Volume 29, Number 4 • Winter 2025

ISSN: 1087-3430

Electronic Journal for Research in Science & Mathematics Education

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

ICRSME

PUBLISHER

ICRSME, International Consortium for Research in Science & Mathematics Education

Cover designed by Dusty Crocker

CONTENTS

Volume 29 No. 4 | Winter 2025/2026

RESEARCH / EMPIRICAL

Electronic Journal for Research in Science & Mathematics Education (EJRSME)

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 29, NO. 4, 1-28

Authority, Autonomy, and Agency in Mathematics Education Research: A Systematic Review of Conceptualizations

Daniel Edelen
Georgia State University

Sarah B. Bush
University of Central Florida

Karen S. Karp
University of Louisville

Audra Skukauskaitė
University of Central Florida

Farshid Safi
University of Central Florida

Sherron Killingsworth Roberts
University of Central Florida

ABSTRACT

In this systematic review, we examine the conceptualization and historical grounding of the terms authority, autonomy, and agency within mathematics education research. These constructs are central to understanding power dynamics and fostering equitable participation in mathematical learning environments. Our review includes 36 empirical studies published up to 2021, analyzing their definitions, theoretical foundations, and intertextual references. Through a taxonomic and domain analysis, we identify seven distinct domains: mathematical authority, authority structures, authority relationships, autonomy as choice, sociomathematical autonomy, agency of the self, and agency through racial identities. Findings highlight the field's reliance on foundational theories, such as Weber's framework of authority, Piaget's developmental perspectives on autonomy, and Bandura's conceptualization of agency, often without deep engagement with their implications for contemporary educational contexts. While these constructs are frequently invoked, their inconsistent definitions and overlapping usage create conceptual ambiguity. Our analysis underscores the need for greater theoretical clarity and attention to the collective dimensions of autonomy and agency, which remain underexplored. We call on researchers to critically engage with the historical and epistemological roots of these constructs, explore their intersections, and prioritize equity-focused research. By offering a detailed taxonomy, this review provides a foundation for advancing theoretical precision and practical application in mathematics education.

Keywords: Authority, autonomy, agency, term analysis

Introduction

In mathematics education, relationships of power profoundly shape who is seen as knowledgeable, competent, and capable of autonomous action (Langer-Osuna & Esmonde, 2017). These dynamics influence classroom interactions within the broader development of learner identities which are constructed moment-by-moment and over time (Dunleavy, 2015; Gresalfi & Cobb, 2006; Gresalfi et al., 2009). Addressing these relationships requires an understanding of how power is distributed, negotiated, and contested within mathematics classrooms. To this end, researchers have explored power dynamics from diverse theoretical and methodological perspectives, including positional analyses (e.g., Wagner & Herbel-Eisenmann, 2014a; Wood, 2016), narrative approaches (e.g., Langer-Osuna, 2016), and interactional perspectives (e.g., Gresalfi & Cobb, 2006; Gresalfi et al., 2009).

Central to these discussions are the constructs of authority, autonomy, and agency, which have substantial implications for understanding power in mathematics education. These constructs are frequently central to efforts to design equitable learning environments. For instance, the National Council of Teachers of Mathematics' *Catalyzing Change* series (NCTM, 2018; 2020a; 2020b) explicitly calls for fostering student agency and shifting authority in ways that support equitable participation. However, despite their widespread use, authority, autonomy, and agency are often poorly defined within the research literature. Their overlapping conceptualizations, frequent interchangeable usage, and lack of clarity contribute to theoretical ambiguity and impede the development of actionable frameworks for understanding and addressing power in mathematics education.

This gap in clarity and precision highlights a pressing need to critically examine how these constructs have been defined, theorized, and operationalized over time. By tracing their histories and identifying their epistemological and ontological underpinnings, researchers can gain a deeper understanding of the assumptions that shape current scholarship and practice. Moreover, clarifying these constructs is essential for advancing equity in mathematics education, as vague or inconsistent definitions risk reinforcing, rather than challenging, existing power hierarchies.

