interpret the geological record and uncover evidence of ancient life forms by studying rock layers and fossils.

Mr. Smith also incorporates physics by exploring the role of physical forces, such as gravity, radiation, and temperature, in shaping the evolution of life on Earth. He discusses how environmental conditions can drive evolutionary change and uses data from ecological studies and climate models to show how physical factors impact the distribution and behavior of species. This helps students understand the complex interactions between organisms and their environment.

To delve deeper into the mechanisms of evolution, Mr. Smith incorporates chemistry by examining the structure and function of DNA. He explains how chemical reactions and molecular interactions drive genetic mutations and the formation of new traits. By presenting data on the molecular composition of different organisms and using graphics to illustrate the processes of DNA replication, transcription, and translation, students gain a deeper understanding of how genetic information is passed down from one generation to the next.

Throughout the lessons, Mr. Smith uses hands-on activities, such as examining fossil specimens, analyzing DNA sequences, and conducting experiments to observe chemical reactions. He also incorporates graphics, such as phylogenetic trees, geological time scales, and diagrams of ecosystem energy flow, to provide students with tangible evidence of evolutionary processes and help them visualize the dynamic changes that have shaped the Earth's surface and its inhabitants. Assessment of students' content knowledge is accomplished through the completion of graphic organizers, labs, and traditional quizzes. The integration of the different activities, examples, and manipulatives reflects how all science disciplines can be used by one teacher to help students understand an interdisciplinary socio-scientific theme.

Case Discussion

Through this integrated interdisciplinary approach, Mr. Smith engages his students in studying evolution. The activities, labs, and representations used by the teacher help enforce the concepts of evolution. He provides an interdisciplinary, comprehensive, and holistic understanding of the topic by integrating data, examples, and graphics from biology, geology, chemistry, and physics. This method not only enhances students' scientific literacy but also fosters an appreciation for the interconnectedness of the natural world.

Mr. Smith uses subject knowledge from all the science disciplines and knowledge about the different representations (e.g., graphics, illustrative examples, and pictures) that can enhance students' understanding of evolution. He also has knowledge of the environment and context of where the different disciplines intersect with the students' lives, and which examples allow him to get to that intersection. For example, while Mr. Smith could also have chosen to illustrate evolution using examples on much shorter timescales (e.g., the changes in the coloring of the peppered moth in England due to the Industrial Revolution), he did not. By showing this specific example, he could connect to geological times, denoted in the notation as $PCK(\text{evolution}|B \otimes G)$.

To develop the students' 21st-century skills, Mr. Smith has students use technology to download data, determine the validity of the data, and analyze the data. During the data analysis, students engage in critical thinking and problem-solving to answer the guiding questions posed by the teacher.

Combined, Mr. Smith's integrated interdisciplinary PCK regarding the topic of evolution can be denoted as

$$\begin{aligned}
 PCK(\text{Evolution}|B \otimes G \otimes P \otimes C) \\
 = PCK(\text{evolution}|B) + PCK(\text{geological time}|B \otimes G) + PCK(\text{evolution}|G) \\
 + PCK(\text{evolution}|G \otimes P) + PCK(\text{evolution}|P) + PCK(\text{evolution}|P \otimes C) \\
 + PCK(\text{evolution}|C)
 \end{aligned}$$

The teacher's interdisciplinary PCK about evolution is the equivalent of the PCK about evolution of each science domain and the PCK about evolution at the interface of each science domain with another science domain.

Toy Case 4b - Transdisciplinary PCK: Evolution

Case

Mr. Smith has had several students discuss the validity of the theory of evolution by natural selection and is aware of discussions on social media and in school board meetings on evolution. He thinks it is important that students are aware of various arguments in these discussions, their history, and implications. Hence, in the following year, Mr. Smith expanded his integrated approach to additional disciplines beyond the natural sciences. It is now an integrated transdisciplinary approach to teaching the topic of evolution that integrates data, examples, and graphics from various science disciplines (biology, geology, chemistry, and physics) and other academic disciplines (e.g., sociology, English, history, public health, mathematics, and economics).

Mr. Smith now begins the unit by introducing the concept of evolution through a story set in the historical context of Charles Darwin's journey on the HMS Beagle. It is the first introduction to the central ideas in the theory of evolution by natural selection. However, this historical narrative is enriched with excerpts from Darwin's writings, and Mr. Smith dedicates some time to exploring the formal language used.

To illustrate the biological principles of evolution, Mr. Smith uses the example of the peppered moth during the Industrial Revolution. Students analyze genetic data and create phylogenetic trees to understand the relationships between different species. They also examine the fossil record and geological time scale to see how organisms have changed over time. In chemistry, students learn about the molecular basis of evolution, including DNA, RNA, and protein synthesis. These hands-on activities help students understand how mutations in genetic material can lead to evolutionary changes.

To connect evolution to sociology, Ms. Smith examines the social implications of evolutionary theory with case studies. Students are organized into groups with specific roles for each member. They discuss how evolutionary concepts have been used to justify social hierarchies and discrimination, and the ethical considerations of such applications. They also explore the impact of human evolution on social structures and behaviors. Mr. Smith provides guiding questions for exploration and discussion for each role within the groups.

In connection with a public health case study, students learn about pathogens' evolution and antibiotic resistance development. They analyze case studies of disease outbreaks and explore the role of evolution in shaping human health. Students are encouraged to be creative in developing solutions for how public health can prevent disease outbreaks. This transdisciplinary approach helps students understand the practical applications of evolutionary theory in addressing public health challenges.

Mathematics is integral to understanding evolutionary processes. Mr. Smith teaches students how to use mathematical models to simulate population dynamics and predict evolutionary outcomes. They learn about statistical methods for analyzing genetic data and the mathematical principles underlying natural selection and genetic drift. Students must solve problems using data related to real-world scenarios. For example, Mr. Smith explores the economic implications of evolution. Students discuss how evolutionary principles can be applied to understand economic behavior and market dynamics. They examine case studies of industries, such as agriculture and biotechnology, that are influenced by evolutionary processes.

Case Discussion

By using a transdisciplinary approach, Mr. Smith provides his students with a holistic understanding of evolution from multiple perspectives due to the impact this understanding has on the many different aspects of the real world. The PCK components used by Mr. Smith in this transdisciplinary example include using multiple representations in the form of case studies and activities. His depth of knowledge of the topic allowed him to plan the different activities, labs, and case studies. The formation and use of groups with group members playing different roles reflects his general pedagogy and classroom management. Making the case studies parallel to real-life issues shows that the topic of evolution is a socio-cultural issue. This method enhances their knowledge of the subject and helps them develop critical thinking, communication, and problem-solving skills. By integrating multiple disciplines, students gain a deeper appreciation for the complexity and interconnectedness of the natural world.

$$\begin{aligned} PCK(\text{Evolution}|\text{all disciplines}) \\ = PCK(\text{evolution}|B \otimes G \otimes P \otimes C) + PCK(\text{evolution}|\text{history}) \\ + PCK(\text{evolution}|\text{history} \otimes \text{language}) + \dots \end{aligned}$$

Mr. Smith's transdisciplinary PCK about evolution is equivalent to the PCK about evolution at the interface of all science domains, see discussion of Toy Case 4a, and the PCK about evolution at the interface of all the different and relevant subject areas (e.g., history, language, and public health).

Toy Case 5 Transdisciplinary PCK: Nuclear Bomb

Case

In a physics class, a teacher introduces the working principles of a nuclear bomb at the end of a lesson series on nuclear energy. Nuclear weapons are a particular interest of the teacher, and he thinks it is important to introduce students to more than the physics. The class reads a text on the Hiroshima and Nagasaki nuclear bombs. They read out loud, taking turns, in line with a school-wide effort to increase literacy (the principal is adamant that 'every teacher is a language teacher'). The text introduces the students to mass-to-energy conversion and the historical context leading to the dropping of the bombs. Throughout the reading of the text, the physics teacher probes the students' understanding of the technical and historical terms in the text. After reading the text, the teacher assigns homework where the students look for additional resources online. He assigns specific topics to research to each student based on questions from the students and some additional questions from the teacher. He also provides online resources (e.g., articles, databases, and authors) as a starting point. In the following lesson, the class discusses the effect of the bombs on the cities and the surrounding environment (fauna and flora), ethical considerations, and current global politics.

Case Discussion

A teacher can't be well-versed in all disciplines. However, for a single topic, a teacher is likely to develop enough knowledge (both CK and PCK) to introduce students to (parts of) biology, earth sciences, history, philosophy, political science, etc. The teacher introduces the effects of radiation on biological tissues as an impact from the bombs. The connections between the historical events and the social geography are presented through accounts from the diaries of survivors that students read and share. The teacher extends the discussion about nuclear bombs by presenting current, potential, global, and political case studies. In this toy case, the physics teacher has developed transdisciplinary PCK about nuclear bombs *NB*.

$$PCK_{T_P}(NB) = PCK(NB|P) + PCK(NB|B) + PCK(NB|G) + PCK(NB|history) + PCK(language|NB)$$

The transdisciplinary PCK about nuclear bombs of the physics teacher is equivalent to the PCK about nuclear bombs in physics, biology, and geosciences plus the PCK about nuclear bombs in other subject areas (e.g., history and language). This transdisciplinary approach also develops students' information and technology literacy by supporting students in looking for information online and their critical thinking about the role of nuclear weapons in global politics.

Discussion

The toy cases were developed to show the impact of an approach to science education that embraces inter- and transdisciplinarity and develops 21st-century skills in science teachers' PCK. The cases show how different topics can be taught, ranging from a multidisciplinary approach to an interdisciplinary approach to a transdisciplinary approach. Three themes emerged from the development of the toy cases. The first theme that emerged early in the development was that inter- and transdisciplinary science education offers great opportunities to develop students' 21st-century skills. It could also be read inversely, i.e., the influence of 21st-century skills forces science teachers to broaden their science background among science domains and be able to apply the new interdisciplinary content knowledge to real-world applications. Either way, the combination of inter- and transdisciplinary science and 21st-century skills forces science teachers' PCK to expand. Adding inter- and transdisciplinary PCK as separate grain sizes of PCK is the second theme. Finally, as a byproduct of this newly expanded PCK, these toy cases reflect how preservice and in-service science teachers should be trained. These three themes are discussed in more detail below.

PCK Components and 21st-century Skills

The toy cases presented introduced inter- and transdisciplinary aspects of science education, and thus also of science teachers' PCK. Moreover, the cases also introduced different ways to include the 21st-century skills in science education. Table 3 summarizes which elements of PCK and which 21st-century skills are present in the different cases. The cases are complementary as each shows a different aspect of introducing inter- and transdisciplinarity into science education. Moreover, the cases are presented in an order that adds layers of increasing complexity.

The first introductory case, where physics and biology teachers introduce students to optics in the eye, shows how classical science education already has elements of multidisciplinary. The teachers remain firmly within their own discipline but are aware of their colleagues' content. Indeed, a teacher should have and apply curriculum knowledge, both vertically (What content do students encounter in other years?) and horizontally (What content do students see in other courses?). However, it is worthwhile to expand from multidisciplinary to interdisciplinarity. This happens in the second case by illustrating how a teacher team guides students through a (typical) STEM project.

While they started multidisciplinary, each respecting the boundaries of the disciplines, the project and students pushed them to adopt an interdisciplinary approach. The PCK components and the 21st-century skills used and exemplified in these cases are few. The interaction among the knowledge components and competencies lacks specificity.

The following two cases expand this notion even further to transdisciplinarity. Cases 3 and 4 illustrate how a team could approach topics in science in both an interdisciplinary and transdisciplinary way. For climate change (toy case 3), this is obvious and perhaps even expected. Climate change is, after all, a truly transdisciplinary problem. The increase in the number of PCK components and 21st-century skills increases due to the nature of the cases being interdisciplinary and transdisciplinary. However, the fourth case shows that it is also possible to develop a topic that is typically taught in a unidimensional way, here evolution, to an inter- and even transdisciplinary approach. It is by expanding the approach to be inter- and transdisciplinary that students' 21st-century skills could be developed and expanded. This would necessitate the teacher developing more PCK components.

The fifth case adds a final layer of nuance. Where in all previous cases several teachers work together, the final case shows that it is possible for a single teacher to work in a transdisciplinary way. As argued in the case discussion, it cannot be expected that a single teacher has in-depth CK and PCK for all topics. However, diving deep into a single topic and exploring it from different angles, whilst paying attention to 21st-century skills, could be enriching for both the students and the teacher.

Interdisciplinary and Transdisciplinary PCK

The societal issues and application of science content necessitate teachers to understand science content beyond their specialty and their impact on society. Science teachers at the secondary level should still have a depth of science content knowledge within their own domain. However, they should also be able to expand their science knowledge beyond their specialty area. The nature of societal issues and wicked problems forces citizens to understand and critically think about multiple variables that might influence how society solves problems. Toy cases 3a and b reflect the idea that science content, when related to a social issue, needs to be taught in an interdisciplinary and not a multidisciplinary manner. In many instances, teachers in a school will approach a collaborative project by siloing their respective subject areas and not communicating and sharing knowledge. As a multidisciplinary approach, science content is not understood or learned as relating to each other. If the science teachers in the interdisciplinary evolution toy case collaborate and show how the different science domains provide a holistic approach to an issue, the students would have an easier time applying the science principles to real-world applications. Science teachers and students benefit from the seamless integration and crossover of science content to help contextualize evolution and how it impacts and can be explained using multiple science domains.

In the transdisciplinary toy case for evolution, students take the next step in realizing how evolution applies to their lives. Learning is increased when science is personal or applicable. Applying science to real-world problems or examples solidifies the science content for many students. The transdisciplinary approach to learning science content provides the societal and personal contexts for enhanced learning. Science facts become real, and epistemic knowledge is solidified.