In this paper, we seek to illuminate the histories, conceptualizations, and theoretical groundings of authority, autonomy, and agency in mathematics education. Using a systematic review of 36 empirical studies, we analyze how these three constructs have been defined and employed in the field. By categorizing these studies and identifying patterns across time and contexts, we aim to provide clearer distinctions and definitions that can guide future research and practice.

Aim of the Paper and Research Question

This paper aims to provide a systematic review of the math education research literature to illuminate the specific histories, traditions, and approaches to authority, autonomy, and agency in mathematics education research. By analyzing how these constructs have been conceptualized over time, we seek to clarify their definitions, theoretical groundings, and the implications for understanding relationships of power in mathematics education.

Research Question

We center our review on the research question: *How have authority, autonomy, and agency been conceptualized in mathematics education research over time?*

Through this systematic review, we define the varying conceptualizations of authority, autonomy, and agency, highlighting distinctions and overlaps across empirical studies. Additionally, we report on the epistemological underpinnings of each construct and explore their implications for research on power dynamics in mathematics education.

Methods

Literature Search Procedures

To conduct this systematic review, we adhered to the guidelines of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Alexander, 2020). We searched five primary academic databases: Academic Search Complete, Education FullText, ERIC, JSTOR, and ProQuest. These databases were selected to ensure comprehensive coverage of mathematics education research while maintaining focus on high-quality, peer-reviewed sources. Academic Search Complete, Education FullText, and ERIC were chosen based on Alexander's (2020) recommendations to ensure saturation of the literature. JSTOR was used to supplement these databases as it houses key journals in mathematics education (e.g., *Journal of Research in Mathematics Education, Educational Studies in Mathematics,* and *For the Learning of Mathematics*). ProQuest was intentionally included to capture published dissertations and expand the dataset beyond traditional journal articles.

We included studies published up to 2021, as our goal was to capture the historical development of the terms "authority," "autonomy," and "agency" in mathematics education research. The search was conducted using these terms as keywords, requiring their presence in the title, abstract, or keywords of the identified studies. While books, theoretical/philosophical articles, and conference proceedings were not included in the empirical dataset, their contributions to the conceptual framing of authority, autonomy, and agency are acknowledged in the discussion section of this paper. The exclusion of these sources was guided by the focus of this review on empirical studies that provide direct evidence of how these constructs are conceptualized in educational practice. We discuss this more in the following section.

Inclusion/Exclusion Criteria

In selecting studies for this review, we developed clear criteria to ensure our analysis remained on the constructs of authority, autonomy, and agency in mathematics education research conceptualized and operationalized in empirical research. Our primary inclusion criteria was whether these terms were explicitly central to the study. Thus, a study must place one or more of these constructs at the center of its research questions, design, or analysis. We included only empirical studies that provided sufficient methodological detail (e.g., on data collection, study context, and analytic approach) to allow for systematic comparison and synthesis across studies.

To maintain a coherent and methodologically rigorous dataset, we excluded non-empirical sources, such as theoretical or philosophical papers, policy documents, and conceptual essays. These works, while often important for understanding the constructs in question, do not offer the kind of empirical grounding required for our domain and taxonomic analysis. However, because many of these texts are frequently cited within the 36 empirical studies, we will highlight select non-empirical works in our discussion to support the conceptual framing of our findings. In this way, our engagement with non-empirical sources is more interpretive, since they were not included in the analysis of the dataset. Yet, we drew upon select non-empirical works to help situate the patterns across empirical studies within broader theoretical conversations.

We also excluded studies that only referenced authority, autonomy, or agency in passing. For example, when some articles only mention the constructs within their implications or conclusions, without meaningfully engaging with constructs as part of their analytic focus. In addition, we chose to exclude conference proceedings, given the variability in their peer review standards, and instead focused on journal articles and dissertations to ensure high-quality methodological rigor.

Finally, while our search was limited to studies written or translated into English, and was focused primarily on journal articles and dissertations, we applied no restrictions based on grade level, population, or geographic location. These parameters allowed us to capture a wide range of learning contexts, including studies involving children, teachers, and families, to reflect the diverse settings in which these constructs are often negotiated in mathematics education.

Dataset Construction

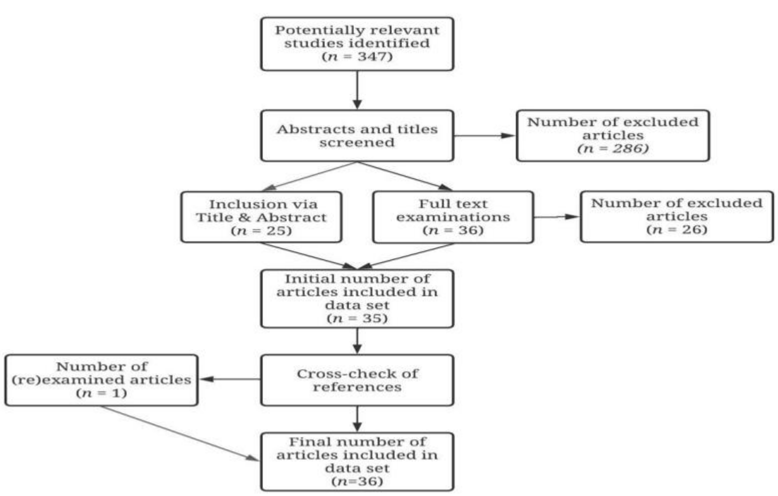
Our initial search, guided by the inclusion/exclusion criteria outlined above, identified 347 studies as potential candidates for this systematic review. These studies were screened in multiple stages to ensure alignment with the research question and focus of this review. Figure 1 outlines the step-by-step construction of the final dataset.

In the first stage, we screened the titles and abstracts of all 347 studies. At this stage, 286 studies were excluded for not meeting one or more of the inclusion criteria, such as the absence of an explicit focus on authority, autonomy, or agency. This process resulted in a subset of 61 studies deemed potentially relevant. The second stage involved a more detailed review of the abstracts of these 61 studies, with particular attention to whether authority, autonomy, or agency was central to their focus. This step identified 25 studies in which one or more of these constructs were explicitly discussed in the title or abstract. These 25 studies were immediately included in the final dataset. For the remaining 36 studies, we conducted a full-text review to determine their alignment with our inclusion criteria. This deeper examination resulted in the inclusion of 10 additional studies that explicitly addressed authority, autonomy, or agency, bringing the total to 35 studies.

To ensure saturation and mitigate the possibility of overlooking key studies, we cross-checked the reference lists of these 35 articles. This cross-referencing process identified one additional study that met the inclusion criteria, which was then added to the final dataset. This step also helped confirm that no major studies within the scope of this review were missed. The final data set comprised 36 articles. Through this process, we ensured that the final dataset reflects the empirical research explicitly centered on authority, autonomy, or agency in mathematics education. The rigorous screening and cross-referencing process provided confidence in the comprehensiveness of the dataset, while also highlighting key works outside the inclusion criteria that contribute to the conceptual understanding of these constructs, which are discussed in subsequent sections.

Figure 1

Flow diagram of study selection process



Logic of Analysis

To explore how authority, autonomy, and agency have been conceptualized in mathematics education research, we adopted a multi-faceted ethnographic research perspective (Green et al., 2015). This approach allowed us to examine published articles as artifacts, or textual representations of the theoretical and methodological choices made by researchers. By positioning ourselves as "readers-as-ethnographic-analysts," (Green et al., 2015, p. 27) we sought to uncover the epistemological roots, theoretical orientations, and conceptual frameworks embedded within these studies. This perspective guided our analysis, enabling us to trace the histories and relationships that underpin the conceptualizations of these three constructs.

Our analysis began with a categorization of studies, grouping them by their primary focus on authority, autonomy, or agency. This initial step provided a foundation for organizing the literature and identifying patterns of emphasis within the field. We then turned to the temporal dimension, constructing a timeline of the included studies, as suggested by Green et al. (2015). Mapping these studies chronologically revealed how the conceptualizations of authority, autonomy, and agency have evolved over time, as well as how certain ideas have shaped, intersected, or diverged across the literature.