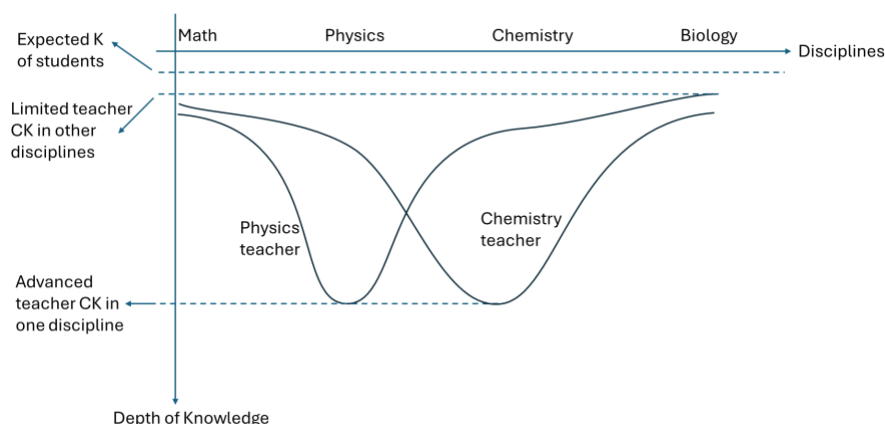
Table 3*Alignment of PCK components and 21st-century skills with toy cases*

	Case 1 Eye inter	Case 2 Health monitor inter	Case 3a CC inter	Case 3b CC trans	Case 4a Evol inter	Case 4b Evol trans	Case 5 Nuclear bomb trans
PCK components							
Representations	X		X	X	X	X	
Students' Understanding		X	X		X		
Subject Matter	X	X	X	X	X	X	X
Curriculum	X		X		X		
Assessment				X	X		
Environment/Context				X	X		X
Socio-cultural issues			X	X		X	X
General Pedagogy						X	X
Classroom Management		X				X	
21st-Century Skills							
Technology literacy		X	X		X		X
Media literacy			X	X	X		
Information literacy				X			X
Critical thinking			X	X	X		X
Problem-solving		X	X	X	X	X	
Creativity & Flexibility			X	X		X	
Collaboration & Communication		X	X	X		X	

So, given the premise that societal challenges require complex solutions, that science is becoming more inter- and transdisciplinary, and assuming that science education should or will follow suit, this evolution has implications for the knowledge of science teachers. As argued above, the teachers' content knowledge will have to expand. Figure 4 describes teacher content knowledge across domains, which is deepest in their own domain and is deeper than what is expected from students for neighboring domains. As teachers are expected to teach in a more interdisciplinary or transdisciplinary way, their content knowledge will need to deepen in multiple domains of science. However, they can and should still maintain the depth of their knowledge in their primary science domain. For example, a physics teacher will maintain deep content knowledge in physics but will also need to build a deeper insight into mathematics, chemistry, and biology without being expected to become a full expert in those neighboring domains. To be clear, the positioning of the disciplines on a single (horizontal) axis is a gross simplification of the relation between the disciplines. As such, the figure is schematic and should not be read as a true two-dimensional graph.

Figure 4

Disciplinary Knowledge of Science Teachers across Domains



Additionally, as the toy-cases illustrate, teachers need knowledge of how their domain, and the topics within it, are related to other domains within (inter) and beyond (trans) science. Moreover, and specifically for teachers, they must know mistakes students make in the secondary domain that are connected to their domain, they need to know how to connect both domains in different ways to support students in their understanding, they need to know which representations (with elements of both domains) allow this connection. In other words, the PCK of science teachers will also need to expand. Specifically, the model of Veal and Makinster (1999) that introduced PCK grain sizes should be expanded to include interdisciplinary (ID-PCK) and transdisciplinary (TD-PCK) PCK. Figure 5 shows a tentative and expanded framework for PCK based on Veal and Makinster (1999). It should be clear that teachers, of course, still need general, domain-specific, and topic-specific PCK. This expanded PCK will allow science teachers to understand and introduce students to how the topics, concepts, and issues relate to other domains of science and impact society and the real world. Science teachers must broaden their scope and understanding of their knowledge base to make learning more accessible and personal, thus providing students with the opportunities to develop 21st-century skills.

Figure 5 presents the remodeled grain size for PCK. This model reflects the external and societal pressures on science teachers due to socio-scientific issues necessitating an interdisciplinary and transdisciplinary approach to solving them. Science teachers must develop ID-PCK and TD-PCK to teach science appropriately while integrating 21st-century skills. Interdisciplinary PCK sits atop the domains of science, as all content and areas of science are blended rather than being individually assembled with discrete knowledge, as in a multidisciplinary approach. Transdisciplinary PCK resides atop the traditional school subjects and domains of knowledge. When solving more complex 21st-century socio-scientific issues, teachers must smoothly integrate other subject areas to prepare effective solutions for these issues. Both new knowledge bases require extending the knowledge and skills of the science teachers. The toy cases serve as a blueprint for understanding how ID-PCK and TD-PCK can be conceptualized in the preparation and development of excellent science teachers.

Figure 5

A Revised Framework for PCK based on Veal & MaKinster (1999)

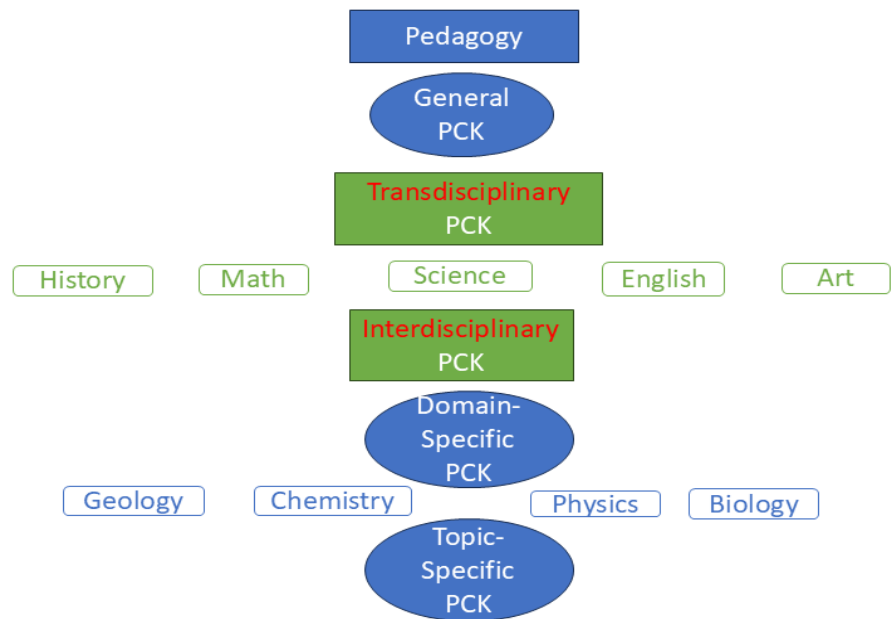
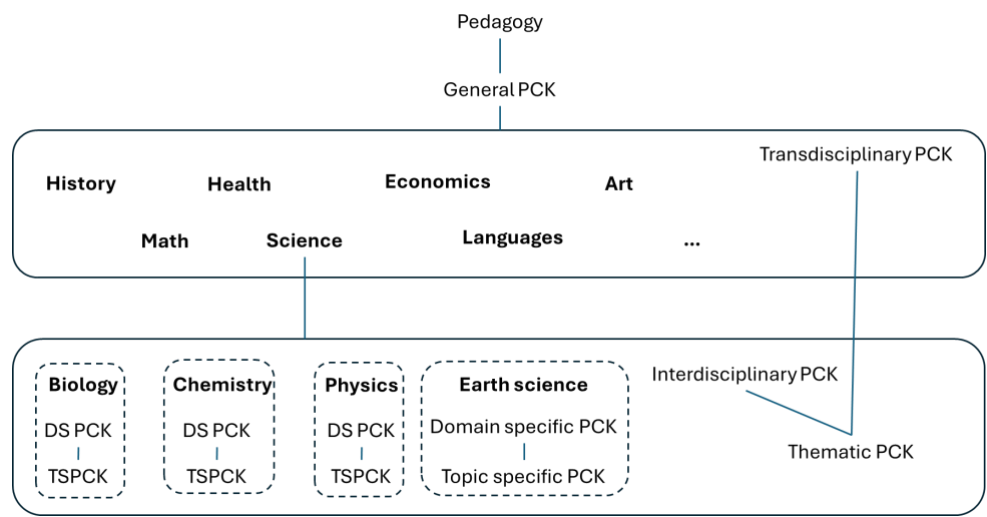


Figure 6 represents an explanatory figure for showing how ID-PCK and TD-PCK interact within and beyond the domains of science and including the knowledge of other subject areas. Interdisciplinary PCK encompasses the domains of science. The boundaries of the domains are depicted as dashed lines to reflect how all science domains contribute to solving the thematic problem or issue. This presents a potentially new idea of Thematic-PCK, which includes a broader vision of a theme that includes facts, concepts, domains, and extensions. Transdisciplinary PCK encompasses all school subjects and domains of knowledge.

Figure 6

An Explanatory Figure for PCK Grain Size



Professional Development

The impact of this new framework dictates that science teachers, new and old, must adapt their knowledge structures so that they can teach science differently. For example, how does a physics teacher present information about climate change? Climate change is a wicked problem that mandates a broad understanding of the science behind it and its potential impacts on different facets of society. For a science teacher to appropriately and logically teach a unit on climate change, this science teacher would have to have interdisciplinary PCK in which science facts and concepts are intertwined to present the scientific case for climate change (Toolin et al., 2022). If the science teacher wants to present the practical application, causes, and effects of climate change, they must develop and show how it influences other aspects of the students' lives with examples and applications to the economy, communication, public health, and history. By designing lesson plans in this transdisciplinary manner, science teachers should model how scientists and society attack wicked problems.

At the preservice level, budding new science teachers must be trained in how science can be positioned with other science domains and academic subjects. This may influence how preservice teachers are certified to teach (Hermann, 2016). For example, in the United States, each state has its own laws on how to certify middle and secondary science teachers. In most states, science teachers can be certified in one science domain (e.g., Chemistry, Biology, Physics, or Earth Sciences) or as a general science teacher. Applying this new framework to PCK necessitates that preservice teachers be certified as science teachers with strong science content knowledge in at least one domain but also having science content knowledge across other science domains. Science teachers will need to be certified with the new standards for teaching, while the idea of just being a chemistry or physics teacher will no longer be recognized. By contrast, in Flanders, Belgium, it is generally accepted that teachers need in-depth domain-specific PCK. Indeed, this paper does not negate the value of having dedicated physics or chemistry teachers, for example. Rather, the toy cases illustrate that it might no longer be sufficient for them to have only deep content knowledge within their own discipline. Instead, teacher training programs should require students to get at least some in-depth science content knowledge in other disciplines. Additionally, the teacher training programs themselves should emphasize developing, in addition to in-depth PCK in a single domain, PCK in other domains as well as inter- and transdisciplinary PCK.

Of course, this argument about preparing science teachers with ID-PCK and TD-PCK extends to the professional development and expectations of science teacher educators. As Abell (1997) stated, science teacher educators must also have PCK and all the associated knowledge bases that Shulman described. The argument extends to science teacher educators at the university, state, and district levels, in that they must also have ID-PCK and TD-PCK to prepare future science teachers.

Conclusion

This paper argues for two distinct and related points that can influence how science teachers use inter- and transdisciplinary topics or themes, including socio-scientific issues, in their classes while developing students' 21st-century skills. First, the grain size of PCK needs to be expanded to include Interdisciplinary PCK (ID-PCK) and Transdisciplinary PCK (TD-PCK). This results from how scientists look at societal issues and real-world problems. Teaching these concepts, known in many instances as wicked problems, necessitates a new teaching method that a science teacher must use. The 21st-century skills that students need influence the movement from isolated science content to an integrated approach. By developing these 21st-century skills, students will be more prepared to

understand and solve socio-scientific issues. The interdisciplinarity of science within science domains and with other academic subjects reflects how science should be taught and learned. Silos and isolated science content cannot continue to be the norm if students are prepared to advance society in the 21st-century.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

William R. Veal (vealw@cofc.edu) is a professor of teacher education and chemistry at the University of Charleston. His research interests focus on reform-based science education and its impact on preservice teacher accreditation. He also researches the intersection of pedagogical content knowledge and creativity in science teaching and learning.

Jan Sermeus (jan.sermeus@kuleuven.be) is an associate professor at KU Leuven and the Brussels Planetarium, researching astronomy education and science education. He is interested in students' understanding of astronomical concepts and their understanding of the nature of science.

References

- Abell, S.K. (1997). The professional development of science teacher educators: Is there a missing piece? *The Electronic Journal for Research in Science & Mathematics Education*, 1(4). <https://ejrsmc.icrsmc.com/issue/view/711>
- Al-Abdali, N. S., & Al-Balushi, S. M. (2016). Teaching for creativity by science teachers in grades 5-10. *International Journal of Science and Mathematics Education*, 14(2), 251–268. <https://doi.org/10.1007/s10763-014-9612-3>
- AniMed (2017), *How the Eye Works Animation - How Do We See Video - Nearsighted & Farsighted Human Eye Anatomy*. Retrieved March 6, 2025, from <https://www.youtube.com/watch?v=YcedXDN6a88>
- Barthel, R., & Seidl, R. (2017). Interdisciplinary collaboration between natural and social sciences—status and trends exemplified in groundwater research. *PloS one*, 12(1), <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0170754>.
- Bielik, T. & Krell, M. (2025). Developing and evaluating the extended epistemic vigilance framework. *Journal of Research in Science Teaching*, 62(3), 869–895. <https://doi.org/10.1002/tea.21983>
- Bliss, A. C. (2019). Adult Science-Based Learning: The Intersection of Digital, Science, and Information Literacies. *Adult Learning*, 30(3), 128–137. <https://doi.org/10.1177/1045159519829042>
- Bornmann, L., Haunschild, R., & Mutz, R. (2021). Growth rates of modern science: a latent piecewise growth curve approach to model publication numbers from established and new literature databases. *Humanities and Social Sciences Communications*, 8(1), 1-15. <https://doi.org/10.48550/arXiv.2012.07675>
- Carlson, J., Daehler, K. R., Alonzo, A. C., Barendsen, E., Berry, A., Borowski, A., Carpendale, J., Chan, K. K. H., Cooper, R., Friedrichsen, P., Gess-Newsome, J., Henze-Rietveld, I., Hume, A., Kirschner, S., Liepertz, S., Loughran, J., Mavhunga, E., Neumann, K., Nilsson, P., Park, S., Rollnick, M., Sickel, A., Schneider, R. M., Suh, J. K., van Driel, J., & Wilson, C. D. (2019). The refined consensus model of pedagogical content knowledge in science education. In A. Hume, R. Cooper, & A. Borowski (Eds.), *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 77–92). Springer Nature Singapore Pte Ltd. https://doi.org/10.1007/978-981-13-5898-2_2

- Chan, K. Van Driel, J., Henze, I., Stammes, H., & de Vries, M. (August 24, 2024). *History and Development of PCK* [Meaning of teacher expertise and PCK for 21st century science education]. 21st century science education: What does it mean for teachers? Leiden, The Netherlands.
- Chang, C. C., & Chen, Y. (2020). Cognition, attitude, and interest in cross-disciplinary i-STEM robotics curriculum developed by thematic integration approaches of webbed and threaded models: A concurrent embedded mixed methods study. *Journal of Science Education and Technology*, 29(5), 622-634. <https://doi.org/10.1007/s10956-020-09841-9>
- Chen, D. (2021). Toward an understanding of 21st-century skills: From a systematic review. *International Journal for Educational and Vocational Guidance*, 23, 275-294. <https://doi.org/10.1007/s10775-021-09511-1>
- Cochran, K. F., DeRuiter, J. A., & King, R. A. (1993). Pedagogical content knowing: An integrative model for teacher preparation. *Journal of Teacher Education*, 44(4), 263-272. <https://doi.org/10.1177/0022487193044004004>
- Deák, C., Kumar, B., Szabó, I., Nagy, G., & Szentesi, S. (2021). Evolution of new approaches in pedagogy and STEM with inquiry-based learning and post-pandemic scenarios. *Education Sciences*, 11(7), 319. <https://doi.org/10.3390/educsci11070319>
- English, L. D. (2023). Ways of thinking in STEM-based problem solving. *ZDM—Mathematics Education*, 55(7), 1219-1230. <https://doi.org/10.1007/s11858-023-01474-7>
- Facione, P.A. (1990). *Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations*. American Philosophical Association: Newark, Delaware. <https://eric.ed.gov/?id=ED315423>
- Friedrichsen, P. J., van Driel, J. H., & Abell, S. K. (2011). Taking a closer look at science teaching orientations. *Science Education*, 95(2), 358-376. <https://doi.org/10.1002/sce.20428>
- Furner, J. M. (2024). The best pedagogical practices for teaching mathematics revisited: Using math manipulatives, children's literature, and GeoGebra to produce math confident young people for a STEM world. *Pedagogical Research*, 9(2), em0193. <https://doi.org/10.29333/pr/14194>
- Galili, I. (2019). Towards a refined depiction of nature of science: Applications to physics education. *Science & Education*, 28, 503-537. <https://doi.org/10.1007/s11191-019-00042-4>
- Gess-Newsome, J. (2015). A model of teacher professional knowledge and skill including PCK: Results of the thinking from the PCK summit. In A. Berry, P. Friedrichsen, & J. Loughran (Eds.), *Re-Examining pedagogical content knowledge in science education* (pp. 28–42). Routledge. <https://www.taylorfrancis.com/chapters/edit/10.4324/9781315735665-4/model-teacher-professional-knowledge-skill-including-pck-gess-newsome-julie>
- Gleason, D. W. (2020). The humanities meet STEM: Five approaches for humanists. *Arts and Humanities in Higher Education*, 19(2), 186-206. <https://doi.org/10.1177/1474022218806730>
- González-Pérez, L.I.; & Ramírez-Montoya, M.S. (2022). Components of education 4.0 in 21st century skills frameworks: Systematic review. *Sustainability*, 14, 1493. <https://doi.org/10.3390/su14031493>
- Grossman, P. L. (1990). *The making of a teacher: Teacher knowledge and teacher education*. New York: Teachers College Press.
- Hermann, R.S. (2016). Elementary education majors' views on evolution: A comparison of undergraduate majors understanding of natural selection and acceptance of evolution. *The Electronic Journal for Research in Science & Mathematics Education*, 20(6). <https://ejrsme.icrsme.com/article/view/16305/10375>

- Hilton, M. (2010). Exploring the Intersection of Science Education and 21st Century Skills: A Workshop Summary.
https://doi.org/https://www.nap.edu/catalog.php?record_id=12771
- Hubber, P., Widjaja, W., & Aranda, G. (2022). Assessment of an interdisciplinary project in science and mathematics: Opportunities and challenges. *Teaching Science*, 68(1), 13–25.
<https://doi.org/https://asta.edu.au/resources/teaching-science-journal/>
- Humphreys, D., & Kelly, P. (2014). *How Liberal Arts and Science majors Fare in Employment: A Report on Earnings and Long-term Career Paths*. Washington, DC: Association of American Colleges and Universities.
- Jin, Q., Kim, M., & Wong, S. (2023). Elementary students' online information problem solving (IPS) in a science classroom. *Alberta Journal of Educational Research*, 69(1), 1–19.
<https://doi.org/https://journalhosting.ucalgary.ca/index.php/ajer/article/view/71768>
- Kagan, D. M. (1993). Contexts for the use of classroom cases. *American Educational Research Journal*, 30(4), 703–723. <https://doi.org/10.3102/00028312030004703>
- Karaca-Atik, A., Meeuwisse, M., Forgievski, M., & Smeets, G. (2023). Uncovering important 21st-century skills for sustainable career development of social sciences graduates: A systematic review. *Educational Research Review*, 39, 1–15.
<https://doi.org/10.1016/j.edurev.2023.100528>
- Klein, J.T. (2010). The taxonomy of interdisciplinarity. In Frodeman, R., J.T., Klein & Mitcham, C., *The Oxford Handbook of Interdisciplinarity* (pp. 15–30). Oxford University Press.
- Kotsiou, A., Fajardo-Tovar, D.D., Cowhitt, T., Major, L. & Wegerif, R. (2022). A scoping review of future skills frameworks. *Irish Educational Studies*, 41:1, 171–186,
<https://doi.org/10.1080/03323315.2021.2022522>
- Luczak, J. (2017). Talk about toy models. *Studies in History and Philosophy of Science Part B: Studies in history and philosophy of modern physics*, 57, 1–7. <https://doi.org/10.1016/j.shpsb.2016.11.002>
- Magnusson, S., Krajcik, J., & Borko, H. (1999). Nature, sources, and development of pedagogical content knowledge for science teaching. In J. Gess-Newsome & N. G. Lederman (Eds.), *Examining pedagogical content knowledge: The construct and its implications for science education* (pp. 95–132). Kluwer Academic.
- Marcoci, A., Thresher, A. C., Martens, N. C., Galison, P., Doeleman, S. S., & Johnson, M. D. (2023). Big STEM collaborations should include humanities and social science. *Nature Human Behaviour*, 7(8), 1229–1230. DOI: [10.1038/s41562-023-01674-x](https://doi.org/10.1038/s41562-023-01674-x)
- Marks, R. (1990). Pedagogical content knowledge: From a mathematical case to a modified conception. *Journal of Teacher Education*, 41(3), 3–11. DOI: [10.1177/002248719004100302](https://doi.org/10.1177/002248719004100302)
- Martin, J. (2018). *Skills for the 21st century: Findings and policy lessons from the OECD survey of adult skills*. (Working Paper No. 166). OECD Publishing. <https://doi.org/10.1787/96e69229-en>.
- Osborne, J. (August 19, 2024). *21st Science Education: Desired Student Outcomes* [21st Century Science Education, Student Outcomes, and Consequences for Teachers]. 21st Century Science Education: What does it mean for Teachers?, Leiden, The Netherlands.
- Redish, E. (2021). Using math in physics: 4. Toy models. *The Physics Teacher*, 59(9), 683–688.
<https://doi.org/10.1119/5.0025840>
- Rodriguez, S., Allen, K., Harron, J., & Qadri, S. A. (2019). Making and the 5E Learning Cycle. *Science Teacher*, 86(5), 48–55.
https://doi.org/https://www.nsta.org/publications/browse_journals.aspx?action=issue&id=116215
- Rollnick, M., & Mavhunga, E. (2015). The PCK Summit and its effect on work in South Africa. In *Re-examining pedagogical content knowledge in science education* (pp. 135–146). Routledge.
- Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, 15(2), 4–14. [http://dx.doi.org/10.3102/0013189X015002004](https://doi.org/10.3102/0013189X015002004)

- Shulman, L. S. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1–23. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Sikand, M., Mazzatenta, C., Wong, K., Bush, J., & Socha, A. M. (2021). Two year community: Sustainability, energy, and the green economy: An interdisciplinary course on environmental sustainability and life cycle analysis. *Journal of College Science Teaching*, 50(3), 8-16. <https://doi.org/10.1080/0047231X.2021.12290502>
- Toolin, R.E., Jorgenson, S., & Ratmeyer, S. (2022). STEM touchstones for teacher professional learning: An analysis of teacher content and pedagogical content knowledge in a place-based professional development program. *Electronic Journal for Research in Science & Mathematics Education*, 26(2), 32-55. <https://ejrsme.icrsme.com/article/view/21986/14614>
- Urone & Hinrichs (2025), *College Physics 2e*, <https://openstax.org/books/college-physics-2e/pages/25-6-image-formation-by-lenses>.
- Van Noorden, R. (2015). Interdisciplinary research by the numbers. *Nature*, 525(7569), 306-307. <https://doi.org/10.1038/525306a>
- Veal, W.R., & Kind, V. (2025, August 25-29). *Connecting creativity with PCK: Creating a framework for science teacher education* [Paper presentation]. European Science Education Research Association Biannual Conference, Copenhagen, Denmark.
- Veal, W.R., & MaKinster, J.G. (1999). Pedagogical content knowledge taxonomies. *The Electronic Journal for Research in Science & Mathematics Education*, 3(4). <https://ejrsme.icrsme.com/article/view/7615>
- Weart, S. (2013). Rise of interdisciplinary research on climate. *Proceedings of the National Academy of Sciences - PNAS*, 110(Supplement 1), 3657–3664. <https://doi.org/10.1073/pnas.1107482109>
- Widya, Rifandi, R., & Laila Rahmi, Y. (2019). STEM education to fulfil the 21st century demand: A literature review. *Journal of Physics. Conference Series*, 1317(1). <https://doi.org/10.1088/1742-6596/1317/1/012208>

Beyond Problem Solving: Designing Concept-Building Worked Examples for Chemistry Education

Markus Emden 

University of Teacher Education Zurich

Elke Sumfleth 

University of Duisburg-Essen

Katrin Schuessler 

University of Duisburg-Essen

ABSTRACT

Science teaching in schools attends to two major varieties of knowledge: procedural and conceptual. While studies on how to support procedural learning in science have abounded over the past two decades, the situation regarding conceptual learning has been markedly calmer. Yet, as conceptual learning is a necessary requisite for any learning, viable methods to support students in their learning are needed. The present article reviews two approaches to conceptual learning from texts—expository and refutation—to identify potential. It merges these findings with research on learning with worked examples to suggest the design of an innovative format of learning materials: concept-building worked examples. A set of these is designed relying on the Cognitive Theory of Multimedia Learning, while capitalizing on suggestions from research on learning with texts. These concept-building worked examples allow students to develop scientific concepts at their own pace—something hardly accomplished by traditional textbooks. The materials are meant to be used for self-study. They have been published online free of charge. Overall, 34 concept-building examples help students in lower secondary education (ISCED-level 2) learn nine chemical concepts. These concepts are developed stepwise in narrative texts between characters that students can identify with. The worked examples introduce expert and lay characters, where the latter serve as stand-ins for students sharing the character's preconcepts and allow for an explicit take-up of these. Self-explanation prompts and concept-related tasks ensure active processing of the worked examples' content. Illustrative examples are provided to renew discourse on concept-learning in chemical education.

Keywords: chemistry education, concept-learning, worked examples, refutation text, learning from text

Introduction

Recent science education has devoted an extended focus on domain-specific problem-solving abilities, be it in the form of scientific inquiry (National Research Council [NRC], 1996) or in the form of scientific and engineering practices (NRC, 2011). This development is mirrored internationally in curricula, and much research has addressed novel and improved ways to facilitate student growth in their respective problem-solving abilities. One argument for this emphasis on “doing science”—as opposed to, *e.g.*, “knowing science” (Hodson, 2014)—is that the conceptual knowledge base of science

is expanding more rapidly than schools could ever manage to accommodate within their limited syllabi. Therefore, the idea is to introduce students to prototypical ways of thinking and acting in science, as this could empower them to generate their own scientific knowledge as well as to reflect on how authentic science generates its knowledge.

Science education has come to orient itself towards three visions of scientific literacy with different emphases: Vision I scientific literacy might be termed the “vocational approach” enabling students to take up science related work or studies after school; Vision II scientific literacy is a “pragmatic approach” that wants students to realize where science features helpfully in their everyday lives (both: Roberts, 2007; Roberts & Bybee, 2014); Vision III scientific literacy acknowledges that modern societies are dependent on scientific resources, so that members of these societies need to have command over some “democratic scientific literacy” that allows them to evaluate and choose between several solution suggestions, to distinguish fake news from real ones, and to identify deliberate disinformation concerning scientific facts (Sjöström & Eilks, 2018). Each of these visions relies on conceptual scientific knowledge that students are expected to master. While much work has been done in the past decades on supporting students’ progress in the processes of science, it appears as if it has been missed to some degree that science is an extremely theory-laden endeavor (Lederman, 2007). Students cannot rely solely on being introduced to “doing science,” but their enactment of science needs to fall back on a conceptual basis, *i.e.*, they need to “know science”, too. Klahr and Dunbar (1988) have shown in their Scientific Discovery as Dual Search approach that the generation of knowledge necessarily draws on available conceptual (search hypothesis space) and procedural knowledges (search experiment space). If either of these knowledges is atrophied or non-existent, there can be no scientific discovery.

So, while teachers and science educators are laudably turning towards a flexibilization of science literacy by emphasizing exemplary and transferable procedural abilities and skills, they must not overlook the likewise essential task to lay conceptual foundations on which these abilities can work (see also Sinatra & Broughton, 2011). The present article aims to introduce an innovative task format that is devoted to achieving conceptual change at students’ pace. The development of self-learning materials is considered to be a crucial element in contemporary chemistry education, which is chronically short of time and faces challenges of cumulative learning as few other school subjects do. Disregarding student needs or barring them from catching up on their own will leave many students stranded when it comes to matters of scientific literacy. Developing self-learning materials is time-consuming and teachers cannot be expected to each take on this task—especially with regard to extensive sets of materials. Tested materials, therefore, need to be introduced in a way that science educators and teachers from other locales can adopt them (considerably) easily. The current article introduces such a set of tested materials for self-study of fundamental concepts in chemistry. It hopes to encourage colleagues from different contexts to identify potentials how to bring the materials into their classrooms, how to adopt them—*e.g.*, using artificial intelligence technologies—, or how to develop their own learning materials building on the same design principles.

The current article will therefore start by relating what we understand about how concepts are learned. It will, then, go on to discuss how concept-learning can be supported by learning with texts—it discusses learning from expository texts and refutation texts before turning to learning with worked examples. Each of these perspectives will contribute insight into learning concepts either as challenges or as potentials. These will be employed to design an innovative instructional text format supporting concept learning in chemistry at the lower secondary level (ISCED-level 2). Research findings on its effectiveness are briefly introduced, and perspectives for adoption of the worked examples are discussed.

Theoretical Background

The theoretical background will introduce relevant theory and research from science education on concept-learning, learning from texts, and learning from worked examples. It aims at highlighting the benefits and challenges of each approach regarding student learning. “Lessons learned” from these lines of research (Table 1) will serve to gauge hitherto unchanneled synergies and inform the design of an innovative format: concept-building worked examples.

Scientific Concepts in International Curricula

Most international curricula are oriented towards a limited number of crucial scientific concepts that need to be introduced. The NRC-standards (NRC, 2011) refer to these as (a) Core Ideas, *i.e.*, central domain-specific concepts (*e.g.*, evolution, chemical equilibrium), and (b) Crosscutting Concepts, *i.e.*, those concepts spanning scientific domains (*e.g.*, energy). Another example are German educational standards for chemistry education (KMK, 2024) that give these basic concepts: (1) composition and properties of matter and its particles, (2) chemical reaction, (3) energy changes. Within this framework, any chemical phenomenon can be approached from each of the three perspectives, which, in turn, can be further differentiated, *e.g.*, introducing donor-acceptor concepts in the scope of chemical reactions from the second basic concept.