To deepen our understanding, we conducted a line-by-line analysis of each article, drawing on Green's (1983) domains to systematically examine key elements such as the study's purpose, definitions, settings, theoretical orientations, and methodologies. This process allowed us to engage closely with the text, uncovering both explicit and implicit ways these constructs were defined and operationalized.

A critical component of our analysis was intertextual mapping (Baron, 2019), which we used to better understand the over-time conceptualization of authority, autonomy, and agency in mathematics education. This method, rooted in the work of Bloome and Egan-Robertson (1993), refers to the juxtaposition of texts, words, and phrases, such as citations or quotations, that appear within and across documents to construct meaning. In academic writing, intertextuality is most literally visible in how authors cite, build upon, or challenge one another's work. We traced these relationships across the dataset by systematically documenting who cited whom, in what ways, and for what purposes. This process allowed us to uncover the interconnections between studies, revealing how constructs were taken up, defined, and evolved across time.

Through mapping, we were able to identify which studies functioned as seminal conceptual anchors, which were cited most frequently for definitional purposes, and how newer studies extended or contested earlier work. Intertextual mapping helped us determine the influence of individual studies and the patterns of conceptual borrowing and alignment that shaped the field's understanding of authority, autonomy, and agency. This lens enabled us to visualize the development of these constructs as an unfolding dialogue rather than a set of isolated contributions, adding depth to our analysis of how meanings have been constructed and sustained over time. Our maps are provided as figures in the following sections.

To identify and analyze across the many conceptualizations of authority, autonomy, and agency within the constructed dataset, we employed a domain and taxonomic analysis following Spradley's (1979/2016) ethnographic methods. Central to this approach is the logic of semantic relationships; we particularly used the relationship of strict inclusion ("X is a kind of Y") to structure each domain within the taxonomies. Domains were constructed through an iterative, recursive, and abductive logic (Agar, 2006), grounded in close textual engagement with each article. As we read across studies, we attended to the language researchers used to name, define, and distinguish constructs and examined the semantic boundaries that authors set between related terms. This analytic approach allowed us to systematically trace how specific conceptualizations of each construct (e.g., social authority or sociomathematical autonomy as a domain) were languaged by researchers within to fit within broader conceptual taxonomic categories (i.e., authority, autonomy, or agency). To identify the specific boundaries authors used to differentiate one form of a construct from another, we posed the same questions for each text: What kind of authority is being described? What are its attributes? How is it situated in relation to other forms of the construct? Through comparative reading, we noted where constructs overlapped, diverged, or evolved across contexts and time.

From these questions and comparative insights, we developed taxonomies to capture the internal organization of each construct and highlighted how conceptual distinctions were constructed, maintained, or refigured within the literature. This process reflects an emic, text-centered logic of inquiry grounded in our ethnographic stance, one that privileges the conceptual language and distinctions visible in the field's own discourse and honors how scholars have come to define authority, autonomy, and agency in mathematics education research.

By combining these analytic approaches, we were able to construct a comprehensive and nuanced picture of the field's engagement with these constructs. This multi-layered process illuminated the epistemological and theoretical underpinnings of the studies to provide insights into how these ideas have been shaped by and have contributed to broader discussions within mathematics education research.

Findings

The findings are based upon the analysis of the 36 reviewed studies. We organized findings around the three main constructs, authority, autonomy, and agency, as conceptualized in the field of mathematics education. We use both the plural and singular "they" when referring to authors of the included studies.

Taxonomy 1: Authority in Mathematics Education

Of the 36 included studies, 21 studies focused on researching and understanding authority as it relates to mathematics education. Based on a domain and a taxonomic analysis, we identified three kinds of authority domains: Mathematical Authority, Authority Structures, and Authority Relationships. Table 1 outlines the included studies within each domain. In the following section, each domain is described, and the characteristics of the findings are articulated.