Still, most international curricula specify—across these core ideas or basic concepts—more comprehensive canons of conceptual scientific knowledge that should be introduced (*e.g.*, atomic models, acid-base reactions, redox reactions). These topics form the backbone or narrative line for developing science classes in which students can also develop their procedural abilities. A major problem in this endeavor is that of too little classroom time—too frequently introductions by teachers need to move at a quicker pace than is beneficial for all the students (van Vorst, 2018). Some might even get lost and have little chance to find their way back by themselves. This could prove problematic in a subject such as chemistry, where understanding one concept often is prerequisite for learning a new one.

Concept-learning

Concepts can theoretically be described as mental representations of fundamental knowledge structures (Krajcik & Shin, 2023). One can distinguish between psychological concepts and cognitive concepts, where the former term relates to more subjective, intrapersonal mental representations and the latter to public, interpersonal representations (Gilbert & Watts, 1983; Rusanen & Pöyhönen, 2013). Formal education aims at developing students’ preexisting psychological concepts (*i.e.*, preconceptions) to align better with the consensual, definitory concepts held in a scientific discipline—the process for this development has been referred to as Conceptual Change for roughly the past four decades (Krajcik & Shin, 2023; Posner et al., 1982; Scott et al., 2007; Strike & Posner, 1982). This can be regarded as a fundamental goal of science education and is, thus, recorded in Table 1, which juxtaposes the most essential aspects drawn from the theory review to inform the design of an innovative format for self-learning.

Learning a concept involves adding to existing cognitive structures and qualifies as “meaningful learning” as opposed to routine-facilitating rote-learning (Ausubel, 1969). When introducing a new concept, some fundamental rules that cluster around the idea of crucial features of a concept should be obeyed (*e.g.*, K. Johnson & Bulla, 2021; Kean, 1982; Layng, 2019; Tennyson & Park, 1980):

- Each concept is characterized by a set of critical features that pertain to each of its cases (*e.g.*, metals are oxidized rather than reduced, metals show metallic shine, metals are good electrical and thermal conductors)—sometimes one concept may embrace several sets of critical features as, for example, in coexisting acid-base-concepts (Kean, 1982).
- Several cases of the same concept may be differentiated from each other by sets of variable features (*e.g.*, noble metals are harder to oxidize, base metals react readily with water to form lye, thermal and electrical conductivity are functions of crystalline structure).
- Several examples of the concept must be introduced that exhibit the critical features (*e.g.*, gold as a noble metal, sodium as a base metal, aluminum *vs.* copper regarding conductivity).
- Non-examples violating the concept features must be provided (*e.g.*, oxygen as an electric and thermal insulator, sulfur that does not have a shine, reduction reactions with chlorine).

However, concept-learning need not restrict itself to introducing new concepts but can also extend existing ones (Scott et al., 1992). This is especially important in science where concepts do not exist in isolation but interact and co-define each other (Kintsch, 1988) or combine to form principles (Gagné, 1965). Also, entities that are colloquially referred to as concepts are, in fact, formed by related sets of concepts (Herron et al., 1977): the metals example above draws concurrently on numerous individual concepts (*e.g.*, oxidation, conductivity). This sublime complexity of concepts is characteristic of science and therefore relevant for science education, not least when addressing preconceptions (Asterhan & Resnick, 2020). The viability of the above sketched method of introducing a new (and simple) concept is, thus, put into question. There appears to be no valid description of how complex concepts develop (Rusanen & Pöyhönen, 2013) as well as there appears to be no viable cognitive model to capture conceptual change processes (Clement, 2013).

Science teachers have traditionally relied on one crucial means to support students' concept-learning: textbooks (Bergqvist & Chang Rundgren, 2017; Martin et al., 2012; Thompson et al. 2023). At the same time, it appears to be increasingly problematic for student learning to rely on science textbooks (C. E. Johnson & Boon, 2023; Kahveci, 2010). Standard science textbooks are most often not written for those students with little prior knowledge, as they often exhibit a linguistic complexity and factual density that defies study by naïve readers (Mikkilä-Erdmann, 2002). And even teachers appear to be skeptical of them, considering that the most reported use of textbooks is for teachers' preparation but not for student learning (*e.g.*, Thompson et al., 2023).

Learning from Text

Much of student learning is mediated through written text, from which students are required to (re-)construct conceptual knowledge. Several formats of text have been introduced for this purpose, each having advantages and disadvantages. The following section discusses three formats for text-based learning opportunities: expository text, refutation text, and worked examples.

Expository text in textbooks

Texts that are meant to convey information rather than to entertain are called “expository text”, a classic example of which is the textbook (Mar et al., 2021). The textbook is a traditional student resource for learning, which students consider to be reliable and accurate (Thompson et al., 2023)—even if they are known to rarely use textbooks spontaneously (Guzzetti et al., 1997). Several studies have investigated science textbooks with regard to linguistic features; they consistently find that texts in science textbooks tend to be too difficult for the addressed readership due to complicated syntax and inflated vocabulary (Guzzetti et al., 1997; Pyburn & Pazicni, 2014). The suitable age that

conventional text readability indices estimate for a given science textbook appears to be one to two years above students biological ages (Calhoun & Rubba, 1993; Puspaningtyas et al., 2020; Soyibo, 1996). Although some unavoidable, linguistic peculiarities of science language (*e.g.*, passive voice and an increased frequency of compound nouns) are responsible for inflated estimation of the conventional readability indices (Calhoun & Rubba, 1993), one must not dismiss the issue easily. Some researchers attribute the rise of learning difficulties in a departure from more narrative-oriented texts when starting science education at the secondary level (Fang, 2006), while narrative texts seem to benefit concept-learning at the primary level (Guzzetti et al., 1993). Note that readability studies—which analyze a given text only structurally—have more recently given way to “comprehensibility” studies (Meyer & Pietzner, 2022; Vojříř & Rusek, 2019) that investigate how students interact with a given text. Students’ interactions can be supported (*e.g.*, Leopold & Leutner, 2012) so that a structural, linguistic criterion such as readability need not be an insurmountable obstacle. Comprehensibility support can take two forms—top-down (readers infer sense) or bottom-up (text guides sensemaking; Meyer & Pietzner, 2022)—suggesting that the actual text—notwithstanding interactions with the readers—exerts an influence on comprehension. Therefore, textual complexity continues to be a bottleneck that needs to be taken into consideration when learning in science (Safari et al., 2017). When science texts become incomprehensible to students, they lose their most crucial function. Unsurprisingly, there is research that shows that teachers are skeptical of the potential a textbook has for student learning (Beerenwinkel & Gräsel, 2005; Thompson et al., 2023). Students, when asked, express a wish for texts to be comprehensible, interesting, and rich in examples (Knecht & Najvarová, 2010). Authors suggest that interest could be furthered “by an attractive style of writing that motivates students toward further reading and deeper thought” (Knecht & Najvarová, 2010, p. 10). Having said this, there is also conflicting evidence that students do not prefer narrative texts (Guzzetti et al., 1993; Guzzetti et al., 1997), or that they even prefer expository actualizations of a conceptual change text format—refutation text (see below)—over its *narrative* variation (Hynd, 2001).

Common science texts tend to be heavy on technical terminology, thereby creating a density of information that makes it harder for students to realize where their preconceptions are no longer adequate and how they could align better with a scientific concept (van Loon et al., 2015). Science textbooks have been criticized for introducing more technical terms in a single reading than foreign language teaching would suggest in its domain (Groves, 1995), or for being less readable than textbooks for English as a Second Language for the same age group (Hu et al., 2021). While this situation has ameliorated to some degree, it still was shown to be, on the whole, valid twenty years later (Groves, 2016). Moreover, the introduction of new terms rarely supports a systematic construction of concepts (Härtig, 2014). This all contradicts a tacitly held, valid demand that science textbooks should be suited for student learning (Lodge, 2021), for example, when trying to catch up. Mikkilä-Erdmann (2002) advocates for science texts to actively support concept-building instead of collecting facts; she favors metaconceptual text that gives “direct comment on how to understand a phenomenon” (p. 343). Essentially, this calls for a different structuring of science texts that is better aligned with students’ needs when attempting conceptual change in learning.

Table 1 collects Lessons Learned from different branches of science education research to inform an innovative format for supporting student learning. Summarizing from the above, essential design suggestions for expository texts in textbooks are: (1) reduce syntactic complexity, (2) use fewer technical terms, (3) use technical terms to develop a concept, (4) employ narrative text for motivational reasons.

Refutation text

An alternative to expository text from traditional textbooks for supporting students in

developing their preconceptions into scientific concepts is conceptual change text. One such format is refutation text that differs from standard expository science text insofar as it explicitly addresses specific preconceptions and then corrects them (*e.g.*, Atasoy et al., 2009; Guzzetti et al., 1997). Relating explicitly to the preconcept can take on different guises: it may be realized in a direct contrast, *i.e.*, by explicitly stating the preconcept and its inadequacy, or through creating discrepant events, *i.e.*, through promoting dissonance by experiencing a contradictory phenomenon and introducing a controversial question (Clement, 2013). Refutation text may start from scratch by creating a cognitive conflict or by extending pre-existing ideas (Scott et al., 1992). These texts are especially valuable for students with little prior knowledge (Ariasi, 2012; Diakidoy et al., 2011; Diakidoy et al., 2016).

It is assumed that learning from refutation texts supports concept-learning by attending to some cognitive principles that subconsciously frame concept-building and cannot be avoided, so one should aim to make use of them (Kendeou et al., 2014):

- (1) encoding principle: Once information has been stored in long-term memory, it cannot be undone and might resurface at any time.
- (2) passive activation principle: Stimuli in working memory may unconsciously draw up formerly encoded information that can interfere.
- (3) coactivation principle: Previously encoded information (preconcept) and updated information need to be juxtaposed, *i.e.*, they need to be present at the same time in working memory if the one is supposed to 'overwrite' the other.
- (4) integration principle: Preconceptual information needs to construct a coherent whole together with 'new' information.
- (5) competing activation principle: As a consequence of the encoding principle, new stimuli in working memory are prone to activate both the incorrect and correct information, at the same time.

The crucial principles to exploit in teaching appear to be the coactivation-principle—by ensuring that a preconcept is triggered at the same time when the scientific concept is introduced—and the integration-principle—by explicitly advising students how to construct a better aligned concept from both the present concepts. Standard expository text does not usually call out specific preconceptions and therefore cannot live up well to the challenge of conceptual change.

Apart from making use of the above principles, another reason for refutation texts' efficiency could lie in extended time-on-task compared to cursory reading of expository texts, as evidence from eye-tracking studies suggests (Ariasi, 2012).

Conceptual change learning from refutation texts benefits from providing curricular coherence and from introducing meaningful and engaging contexts, which allow students to link their own ideas and experiences (Krajcik & Shin, 2023). More added value lies in integrating analogies and graphics into refutation text (Danielson et al., 2016). Furthermore, the coactivation and integration processes may benefit from encouraging students to reflect on their learning process (Krajcik & Shin, 2023), and from requiring elaborated explanations (Danielson et al., 2016; Kendeou et al., 2014). Where possible, the implementation of analogies promises to support conceptual change as analogies rely on everyday knowledge and can therefore coactivate complex preconcepts in students (Clement, 2013).

Students prefer learning from refutation text while hardly ever resorting to expository text from schoolbooks (Guzzetti et al., 1997; see also Sinatra & Broughton, 2011). The effects from learning with refutation text are typically of medium size and tend to be most pronounced for learning in science education (Schroeder & Kucera, 2022). Using refutation instead of expository text in classes shows improved learning of the targeted concepts (Diakidoy et al., 2003). Changing text flow towards a narrative style does not improve learning from refutation text (Guzzetti et al., 1993), nor does it increase students' acceptance of the texts (Guzzetti et al., 1997).

Again, the most essential design implications from this line of research are to be banked in Table 1: (1) raise awareness of preconcepts, (2) juxtapose and integrate preconcepts with scientifically more adequate concepts, (3) employ contexts that are engaging and well-aligned with the curriculum, (4) introduce analogies as preconcepts, (5) encourage students to elaborate on their concept-building.

Worked examples

One antipode to concept-learning in science education is algorithmic problem-solving. Algorithmic problem-solving can be either considered to belong to lower-order cognitive skills (LOCS, *i.e.*, recalling, comprehending, applying in Bloom's (1956) terms)—as opposed to higher-order cognitive skills (HOCS, *i.e.*, analyzing, synthesizing, evaluating)—, or to be even on a level below LOCS (Zoller, 2002). Student performance in algorithmic problem-solving is lower than performance on LOCS tasks, on which students, in turn, perform lower than in HOCS tasks, which are sometimes referred to as conceptual tasks (*e.g.*, Bius, 2016; Surif et al., 2014; Zoller, 2002). Therefore, neglecting concept-learning and overemphasizing algorithmic problem-solving should be discouraged (“algorithmic calisthenics”), as this does not support students’ conceptual understanding (Nakleh & Mitchell, 1993). Duschl (2008) likewise points out that conceptual and epistemic (*i.e.*, problem-solving) learning need to form a symbiosis. Therefore, innovative formats of learning conceptual knowledge need to be explored. It might be worthwhile to explore some of the techniques employed in problem-solving learning and to adapt these to conceptual learning, such as learning from worked examples.

Problem-solving is an approach to teaching that aims at equipping students with flexible strategies and methods to apply to a diversity of analogous problems (*e.g.*, Chen et al., 2023). Its basic tenet is that a powerful problem-solution strategy can be abstracted from the situation in which it was acquired and can be generalized to apply in new, similar problem situations. Research on transfer from learning with worked examples is unfavorable regarding far-transfer effects (Moreno, 2006).

Concrete examples support learners when interpreting general explanations, while at the same time they depend on general explanations to classify examples (Reimann, 1997). During initial skill acquisition, learners improve their understanding better by studying worked examples—this is known as the worked example principle (Renkl, 2005, 2014b, 2021; Sweller & Cooper, 1985). Worked examples consist of a problem statement, a problem solution which is provided in a step-by-step manner, and the solution to the problem (Hilbert et al., 2008; Renkl, 2021). Hence, worked examples present a complete problem-solving cycle, in which students are frequently encouraged to anticipate solution steps (Stark, 1999) and which provide explanations for solution steps (for an overview see Renkl, 2014b, 2021).

Traditionally, worked examples are used in algorithm-based areas of mathematics (*e.g.*, Gerjets et al., 2006; Leppink et al., 2014), physics (*e.g.*, Kalyuga et al., 2001), or statistics (*e.g.*, Paas, 1992). After introducing a theorem, law, or general principle, worked examples demonstrate how to use this knowledge before learners start working on more open problem-solving tasks (Reiss & Renkl, 2002; Renkl, 2005, 2014b, 2021; Renkl et al., 2009).

Novice learners especially prefer worked examples as they enable them to quickly solve analogous problems on their own (Atkinson et al., 2000). When learners are provided with an instruction *and* an example, they mainly rely on the examples even if the instruction is declared to be important (LeFevre & Dixon, 1986). Learners who study worked examples make fewer mistakes and need less time because they do not process irrelevant operations during performance compared to learners who work on problem-solving tasks during skill acquisition (Sweller & Cooper, 1985). A longitudinal study showed that students who studied worked examples instead of attending a regular math class needed less time to study and performed comparable to or better than the control group (Xinming Zhu & Simon, 1987). Even when worked examples are provided for studying at home,

students from a worked example group outperformed students from a problem-solving group on subsequent performance tests (Ward & Sweller, 1990).

However, some major limitations for the worked example principle are: (1) that worked examples need to be carefully designed according to design guidelines to be effective (*e.g.*, Renkl, 2014a, 2014b, 2021), (2) that worked examples are primarily beneficial to novice learners as skilled learners might get distracted by a scaffold that they do not need (“expertise reversal principle” see Kalyuga, 2014, 2021), (3) that learners need to actively process worked examples (Renkl, 2014b, 2021; Renkl et al., 1998; “self-explanation principle” see Chi, 2021), and (4) that worked examples should not remain an isolated task format; their strength lies in introducing students to a problem-solving approach but not in advising transfer to novel, more dissimilar problems (*e.g.*, Renkl & Atkinson, 2012; Jalani & Sern, 2015). For this, supplementary measures such as self-explanation are suggested (*e.g.*, Fiorella, 2023) or a combination with problem-solving tasks (*e.g.*, Renkl, 2021; van Gog et al., 2011). Even though the worked example principle originally was investigated in algorithmic domains it has also been shown to be valid in non-algorithmic domains, introducing heuristic task solution (Hilbert et al., 2008; Renkl, 2005). Worked example research is much intertwined with research in Cognitive Load Theory, as worked examples aim at optimizing cognitive load during learning (*e.g.*, Renkl & Atkinson, 2012; Sweller & Cooper, 1985; Ward & Sweller, 1990). Therefore, worked example design has been strongly influenced by research on the Cognitive Theory of Multimedia Learning (*e.g.*, Mayer & Fiorella, 2021) that suggests general design principles and guidelines (Mayer, 2020) for optimizing cognitive processing when learning from materials that include more than one source of information, *e.g.*, text and images.

Regarding students learning heuristic and algorithmic content, research suggests these Lessons Learned to be recorded in Table 1: (1) provide a stepwise development of the learning aim, (2) respect Design Principles and Guidelines of the Cognitive Theory of Multimedia Learning, (3) encourage students to self-explain the presented information.

Lessons learned for concept-learning

Students are expected to acquire new conceptual knowledge through learning in schools. They can be supported in this endeavor by instructional texts. Traditional expository texts are frequently too difficult to support student understanding. Conceptual change texts, such as refutation texts, are advocated for improving students’ alignment of preconcepts with scientifically more adequate ones. These texts raise students’ awareness of preconcepts before introducing and integrating updated concepts.

Furthermore, refutation texts benefit from the introduction of analogies, graphical representations, and self-explanation tasks for students. The very same design strategies are effective in learning with worked examples—that traditionally deal with algorithmic or heuristic learning—to support problem-solving abilities. Worked examples provide smaller steps at the beginning and make strategies explicit until students have gained sufficient expertise to take bigger steps on their own. They support students by requiring them to elaborate on their freshly acquired knowledge, thereby allowing contrasting new knowledge with old ideas.

The above is a comprehensive summary (Lessons Learned) of the more detailed discussions of the theoretical background. Each of the reviewed branches of science education research necessarily brings its own focus to investigating the effectiveness of learning. Assuming that learning is complex and multifactorial in its conditions, novel formats of study should consider as much research evidence as possible. Therefore, essential aspects of the above research are comprehensively considered when designing an innovative format for student self-study. Regarding concept-learning, this review suggests to (see Tab. 1): (1) develop scientifically adequate concepts stepwise, (2) start learning from students’

preconcepts, (3) address preconcepts explicitly where there are misconceptions, (4) use narrative language for motivation and reading ease, (5) require students to concurrently monitor their concept-building by reflection and self-explanation prompts, (6) introduce technical terminology functionally, *i.e.*, to build or to close a concept. Merging evidence from research on learning effectiveness, thus, is expected to inform the design of incremental, student-paced learning opportunities for concept-building. The resulting concept-building worked examples are an innovative format for self-study that decidedly emphasizes an undernourished pillar of scientific literacy: having a knowledge of scientific concepts.

Table 1

Combining insights from different strands of science education research to inform the design of an innovative format for science concept-learning

Challenges of concept-learning	Lessons Learned from research on learning with			Consequences for supporting concept-learning
	expository texts in textbooks	refutation texts	worked examples	
Aligning preconcepts with scientifically more adequate concepts	Support reading comprehension by • Reducing syntactic complexity • Using fewer technical terms • Using technical terms to develop a concept • Employing narrative text for motivational reasons	Support concept-building by • Raising awareness of preconcepts • Juxtaposing and integrating preconcepts with scientifically more adequate concepts • Employing contexts that are engaging and well-aligned with the curriculum • Introducing analogies as ‘preconcepts’ • Encouraging students to elaborate on their concept-building	Support student learning by • Providing a stepwise development of the learning aim • Respecting Design Principles and Guidelines of the Multi-media Theory of Learning • Encouraging students to self-explain presented information	Design learning tasks that • Develop scientifically adequate concepts stepwise • Start in students’ preconcepts • Address preconcepts explicitly where there are misconceptions • Use narrative language for motivation and reading ease • Require students to concurrently monitor their concept-building by reflection and self-explanation prompts • Introduce technical terminology functionally, <i>i.e.</i>, to build or to close a concept

Concept-building worked examples – A development project

The proposed concept-building worked examples stem from a development project funded by the Müller-Reitz Foundation. Its aim was to develop learning materials on basic concepts from secondary chemistry education in Germany (Schübler et al., 2016). The complete learning materials are written in German and were explicitly meant for phases of self-study (*e.g.*, to catch up after an absence due to illness or for remedial purposes). We did not design the materials as primarily in-class teaching materials or for homework. In a word: We designed the concept-building worked examples as a substitute when no teacher is available, and the textbook does not help.

We planned for a publication and distribution free of charge so that use of self-study materials would not be hindered by economic status. This—and later the sheer extent of developed materials (Tab. 2)—prohibited a print-version, and we decided to arrange the concept-building worked examples in a ubiquitous file-format distributable easily via the internet. We designed the concept-building worked examples so that they could viably be displayed as pdf-files. We actualized some interactive elements (*e.g.*, requiring students to provide written answers) by means of form fields (Fig. 1).

Figure 1

Example slide “form fields” (blue tinted boxes: translation and comment for screenshot from Scientific Models-Diffusion (10/63))

„Und wie macht es das?“, fragt Lisa. „Was haben denn die Teilchen mit dem Geruch von den Kerzen in meinem Zimmer zu tun?“, fragt auch Max. „Sicher weißt du, aus welchem Stoff Kerzen bestehen?“, fragt Sinan weiter. „Aus Wachs“, antwortet Max. „Richtig“, sagt Sinan und nickt. „Und mit dem Wachs ist es genau wie mit der Kohle. Wenn du dir vorstellst, dass du das Wachs in immer kleinere Wachsportionen zerteilen könntest ...“ Beschreibe, was passieren würde, wenn du Wachs immer weiter zerteilen könntest, bevor du weiter liest. form field for filling in solutions online and for saving to pdf	"But how can it do that", Lisa asks. "What do those particles have to do with the scent of candles in my room?", Max follows up. "Certainly, you know which substance candles are made of?", Sinan inquires. "Of wax", Max replies. "Correct", Sinan nods affirmingly. "With wax it is just as with the coal from earlier on. Picture cutting wax into ever smaller wax portions ..." Describe what would happen if you cut wax ever smaller, before continuing to read.
--	--

Aiming at access from different portable end devices we organized worked examples in small screens so that scrolling could be kept to a minimum. Therefore, composition of the concept-building worked examples made use of presentation software slides, that also required authors to keep it short and simple. The design of the concept-building worked examples drew on research literature referenced in the theoretical background. The following section will illustrate how worked example research informed the design of the concept-building worked examples.

Cumulative concept-building through sets of worked examples

Instead of presenting an algorithmic or a heuristic solution to a problem, the proposed worked examples serve to develop conceptual knowledge in chemistry by introducing conceptual knowledge necessary to explain a scientific phenomenon. The innovative format of concept-building worked examples strives to combine beneficial features of worked examples with what is known about learning from texts. Concept-building, thus, can be well supported, ultimately relying on the above principles to introduce a concept. These principles are realized in all the concept-building worked examples without forcing them into a rigid pattern. Rather, the narrative flow and context determine how and when they are introduced:

- providing defining features of a concept with examples (*e.g.*, acid molecules contain a polarized covalent bond to at least one hydrogen atom that can dissociate from the molecule; acids turn universal indicator red)
- providing non-examples that do not share defining features (*e.g.*, hydrogen atoms not connected by a polarized bond cause no acid reaction)
- applying the concept to other examples (*e.g.*, translating the dissociation of protons from phosphoric to citric acid)
- differentiating between examples and non-examples in a dialogue between novices and experts (*e.g.*, “Wow, that’s a lot of H-atoms in citric acid.” – “And even more other atoms.” – “Yes, but only a few of the H-atoms are connected through a polarized bond.”)
- enriching the concept with related concepts (*e.g.*, introducing pH as a measure of proton concentration; modelling neutralization as a process to capture dissociated protons)
- extending the concept (*e.g.*, introducing bases as substances that bond dissociated protons; bases turn universal indicator greenish-blue)

Concept-building worked examples were divided into sets of smaller, self-contained examples (see Tab. 2). This structure accounts for the inherent complexity of scientific concepts that frequently combine several concepts into an overarching one. Above this, the division into smaller, self-contained worked examples allows students to plan breaks when studying (Renkl, 2014a, 2014b, 2021). Presenting them with a complete set of slides for an overarching concept could probably be detrimental motivation-wise. Starting with the second example in a set, students are reminded of central ideas from the previous example(s). When studying the examples students are asked to answer questions, to give their ideas on a topic, or to apply their freshly acquired concept knowledge. They enter their replies into form fields of the pdf files enabling saving solutions—this is meant to encourage self-explanation (Fig. 2). Individual examples close with a self-assessment in which students are encouraged to write down their understanding of central ideas from the concept; again, form fields are provided. Example solutions allow for self-assessment.

Figure 2

Example slide “recapitulation” (Water – hydrogen bonds (6/45))



WASSER → SEQUENZ 1

TEIL III: Wasserstoffbrücken

Hugi-Bugi casts a short, triumphant glance over at Jupiter-Saturn and rambles on: "Such a bond, in which the binding electrons are somewhat closer to the atom with the higher electronegativity, is called a **polarized atom bond**. The shift of binding electrons causes both the atoms to carry **partial charges**. Hydrogen's electrons are a little bit further away than usual from hydrogen's nucleus and so the hydrogen atom is charged partially positive. Thus, the oxygen atom doesn't just have its own electrons around. It even attracts the hydrogen's electrons so that it becomes charged partially negative. We note this with **delta minus (δ^-)**, **delta plus (δ^+)** respectively."

Hugi-Bugi wirft noch einen kurzen, triumphierenden Blick zu Jupiter-Saturn und legt dann los: „Eine solche Bindung, bei der die Bindungselektronen etwas näher am elektronegativeren Atom sind, wird **polare Atombindung** genannt. Die Verschiebung der Bindungselektronen führt dazu, dass die beiden Atome **Partiellladungen** tragen. Die Elektronen von den Wasserstoffatomen sind vom Wasserstoffatomkern ein bisschen weiter weg als normalerweise, deswegen sind die Wasserstoffatome partial positiv geladen. Das Sauerstoffatom hat nicht nur seine eigenen Elektronen um sich herum, sondern es hat zusätzlich noch die Außenelektronen von den Wasserstoffatomen nah zu sich gezogen, daher ist das Sauerstoffatom partial negativ geladen. Dafür schreibt man dann so ein **Delta-Minus (δ^-)** oder **Delta-Plus (δ^+)**.“

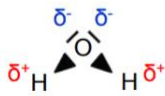


Bild 2: Polare Atombindungen und Partiellladungen im Wassermolekül

[image 2: polarized atom bond and partial charges in a water molecule]

The individual worked examples are divided into meaningful sections, each of which is presented on a single slide (“segmenting principle” see Mayer, 2020; “meaningful-building-blocks guideline” see Catrambone, 1998; Gerjets et al., 2006; Renkl, 2014a, 2014b, 2021). Individual slides typically contain a complete solution step/idea for building the concept—rarely is an idea spread over more than one slide, so that a slide could principally stand on its own, which necessarily leads to some slides that are lean in textual information. This structural scaffold invites learners to take the time they need to adequately process information of the just-read paragraph before reading the next paragraph (pausing and temporal cueing: Mayer & Pilegard, 2014; Spanjers et al., 2012).

Table 2

Overview of covered concepts and topics in concept-building worked-out examples (Schüßler et al., 2016), plus length (slide count)

Concept (set of examples)	Topic (individual example)	Slides (#)
naming of alkanes	straight-chained alkanes	39
	simple-branched alkanes	48
	multiple-branched alkanes (mono-substituting side chains)	46
	multiple-branched alkanes (poly-substituting side chains)	40
atomic models	Rutherford's model	50
	protons and neutrons	42
	isotopes	42
	Bohr's model	52
water	geometry of the water molecule	71
	surface tension	41
	hydrogen-bonds	50
science models	diffusion	67
	changes of state of matter	79
	elements, atoms, molecules	56
	example of a chemical reaction: incomplete combustion	49
	complete combustion	66
redox reactions	oxidation reactions with oxygen	36
	redox reaction of iron oxide	34
	redox reaction of silver sulfide	54
acids (coke)	acid solutions	44
	pH	38
	neutralization	43
acids (red tea)	acid solutions	40
	pH	36
	neutralization	45
acids (wine gum)	acid solutions	62
	pH	41
	neutralization	37
chemical substances	properties	44
	mixtures	41
	separation methods	64
	extraction	22
	introduction to the chemical reaction	36
9 concepts	34 worked examples	1554

Note: “acids” was developed in three different contexts for the sake of an evaluation study of these self-study-materials (Schüßler, 2017). The concepts introduced in the worked example sets are shared between the contexts.

Relating to the students’ perspective in worked examples

A central issue to be addressed in the innovative task for concept-learning was to emancipate students from their textbooks’ motivational pitfalls. Students want to read texts—if they must—that exhibit “an attractive style of writing that motivates students toward further reading and deeper thought,” which can be achieved by making texts comprehensible, interesting, and rich in examples (Knecht & Najvarová, 2010, p. 11). Text comprehension—as estimated by reading indices—can be improved by employing a narrative style of language that is less heavy on passive constructions, impersonal subjects, or hypotaxis (Safari et al., 2017). Bearing in mind that the proposed tasks are meant for students who cannot easily learn during classes or from textbooks, it is assumed that the motivational windfall gained from narrative embedding may lead to the generation of situational interest, which in turn can exert a beneficial effect on learning (Hulleman et al., 2010; Xihe Zhu et al., 2009). Therefore, introducing cover stories with fictional characters for worked examples might allow exploitation of these motivational advantages (see also Avraamidou & Osborne, 2009).

A variation of contexts has been shown to exert an influence on students’ situational interest, favoring everyday-related and uncommon contexts (Habig et al., 2018). Therefore, a variety of contexts was covered through which students encounter the chemical concepts: *e.g.*, teenage boys wondering at the pool why there is phosphoric acid in coke, pre-schoolers learning about mixtures and reactions when baking muffins with their babysitter, a pair of siblings being introduced to oxidation and reduction by their grandfather who is sorting through rusted screws, or a bunch of extra-terrestrial schoolkids learning about earth and that special (*i.e.*, anomalous) substance water. These cover stories allow exploiting the potential of what van Gog and Rummel (2010) term “coping models” and “mastery models” by introducing fictional characters with different levels of expertise. Coping characters are typically novices who try to succeed in something (cope); they can serve as figures of identification or role models; they are the stand-ins to commit typical mistakes and to ask dumb or naïve questions (Fig. 3). Mastery characters on the other hand are more progressed in the matter in question and are understood to be experts in the story; they present relevant knowledge correctly, answer questions competently and correct typical mistakes (van Gog & Rummel, 2010). Thus, worked examples can use cover stories with mastery and coping characters to develop a concept step-by-step by first discovering a problem, then discussing it with each other, and finally arriving at a joint solution. Koenen (2014) used such cover stories in worked examples to foster inquiry learning. Likewise, Kölbach’s (2011) worked examples resorted to cover stories to demonstrate how conceptual knowledge from chemistry could be applied to explain phenomena.

Figure 3

Example slide “naive question” (Science models - Elements, atoms, and molecules (13/ 51))



"If we heated wax in a cup till it boiled – in a pot on the stove, for instance –, there would indeed curl up gaseous wax above the cup. However, when we light a candle, no gaseous wax curls up above it", Sinan tells them.

[image 4: burning candle]

"But where's the difference in that", Max interrupts.

"The difference is, that wax which is heated in a pot *boils*, while wax in a burning candle *combusts*", Sinan says.

Coping characters serve as figures of identification as they have about the same age as the potential learners; they can be modelled on the intended readership regarding sex and ethnical background, thus extending the potential for self-identification and for providing positive role models (“model-observer similarity principle” see Renkl, 2014a). They do not have—from the start—the necessary conceptual knowledge to explain the phenomenon they have discovered. By making assumptions, asking questions, and discussing their ideas with each other, they provide positive role models for an active involvement with the phenomenon and the relevant conceptual knowledge (Fig. 4). They serve as projections of the learners’ assumed preconcepts without necessitating the learners themselves to disclose their insufficient knowledge.

Figure 4

Example slide “dialogue” including preconception (Redox reactions – Oxidation reactions with oxygen (28/ 32))



"But, oxygen is a gas!", Lennart protests. "It doesn't weigh anything." "Each atom has its specific mass", his grandfather corrects, "Even atoms of gaseous substances. Therefore, the screw's mass increases when more iron oxide forms." "And you can really weigh this?", Lia inquires fascinatedly while picturing herself putting the screws onto grandma's ancient scales. As if having guessed at her thoughts, grandfather replies: "Grandma's ancient scales certainly could not do this – the difference is far too minute. But very fine scales could work." "However", he continues, "we would need to prevent parts of the iron oxide from crumbling off when touching the screws." Lia is somewhat disappointed and murmurs, "Pity."

Making preconceptions explicit can prevent learners from making the same mistakes (D’Mello et al., 2014; Große & Renkl, 2004, 2007; “studying errors guideline” see Renkl, 2014a., 2014b, 2021; Roy & Chi, 2005) or correct already existing preconceptions. Coping characters can, therefore, support conceptual change, as their introduction allows students to co-activate relevant preconceptions and integrate them with the intended scientifically adequate conception (Kendeou et al., 2014). Thus, they turn into identification figures as well as learning facilitators that can inquire “about everything you always wanted to know but were afraid to ask”.

Hartman et al. (2020) showed that observing someone engaging in problem-solving can support students as they become aware of their own knowledge gaps. At the same time, coping characters can inspire learners to behave in a similar way and attempt to formulate elaborations, which might to some extent compensate for poor self-explanations (Stark, 1999). Even self-explanations that attend to incorrect information contribute to concept-building: Siegler (2009) showed that students who explained both the correct and erroneous problem solutions of two children outperformed students who explained only the correct solution of another person.

Accordingly, the introduction of a mastery character has the advantage that instructional explanations can be integrated into the dialogue of a worked example (Nokes et al., 2011; Renkl, 2002). Mastery characters provide scientifically adequate explanations for an encountered phenomenon. That means that instead of presenting a mastery performance—as would be the case in traditional heuristic and algorithmic worked examples—the proposed character presents conceptual mastery information in a Socratic manner. This is considered important for learners with low prior knowledge (Renkl, 2014b, 2021) and for learners who do not self-explain information or solution steps to themselves (“explanation-help guideline” see Renkl, 2014a, 2014b, 2021; Roy & Chi, 2005). Roelle et al. (2017) showed that students need to adequately organize instructional explanations into coherent mental representations in order to profit from studying worked examples. Hence, they conclude that learners should be supported not only by self-explanation prompts but also in organizing instructional explanations.

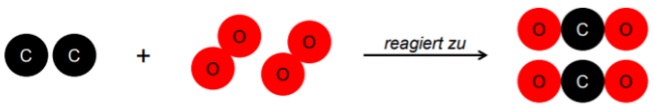
Design principles implemented at the example level

So far, some overarching design criteria have been introduced that pertained to the content of written text rather than to its graphical presentation. This section will change grain size and zoom in on the layout of worked examples to illustrate how some well-researched design principles from the Cognitive Theory of Multimedia Learning (Mayer, 2020) have informed the design of concept-building worked examples. The presentation necessarily has an exemplary character, and not each principle might feature equally prominently in all the concept-building worked examples.

The worked examples exploit the multimedia principle (Mayer, 2020, which requires learning materials to complement textual with pictorial information by presenting text and figures in proximity (“spatial contiguity principle” see Mayer, 2020), *i.e.*, on the same slide. Where textual and pictorial elements correspond, they are given in the same color (color coding); novel information is printed in bold face, while parts of complex figures, which are not explicitly addressed in the respective textual information, may be faded out (greying out; all techniques responding to “signaling principle” see Mayer, 2020, Fig. 5).

Figure 5

Example slide “spatial contiguity, color coding, signaling, self-explanation” (Redox-reactions oxidation reactions with oxygen (17/32))

„Es gibt nicht nur unser Eisenoxid, sondern noch viele andere Oxide“, fährt Lias und Lennarts Opa fort. „Zum Beispiel Kohlenstoffdi-oxid?“, ruft Lia aufgeregt. Den Begriff hat sie letzte Woche im Biologieunterricht gelernt, als es um die Atmung ging. Die Luft, die Menschen ausatmen, enthält mehr Kohlenstoffdioxid als die Luft, die sie zuvor eingeatmet haben. Ihr Opa nickt wieder zustimmend. „Richtig. Kohlenstoffdioxidmoleküle entstehen, wenn Kohlenstoffatome mit Sauerstoffatomen reagieren.“  Bild 3: Oxidation von Kohlenstoff „Und das ist dann auch so eine Oxidation?“, fragt Lennart noch mal nach. Sein Opa nickt erneut. Fasse zusammen, was eine Oxidation kennzeichnet und was ein Oxid ist, bevor du weiter liest.	"There isn't only our iron oxide but many more oxides", Lia and Lennart's grandfather adds. "Carbon di-oxide, for instance", exclaims Lia excitedly. She has learned the term in last week's biology lesson when they were talking about respiration. The air exhaled by people contains more carbon dioxide than the air they inhaled before. Her grandfather nods his head in confirmation, "Correct, molecules of carbon dioxide form when carbon atoms react with oxygen atoms." [image 3: oxidation of carbon] "And this also is such a reaction?", Lennart wants to make sure. His grandfather nods again. Please, summarise, what defines an oxidation and what an oxide is before continue reading.
--	--

Introducing multiple characters into a cover story allows unfolding of the worked example in a dialogue, which helps to better speak to learners (“personalization principle” see Mayer, 2020). Even if the mastery character puts their question to the coping character, the “you” addresses the student reader, too. It helps that questions in the dialogue evolve organically and are, therefore, understood to be necessary as opposed to typical textbook tasks that appear to be disconnected from prior discourse.

Active processing of the worked example by the learner is necessary for successful learning. Self-explanations generate information that goes beyond the text and help to connect new information with each other and with existing concepts in a meaningful way (“self-explanation principle” see Chi, 2021). Since only few learners spontaneously show sufficient self-explanatory behavior (Chi et al., 1989; Renkl et al., 1998), self-explanations should be supported with prompting measures (Stark et al., 2008). These prompts can occur in the form of encouragement or incentives to think something through (Chi, 2021; Roy & Chi, 2005; Wylie & Chi, 2014; see Fig. 5). These incentives do not necessarily require students to give a written answer but should activate students cognitively. Students might be prompted to answer a question, summarize for themselves what they have learned so far, identify crucial information in one of the characters’ utterances, or enter into a thought experiment. They must start an inner dialogue about what they have learned. In this way, new information is derived from the explicitly given information. Comparisons with the mental model reveal differences and initiate concept changes (Chi, 2000). If not taken up actively, received information cannot be used for concept-building; this would be akin to waiting for the flood to heap up a sandcastle.

Design guidelines implemented at the example-set level

Worked examples should not be administered as one-shot interventions but function most effectively when forming a set of examples. Segmenting an overarching concept thus allows for the composition of shorter and, therefore, better to digest entities of study. At the same time, it enables introduction of repetition and recall when spreading study for concept-building over some time (also “self-explanation”).

Providing more than one worked example on a topic proves beneficial for learning from worked examples. Only by encountering several worked examples can learners distinguish shared structural features—*i.e.*, critical features of a concept—from irrelevant surface features (“example-set guideline” see Renkl, 2014a, 2014b, 2021; also Atkinson et al., 2003). In addition, when presenting a series of worked examples, omission of individual solution steps that have been introduced in a prior example encourages learners to fill these steps themselves. It is suggested to start at the worked examples’ end with the omissions so that the familiarized structure of the worked example can serve as a scaffold leading up to the omission (“fading guideline” see Renkl, 2014a, 2014b, 2021; also Atkinson et al., 2003). When designing the worked examples, we tried to take all these design principles into account as much as possible to develop learning materials that help and motivate learners to acquire chemical conceptual knowledge independently.

Discussion

This article introduces an innovative format of learning text: the concept-building worked example. It has been developed from merging research on two crucial aspects of learning: concept-learning and learning from texts. Its basic premises can be outlined thus: Scientific concepts are complex constructs that take time to be constructed; understanding can result neither from hearing them spoken once or twice in classrooms, nor from condensed textbook representations. “Understanding” here refers to a complex process of “selecting key information from the provided learning material, organizing it into a coherent structure in working memory, and integrating it with existing knowledge from long-term memory” (Fiorella, 2023, p. 53). Just like in the case of complex algorithms and heuristics, science concepts can be developed stepwise keeping pace with students’ constructive learning by means of worked examples.

The worked examples in this paper have been evaluated empirically (Schübler, 2017)—the scope of the article prohibits a comprehensive introduction to the study, so that a brief outline of results may suffice: Student learning ($N=436$, $M_{age}=14.1$, 49.8% female) benefits from working with these worked examples with large effect sizes ($d = 0.88$). The learning gains can still be detected in a follow-up test, albeit to a lesser extent ($\eta^2_p = .333$). Students show decreasing mental effort and self-reported perceived task difficulty over the course of working through a set of worked examples, which is suggestive of successful cumulative construction of concepts.

Given that the concept-building worked examples’ basic rationale—allow time to construct a concept—should pertain to all learners, their implementation could benefit minoritized or otherwise disadvantaged students, too. Jones & Burrell (2022) claim that there is a subtle shift in the quality of instruction when looking at teaching minoritized students. Some of this might be traced back to socialized prejudices, some to unequal access to high-quality learning opportunities. Self-study materials might be a remedy in this situation as they are insensitive toward their users’ personalities (at least regarding feedback). Rolling out worked examples online, free of charge, provides unhindered access to science education materials to a substantial number of students. Organizing science education content in smaller, digestible entities—*i.e.*, sets of self-contained worked examples— is a

form of “Microlearning” that supports learners to focus and to learn at their own pace (Mostrady et al., 2025).

Moreover, offering the materials online, allows today—more than at the time of their original development—an easy combination with other tools. Students who are visually challenged can have the example read out to them by assistive technology, moving toward equitable science education (Jones & Burrell, 2022). Automated translations into a multitude of languages using AI-translation engines are virtually only one click away. Even if these translations turned out to be less than perfect, they should still form a sound basis from which to do the fine-tuning of a text. This could ease production of learning materials substantially as opposed to starting from scratch—especially for practicing teachers who are chronically short of time. Having based the materials on recent international, *i.e.*, not culture-specific research on learning and science education suggests that the underlying mechanisms are transferable. There is no reason to assume that the concept-building worked example should work less effectively in a sub-Saharan or a Western suburban locale. Still, this type of material might speak to some students more than to others; but this is the case for each offer science education makes. This should not prevent dissemination but encourage further development with an innovative focus serving the needs of different clienteles.

Having stressed the advantages and potentials of concept-building worked examples, one must not turn a blind eye to their limitations: Does their exclusive online presence not keep them from those students who lack the technological means to access them (Mostrady et al., 2025)? Are the covered concepts still valid considering shifting philosophies of scientific literacy (*e.g.*, Osborne, 2023)? Given students’ growing overreliance on generative AI (*e.g.*, Bauer et al., 2025), would they be as accepting of such a format as students of Schüßler’s (2017)? Being drafted from a white Western perspective, how equitable can they really be (Jones & Burrell, 2022)—even if they try to contradict stereotypes by making girls and migrants the heroines and heroes; or need this today be considered cultural appropriation?

Some of these critical issues do not touch on the *fundamental* design principles but on *superficial* authorial decisions that could easily be reversed should one wish to adopt the concept. Some others might need reflection and reassessment of the goals of science education. Others again require us to self-critically reflect which benefits lie in technological innovation and which risks they bear. For example: Which is the most reasonable way to incorporate AI into learning scenarios, and how can we ascertain that students remain critical of AI outputs (*e.g.*, Bauer et al., 2025)?

Conclusion—Looking back and looking forward

This manuscript, not being a traditional research paper or classical review, needs another kind of ending that still provides summary but also opens perspectives.

Science education has gone a long way to learn how it can support students to learn how to do science. Considering that *doing* something needs *knowing* something in the first place, we still have too little knowledge of how we can help students build the complex concepts of science from which their “doing science” can commence. We, as a discipline, are rich in teaching methods, inquiry practices, and explanation videos. Yet, we are still underdeveloped textbook-wise—approaches to embed science into concepts (Bennett & Lubben, 2006; Parchmann et al., 2023) or to emphasize inquiry practices (Bennett et al., 2007) have made significant contributions, but they have not specifically addressed the issue of learning a concept. Worse still, textbook publishers appear to have succumbed to the challenge of producing textbooks that are as thin as possible since young people allegedly do not read. This is foolish for two reasons: If we consider reading to be essential for leading a reasonable and informed adult life, we must not lower the demands on students’ reading abilities but raise them (Leisen quoted in Meyer & Pietzner, 2022). Secondly, the same young people who allegedly do not read continue to devour the seven *Harry Potter* novels, which come to over 3’000 pages. Young

people do read, but they do not read everything—and least of all, will they read something that does not take them, nor their interests, nor their needs seriously. And there goes the textbook.

By composing texts that young people might deem interesting and entertaining to read, concept-building worked examples invite students to *engage* with instructional text—if a text cannot achieve this, learning will probably not be initiated. The project chose to compose learning texts of considerable length despite a general trend to create text-lean environments—again, if young people read *A Song of Ice and Fire* or the *Lord of the Rings* trilogy, we should not assume that text length is the most crucial criterion for their disinclination towards science texts. If the materials should find adoption (or adaptation) in other locales it would acknowledge the work effort put into developing the innovative format. If, above this, the format should be convincing for other researchers to adopt to different subjects and concepts, the acknowledgment would be taken to a wholly different level. In any case, we are convinced that science education must not forget to find ways of supporting students to get acquainted with scientific fact—googling something or reproducing AI outputs cannot substitute for *bona fide* scientific literacy. “Doing science” can only grow from “knowing science”—neglect either of these and scientific literacy is no longer attainable.

Funding

This work was supported by the Müller Reitz Foundation Grant T0009/23094/2012/KG.

Markus Emden (markus.emden@phzh.ch) has been a professor of science education at Zurich’s University of Teacher Education since 2018. In 2011, he completed his PhD at Duisburg-Essen University (Germany) on process-oriented performance assessment of practical work. His research focus is on science teachers’ continuous professional development, practical work in science classes, and on developing learning materials for chemistry education.

Elke Sumfleth (elke.sumfleth@uni-due.de) was a professor of chemistry education at the University of Duisburg-Essen since 1990 and is now retired. Her research focus was on empirical research in science education, especially in chemistry education.

Katrin Schuessler (katrin.schuessler@uni-due.de) is a post doc researcher at the University of Duisburg-Essen. Her research focus is on empirical research in chemistry education. She is particularly interested in innovative learning materials.

References

- Ariasi, N. (2012). *From eye fixations to scientific contextual learning: A process approach to the refutation text effect* [PhD-Thesis]. Università degli Studi di Padova, Padua.
- Asterhan, C. S. & Resnick, M. S. (2020). Refutation texts and argumentation for conceptual change. A winning or a redundant combination? *Learning and Instruction*, 65, 101265. <https://doi.org/10.1016/j.learninstruc.2019.101265>
- Atasoy, B., Akkus, H., & Kadayifci, H. (2009). The effect of a conceptual change approach on understanding of students’ chemical equilibrium concepts. *Research in Science & Technological Education*, 27(3), 267–282. <https://doi.org/10.1080/02635140903162587>
- Atkinson, R. K., Renkl, A., & Merrill, M. M. (2003). Transitioning from studying examples to solving problems: Effects of self-explanation prompts and fading worked-out steps. *Journal of Educational Psychology*, 95(4), 774–783. <https://doi.org/10.1037/0022-0663.95.4.774>

- Atkinson, R. K., Sharon, J. D., Renkl, A. & Wortham, D. (2000). Learning from examples: Instructional principles from the worked examples research. *Review of Educational Research*, 70(2), 181–214.
- Ausubel, D. P. (1969). A cognitive theory of school learning. *Psychology in the Schools*, 6(4), 331–335 (Paper presented at the meeting of the American Psychological Association, San Francisco, September, 1968).
- Avraamidou, L. & Osborne, J. (2009). The role of narrative in communicating science. *International Journal of Science Education*, 31(12), 1683–1707. <https://doi.org/10.1080/09500690802380695>
- Bauer, E., Greiff, S., Graesser, A. C., Scheiter, K., & Sailer, M. (2025). Looking beyond the hype: Understanding the effects of AI on learning. *Educational Psychology Review*, 37(2), Article 45. <https://doi.org/10.1007/s10648-025-10020-8>
- Beerenwinkel, A., & Gräsel, C. (2005). Texte im Chemieunterricht: Ergebnisse einer Befragung von Lehrkräften. *Zeitschrift für Didaktik der Naturwissenschaften*, 11, pp. 21–39.
- Bennett, J., & Lubben, F. (2006). Context-based chemistry: The Salters approach. *International Journal of Science Education*, 28(9), 999–1015. <https://doi.org/10.1080/09500690600702496>
- Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, 91(3), 347–370. <https://doi.org/10.1002/sce.20186>
- Bergqvist, A., & Chang Rundgren, S.-N. (2017). The influence of textbooks on teachers' knowledge of chemical bonding representations relative to students' difficulties understanding. *Research in Science & Technological Education*, 35(2), 215–237. <https://doi.org/10.1080/02635143.2017.1295934>
- Bius, J. H. (2016). *Increasing the Understanding of Chemical Concepts: The Effectiveness of Multiple Exposures* [PhD-Thesis]. University of Kansas, Lawrence, KA.
- Bloom, B. S. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Longmans, Green, and Co.
- Calhoun, R. S. & Rubba, P. A. (1993). An examination of the conceptual structure and reading level of six sixth-grade science textbooks. *Journal of Elementary Science Education*, 5(2), 21–36.
- Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General*, 127(4), 355–376. <https://doi.org/10.1037/0096-3445.127.4.355>
- Chen, O., Retnowati, E., Chan, B. K. Y., & Kalyuga, S. (2023). The effect of worked examples on learning solution steps and knowledge transfer. *Educational Psychology*, 43(8), 914–928. <https://doi.org/10.1080/01443410.2023.2273762>
- Chi, M. T. H. (2021). The self-explanation principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd edition, pp. 381–393). Cambridge University Press. <https://doi.org/10.1017/9781108894333.040>
- Chi, M. T. H. (2000). Self-explaining: The dual processes of generating inference and repairing mental models. In R. Glaser (Ed.), *Advances in instructional psychology: Educational design and cognitive science*, Vol. 5, pp. 161–238). Lawrence Erlbaum Associates Publishers.
- Chi, M. T., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science*, 13(2), 145–182. https://doi.org/10.1207/s15516709cog1302_1
- Clement, J. J. (2013). Roles for explanatory models and analogies in conceptual change. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (2nd edition, pp. 412–446). Routledge.
- D'Mello, S., Lehman, B., Pekrun, R. & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction*, 29, 153–170. <https://doi.org/10.1016/j.learninstruc.2012.05.003>

- Danielson, R. W., Sinatra, G. M., & Kendeou, P. (2016). Augmenting the refutation text effect with analogies and graphics. *Discourse Processes*, 53(5-6), 392–414. <https://doi.org/10.1080/0163853X.2016.1166334>
- Diakidoy, I.-A. N., Kendeou, P., & Ioannides, C. (2003). Reading about energy: The effects of text structure in science learning and conceptual change. *Contemporary Educational Psychology*, 28(3), 335–356. [https://doi.org/10.1016/S0361-476X\(02\)00039-5](https://doi.org/10.1016/S0361-476X(02)00039-5)
- Diakidoy, I.-A. N., Mouskounti, T., Fella, A., & Ioannides, C. (2016). Comprehension processes and outcomes with refutation and expository texts and their contribution to learning. *Learning and Instruction*, 41, 60–69. <https://doi.org/10.1016/j.learninstruc.2015.10.002>
- Diakidoy, I.-A. N., Mouskounti, T., & Ioannides, C. (2011). Comprehension and learning from refutation and expository texts. *Reading Research Quarterly*, 46(1), 22–38. <https://doi.org/10.1598/RRQ.46.1.2>
- Duschl, R. (2008). Science education in three-part harmony: Balancing conceptual, epistemic, and social learning goals. *Review of Research in Education*, 32(1), 268–291. <https://doi.org/10.3102/0091732X07309371>
- Fang, Z. (2006). The language demands of science reading in middle school. *International Journal of Science Education*, 28(5), 491–520. <https://doi.org/10.1080/09500690500339092>
- Fiorella, L. (2023). Making sense of generative learning. *Educational Psychology Review*, 35(2), Article 50. <https://doi.org/10.1007/s10648-023-09769-7>
- Gagné, R. M. (1965). The learning of concepts. *The School Review*, 73(3), 187–196.
- Gerjets, P., Scheiter, K., & Catrambone, R. (2006). Can learning from molar and modular worked examples be enhanced by providing instructional explanations and prompting self-explanations? *Learning and Instruction*, 16(2), 104–121. <https://doi.org/10.1016/j.learninstruc.2006.02.007>
- Gilbert, J. K., & Watts, D. M. (1983). Concepts, misconceptions and alternative conceptions: Changing perspectives in science education. *Studies in Science Education*, 10(1), 61–98. <https://doi.org/10.1080/03057268308559905>
- Große, C. S. & Renkl, A. (2007). Finding and fixing errors in worked examples: Can this foster learning outcomes? *Learning and Instruction*, 17(6), 612–634. <https://doi.org/10.1016/j.learninstruc.2007.09.008>
- Große, C. S. & Renkl, A. (2004). Learning from worked examples. What happens if errors are included? In P. Gerjets, P. A. Kirschner, J. Elen & R. Joiner (Eds.), *Instructional design for effective and enjoyable computer-supported learning*. (pp. 356–364). Knowledge Media Research Center.
- Groves, F. H. (2016). A longitudinal study of middle and secondary level science textbook vocabulary loads. *School Science and Mathematics*, 116(6), 320–325. <https://doi.org/10.1111/ssm.12183>
- Groves, F. H. (1995). Science vocabulary load of selected secondary science textbooks. *School Science and Mathematics*, 95(5), 230–234.
- Guzzetti, B. J., Snyder, T. E., Glass, G. V., & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, 28(2), 116–159.
- Guzzetti, B. J., Williams, W. O., Skeels, S. A., & Wu, S. M. (1997). Influence of text structure on learning counterintuitive physics concepts. *Journal of Research in Science Teaching*, 34(7), 701–719.
- Habig, S., Blankenburg, J., van Vorst, H., Fechner, S., Parchmann, I., & Sumfleth, E. (2018). Context characteristics and their effects on students' situational interest in chemistry. *International Journal of Science Education*, 40(10), 1154–1175. <https://doi.org/10.1080/09500693.2018.1470349>
- Härtig, H. (2014). The terminology within German lower secondary physics textbook. *Science Education Review Letters*, 1–7. <https://doi.org/10.18452/8210>

- Hartmann, C., van Gog, T., & Rummel, N. (2020). Do examples of failure effectively prepare students for learning from subsequent instruction? *Applied Cognitive Psychology*, 34(4), 879–889. <https://doi.org/10.1002/acp.3651>
- Herron, J. D., Cantu, L. L., Ward, R. & Srinivasan, V. (1977). Problems associated with concept analysis. *Science Education*, 61(2), 185–199.
- Hilbert, T. S., Renkl, A., Kessler, S. & Reiss, K. (2008). Learning to prove in geometry: Learning from heuristic examples and how it can be supported. *Learning and Instruction*, 18(1), 54–65. <https://doi.org/10.1016/j.learninstruc.2006.10.008>
- Hodson, D. (2014). Learning Science, Learning about Science, Doing Science: Different goals demand different learning methods. *International Journal of Science Education*, 36(15), 2534–2553. <https://doi.org/10.1080/09500693.2014.899722>
- Hu, J., Gao, X., & Qiu, X. (2021). Lexical coverage and readability of science textbooks for English-medium instruction secondary schools in Hong Kong. *Sage Open*, 11(1), Article 21582440211001867. <https://doi.org/10.1177/21582440211001867>
- Hulleman, C. S., Godes, O., Hendricks, B. L., & Harackiewicz, J. M. (2010). Enhancing interest and performance with a utility value intervention. *Journal of Educational Psychology*, 102(4), 880–895. <https://doi.org/10.1037/a0019506>
- Hynd, C. R. (2001). Refutational texts and the change process. *International Journal of Educational Research*, 35(7-8), 699–714.
- Jalani, N. H., & Sern, L. C. (2015). The example-problem-based learning model: Applying Cognitive Load Theory. *Procedia - Social and Behavioral Sciences*, 195, 872–880. <https://doi.org/10.1016/j.sbspro.2015.06.366>
- Johnson, C. E., & Boon, H. J. (2023). Identifying and challenging the narrow cognitive demands of science textbooks. In G. P. Thomas & H. J. Boon (Eds.), *Challenges in science education* (pp. 279–303). Springer International Publishing. https://doi.org/10.1007/978-3-031-18092-7_13
- Johnson, K. & Bulla, A. J. (2021). Creating the components for teaching concepts. *Behavior Analysis in Practice*, 14(3), 785–792. <https://doi.org/10.1007/s40617-021-00626-z>
- Jones, T. R., & Burrell, S. (2022). Present in class yet absent in science: The individual and societal impact of inequitable science instruction and challenge to improve science instruction. *Science Education*, 106(5), 1032–1053. <https://doi.org/10.1002/sce.21728>
- Kahveci, A. (2010). Quantitative analysis of science and chemistry textbooks for indicators of reform: A complementary perspective. *International Journal of Science Education*, 32(11), 1495–1519. <https://doi.org/10.1080/09500690903127649>
- Kalyuga, S. (2021). The expertise reversal principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd edition, pp. 171–182). Cambridge University Press. <https://doi.org/10.1017/9781108894333.017>
- Kalyuga, S. (2014). The expertise reversal principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 576–597). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.028>
- Kalyuga, S., Chandler, P., Tuovinen, J. & Sweller, J. (2001). When problem solving is superior to studying worked examples. *Journal of Educational Psychology*, 93(3), 579–588. <https://doi.org/10.1037/0022-0663.93.3.579>
- Kean, E. (1982). Concept learning theory and design of college chemistry instruction. *Journal of Chemical Education*, 59(11), 956–959
- Kendeou, P., Walsh, E. K., Smith, E. R., & O'Brien, E. J. (2014). Knowledge revision processes in refutation texts. *Discourse Processes*, 51(5-6), 374–397. <https://doi.org/10.1080/0163853X.2014.913961>
- Kintsch, W. (1988). The role of knowledge in discourse comprehension: A construction-integration model. *Psychological Review*, 95(2), 163–182.

- Klahr, D. & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, 12(1), 1–48.
- KMK. Kultusministerkonferenz. (2024). Weiterentwickelte Bildungsstandards in den Naturwissenschaften für das Fach Chemie (MSA). https://www.kmk.org/fileadmin/Dateien/veroeffentlichungen_beschluesse/2024/2024_06_13-WeBiS_Chemie_MSA.pdf
- Knecht, P., & Najvarová, V. (2010). How do students rate textbooks? A review of research and ongoing challenges for textbook research and textbook production. *Journal of Educational Media, Memory, and Society*, 2(1), 1–16. <https://doi.org/10.3167/jemms.2010.020101>
- Koenen, J. (2014). *Entwicklung und Evaluation von experimentgestützten Lösungsbeispielen zur Förderung naturwissenschaftlich-experimenteller Arbeitsweisen*. Logos.
- Kölbach, E. (2011). *Kontexteinflüsse beim Lernen mit Lösungsbeispielen*. Logos.
- Krajcik, J., & Shin, N. (2023). Student conceptions, conceptual change, and learning progressions. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds.), *Handbook of research on science education, Volume III* (pp. 121–157). Routledge.
- Layng, T. V. J. (2019). Tutorial: Understanding concepts. Implications for behavior analysts and educators. *Perspectives on Behavior Science*, 42(2), 345–363. <https://doi.org/10.1007/s40614-018-00188-6>
- Lederman, N. G. (2007). Nature of science: Past, present, and future. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 831–879). Erlbaum.
- LeFevre, J.-A. & Dixon, P. (1986). Do written instructions need examples? *Cognition and Instruction*, 3(1), 1–30. https://doi.org/10.1207/s1532690xci0301_1
- Leopold, C., & Leutner, D. (2012). Science text comprehension: Drawing, main idea selection, and summarizing as learning strategies. *Learning and Instruction*, 22(1), 16–26. <https://doi.org/10.1016/j.learninstruc.2011.05.005>
- Leppink, J., Paas, F. G., van Gog, T., van der Vleuten, C. P. & van Merriënboer, J. J. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, 30, 32–42. <https://doi.org/10.1016/j.learninstruc.2013.12.001>
- Lodge, W. (2021). ‘Complex and confusing’: The language demands of school science texts. *Research in Science & Technological Education*, 39(4), 489–505. <https://doi.org/10.1080/02635143.2020.1772740>
- Mar, R. A., Li, J., Nguyen, A. T. P., & Ta, C. P. (2021). Memory and comprehension of narrative versus expository texts: A meta-analysis. *Psychonomic Bulletin & Review*, 28(3), 732–749. <https://doi.org/10.3758/s13423-020-01853-1>
- Martin, M. O., Mullis, I. V. S., Foy, P., & Stanco, G. M. (2012). *TIMSS 2011 International results in science*. TIMSS & PIRLS International Study Center.
- Mayer, R. E. (2020). *Multimedia Learning* (3rd ed.). Cambridge University Press.
- Mayer, R. E., & Fiorella, L. (Eds.). (2021). *The Cambridge Handbook of Multimedia Learning* (3rd edition). Cambridge University Press. <https://doi.org/10.1017/9781108894333>
- Mayer, R. E., & Pilegard, C. (2014). Principles for managing essential processing in multimedia learning: Segmenting, pre-training, and modality Principles. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 316–344). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.016>
- Meyer, D., & Pietzner, V. (2022). Reading textual and non-textual explanations in chemistry texts and textbooks: A Review. *Chemistry Education Research and Practice*, 23(4), 768–785. <https://doi.org/10.1039/d2rp00162d>
- Mikkilä-Erdmann, M. (2002). Science learning through text: The effect of text design and text comprehension skills on conceptual change. In M. Limón & L. Mason (Eds.), *Reconsidering*

- conceptual change: Issues in theory and practice* (pp. 337–353). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-47637-1_17
- Moreno, R. (2006). When worked examples don't work: Is cognitive load theory at an Impasse? *Learning and Instruction*, 16(2), 170–181. <https://doi.org/10.1016/j.learninstruc.2006.02.006>
- Mostrady, A., Sanchez-Lopez, E., & Gonzalez-Sanchez, A. F. (2025). Microlearning and its effectiveness in modern education: A mini review. *Acta Pedagogica Asiana*, 4(1), 33–42. <https://doi.org/10.53623/apga.v4i1.496>
- Nakleh, M. B., & Mitchell, R. C. (1993). Concept learning versus problem solving: There is a difference. *Journal of Chemical Education*, 70(3), 190–192.
- National Research Council [NRC] (Ed.). (2011). *A framework for K–12 science education: Practices, crosscutting concepts, and core ideas*. National Academy Press.
- National Research Council [NRC] (Ed.). (1996). *National Science Education Standards*. National Academy Press.
- Nokes, T. J., Hausmann, Robert G. M., VanLehn, K., & Gershman, S. (2011). Testing the instructional fit hypothesis: The case of self-explanation prompts. *Instructional Science*, 39(5), 645–666.
- Osborne, J. (2023). Science, Scientific Literacy, and Science Education. In N. G. Lederman, D. L. Zeidler, & J. S. Lederman (Eds.), *Handbook of Research on Science Education, Volume III* (pp. 785–816). Routledge. <https://doi.org/10.4324/9780367855758-30>
- Paas, F. G. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84(4), 429–434. <https://doi.org/10.1037//0022-0663.84.4.429>
- Parchmann, I., Hansen, S., Diemel, P., Heyn, T., Grubert, A., & Vollersen, A. (2023). Connecting chemical worlds: Context-based learning co-developed. *Chimia*, 77(10), 668–671. <https://doi.org/10.2533/chimia.2023.668>
- Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66(2), 211–227.
- Puspaningtyas, A. A., Hernani, & Suhandi, A. (2020). Analysis on readability of scientific literacy enrichment book on earth science concept. *Journal of Physics: Conference Series*, 1521(4), 42103. <https://doi.org/10.1088/1742-6596/1521/4/042103>
- Pyburn, D. T. & Pazicni, S. (2014). Applying the multilevel framework of discourse comprehension to evaluate the text characteristics of general chemistry textbooks. *Journal of Chemical Education*, 91(6), 778–783. <https://doi.org/10.1021/ed500006u>
- Reimann, P. (1997). *Lernprozesse beim Wissenserwerb aus Beispielen: Analyse, Modellierung, Förderung*. Huber.
- Reiss, K. & Renkl, A. (2002). Learning to prove: The idea of heuristic examples. *Zentralblatt für Didaktik der Mathematik*, 34(1), 29–35. <https://doi.org/10.1007/BF02655690>
- Renkl, A. (2002). Worked-out examples: Instructional explanations support learning by self-explanations. *Learning and Instruction*, 12(5), 529–556. [https://doi.org/10.1016/S0959-4752\(01\)00030-5](https://doi.org/10.1016/S0959-4752(01)00030-5)
- Renkl, A. (2005). The worked-out example principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 229–245). Cambridge University Press.
- Renkl, A. (2014a). Toward an instructionally oriented theory of example-based learning. *Cognitive Science*, 38(1), 1–37. <https://doi.org/10.1111/cogs.12086>
- Renkl, A. (2014b). The worked example principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 391–412). Cambridge University Press.
- Renkl, A. (2021). The worked example principle in multimedia learning. In R. E. Mayer & L. Fiorella (Eds.), *The Cambridge handbook of multimedia learning* (3rd edition, pp. 231–240). Cambridge University Press. <https://doi.org/10.1017/9781108894333.023>

- Renkl, A., & Atkinson, R. K. (2012). Learning from Worked-Out Examples and Problem Solving. In J. L. Plass, R. Moreno, & R. Brünken (Eds.), *Cognitive Load Theory* (pp. 91–108). Cambridge University Press. <https://doi.org/10.1017/CBO9780511844744.007>
- Renkl, A., Hilbert, T., & Schworm, S. (2009). Example-based learning in heuristic domains: A cognitive load theory account. *Educational Psychology Review*, 21(1), 67–78. <https://doi.org/10.1007/s10648-008-9093-4>
- Renkl, A., Stark, R., Gruber, H., & Mandl, H. (1998). Learning from worked-out examples: The effects of example variability and elicited self-explanations. *Contemporary Educational Psychology*, 23(1), 90–108. <https://doi.org/10.1006/ceps.1997.0959>
- Roberts, D. A. (2007). Scientific literacy/ science literacy. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 729–780). Erlbaum.
- Roberts, D. A., & Bybee, R. W. (2014). Scientific literacy, science literacy, and science education. In N. G. Lederman & S. K. Abell (Eds.), *Handbook of research on science education, Volume II* (pp. 545–558). Routledge.
- Roelle, J., Hiller, S., Berthold, K., & Rumann, S. (2017). Example-based learning: The benefits of prompting organization before providing examples. *Learning and Instruction*, 49, 1–12. <https://doi.org/10.1016/j.learninstruc.2016.11.012>
- Roy, M., & Chi, M. T. (2005). The self-explanation-principle. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 271–286). Cambridge University Press.
- Rusanen, A.-M., & Pöyhönen, S. (2013). Concepts in change. *Science & Education*, 22(6), 1389–1403. <https://doi.org/10.1007/s11191-012-9489-x>
- Safari, M., Montazeri, & Mana Mohaghehg (2017). The effect of reducing lexical and syntactic complexity of texts on reading comprehension. *Journal of Teaching Language Skills*, 36(3), 59–83.
- Schroeder, N. L., & Kucera, A. C. (2022). Refutation text facilitates learning: A meta-analysis of between-subjects experiments. *Educational Psychology Review*, 34(2), 957–987. <https://doi.org/10.1007/s10648-021-09656-z>
- Schübler, K., Emden, M., & Sumfleth, E. (2016). *Lösungsbeispiele zu Inhalten des Chemieunterrichts der Sekundarstufe I*. Müller-Reitz-Stiftung. http://www.unidue.de/chemiedidaktik/09_sonstiges_downloads_loesungsbeispiele.php
- Schübler, K. (2017). *Lernen mit Lösungsbeispielen im Chemieunterricht: Einflüsse auf Lernerfolg, kognitive Belastung und Motivation*. Logos.
- Scott, P. H., Asoko, H. M., & Driver, R. H. (1992). Teaching for conceptual change: A review of strategies. In R. Duit, F. Goldberg, & H. Niedderer (Eds.), *Research in physics learning: Theoretical issues and empirical studies*. (pp. 310–329). Institut für die Pädagogik der Naturwissenschaften.
- Scott, P. H., Asoko, H., & Leach, J. (2007). Student conceptions and conceptual learning in science. In S. K. Abell & N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 31–56). Erlbaum.
- Sinatra, G. M., & Broughton, S. H. (2011). Bridging reading comprehension and conceptual change in science education: The promise of refutation text. *Reading Research Quarterly*, 46(4), 374–393. <https://doi.org/10.1002/RRQ.005>
- Siegler, R. S. (2009). Microgenetic studies of self-explanation. In N. Granott & J. Parziale (Eds.), *Microdevelopment* (pp. 31–58). Cambridge University Press. <https://doi.org/10.1017/CBO9780511489709.002>
- Sjöström, J., & Eilks, I. (2018). Reconsidering different visions of scientific literacy and science education based on the concept of Bildung. In Y. J. Dori, Z. R. Mevarech, & D. R. Baker (Eds.), *Innovations in science education and technology. Cognition, metacognition, and culture in STEM education* (Vol. 24, pp. 65–88). Springer International Publishing. https://doi.org/10.1007/978-3-319-66659-4_4

- Soyibo, K. (1996). A comparison of communication strategies among three Caribbean high-school biology textbooks. *Journal of Biological Education*, 30(3), 190–194.
<https://doi.org/10.1080/00219266.1996.9655502>
- Spanjers, I. A., van Gog, T., Wouters, P., & van Merriënboer, J. J. (2012). Explaining the segmentation effect in learning from animations: The role of pausing and temporal cueing. *Computers & Education*, 59(2), 274–280. <https://doi.org/10.1016/j.compedu.2011.12.024>
- Stark, R. (1999). *Lernen mit Lösungsbeispielen. Einfluss unvollständiger Lösungsbeispiele auf Beispielelaboration, Lernerfolg und Motivation*. Hogrefe.
- Stark, R., Tyroller, M., Krause, U.-M., & Mandl, H. (2008). Effekte einer metakognitiven Promptingmaßnahme beim situierten, beispielbasierten Lernen im Bereich Korrelationsrechnung. *Zeitschrift für Pädagogische Psychologie*, 22(1), 59–71.
<https://doi.org/10.1024/1010-0652.22.1.59>
- Strike, K. A., & Posner, G. J. (1982). Conceptual change and science teaching. *European Journal of Science Education*, 4(3), 231–240. <https://doi.org/10.1080/0140528820040302>
- Surif, J., Ibrahim, N. H., & Dalim, S. F. (2014). Problem Solving: Algorithms and Conceptual and Open-ended Problems in Chemistry. *Procedia - Social and Behavioral Sciences*, 116, 4955–4963.
<https://doi.org/10.1016/j.sbspro.2014.01.1055>
- Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction*, 2(1), 59–89.
https://doi.org/10.1207/s1532690xci0201_3
- Tennyson, R. D. & Park, O.-C. (1980). The teaching of concepts: A review of instructional design research literature. *Review of Educational Research*, 50(1), 55–70.
- Thompson, B., Bunch, Z., & Popova, M. (2023). A review of research on the quality and use of chemistry textbooks. *Journal of Chemical Education*, 100(8), 2884–2895.
<https://doi.org/10.1021/acs.jchemed.3c00385>
- van Gog, T., Kester, L., & Paas, F. (2011). Effects of worked examples, example-problem, and problem-example pairs on novices' learning. *Contemporary Educational Psychology*, 36(3), 212–218.
<https://doi.org/10.1016/j.cedpsych.2010.10.004>
- van Gog, T., & Rummel, N. (2010). Example-based learning: Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review*, 22(2), 155–174.
<https://doi.org/10.1007/s10648-010-9134-7>
- van Loon, M. H., Dunlosky, J., van Gog, T., van Merriënboer, J. J. & Bruin, A. B. de. (2015). Refutations in science texts lead to hypercorrection of misconceptions held with high confidence. *Contemporary educational psychology*, 42, 39–48.
<https://doi.org/10.1016/j.cedpsych.2015.04.003>
- van Vorst, H. (2018). Structuring learning processes by ladders of learning: Results from an implementation study. *Chemistry Education Research and Practice*, 19(4), 1081–1095.
<https://doi.org/10.1039/C8RP00078F>
- Vojříř, K., & Rusek, M. (2019). Science education textbook research trends: A systematic literature review. *International Journal of Science Education*, 41(11), 1496–1516.
<https://doi.org/10.1080/09500693.2019.1613584>
- Ward, M. & Sweller, J. (1990). Structuring effective worked examples. *Cognition and Instruction*, 7(1), 1–39. https://doi.org/10.1207/s1532690xci0701_1
- Wylie, R., & Chi, M. T. H. (2014). The self-explanation principle in multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (2nd edition, pp. 413–432). Cambridge University Press. <https://doi.org/10.1017/CBO9781139547369.021>
- Zhu, X., Chen, A., Ennis, C., Sun, H., Hopple, C., Bonello, M., Bae, M., & Kim, S. (2009). Situational interest, cognitive engagement, and achievement in physical education. *Contemporary Educational Psychology*, 34(3), 221–229. <https://doi.org/10.1016/j.cedpsych.2009.05.002>

- Zhu, X. & Simon, H. A. (1987). Learning mathematics from examples and by doing. *Cognition and Instruction*, 4(3), 137–166. https://doi.org/10.1207/s1532690xci0403_1
- Zoller, U. (2002). Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. *International Journal of Science Education*, 24(2), 185–203. <https://doi.org/10.1080/09500690110049060>