Table 1

Authority Taxonomy

Taxonomy	*Domain*	*Studies*
Authority	Mathematical Authority	Wilson & Lloyd (2000)
		Hamm & Perry (2002)
		Inglis & Ramos (2009)
		Depaepe et al. (2012)
		Wagner & Herbel-Eisenmann (2014b)
		Dunleavy (2015)
		Kinser-Traut & Turner (2020)
		Solomon et al. (2021)
	Authority Structures	Herbel-Eisenmann & Wagner (2010)
		Wagner & Herbel-Eisenmann (2014a)
		Tatsis et al. (2018)
		Andersson & Wagner (2019)
		Ng et al. (2021)
	Authority Relationships	Amit & Fried (2005)
		Gerson & Bateman (2010)
		de Freitas et al. (2012)
		Langer-Osuna (2016)
		Langer-Osuna (2018)
		Langer-Osuna et al. (2020)
		Langer-Osuna et al. (2021)
		Lai & Baldinger (2021)

Intertextual Mapping of Authority Taxonomy

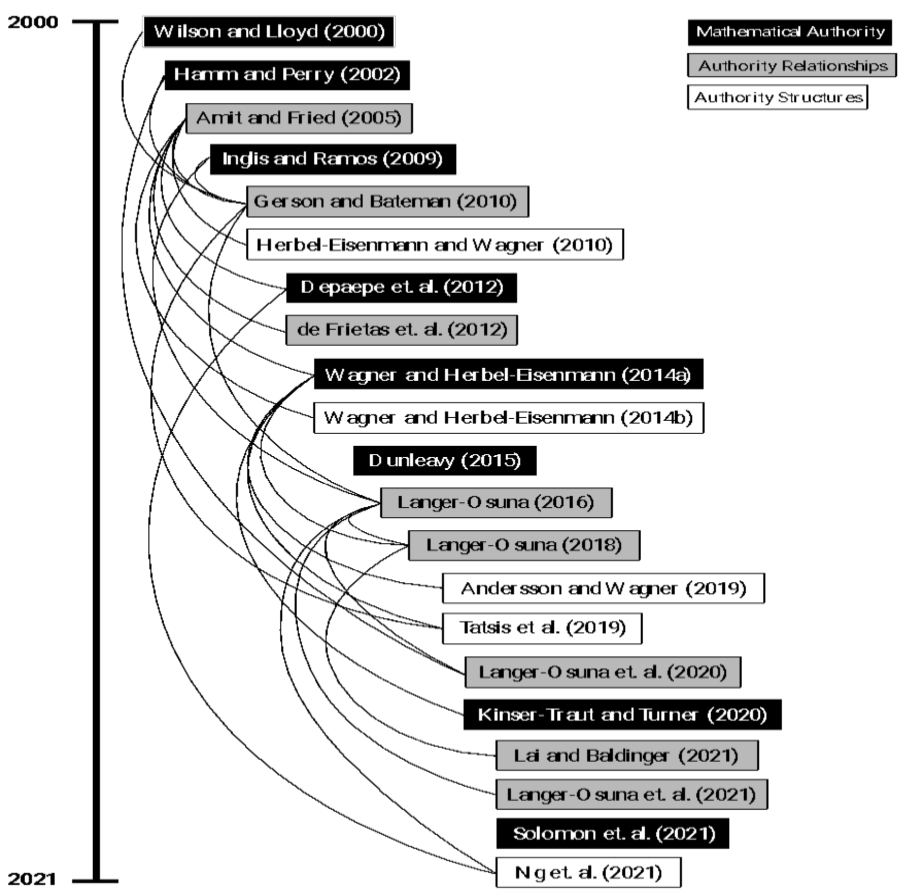
Intertextual mapping of studies made visible several findings of the authority taxonomy. Figure 2 outlines the ways studies intertextually drew on earlier conceptualizations of authority through citations. In the subsequent sections, we outline the specific domains that make up this taxonomy; however, there are several findings that are of interest to the entire taxonomy. Through tracing the citations of studies included in the taxonomy, a clear influence from a single theoretical conceptualization can be seen that has governed how authority has been defined, conceptualized, and studied. That dominant influence is the work of the prominent sociologist Max Weber. Of the 21 studies included in this taxonomy, seven studies directly and indirectly build on Weber's (1947) traditional authority definition. Although sometimes this is identified as a direct citation (i.e., Amit & Fried, 2005; Gerson & Bateman, 2010; Kinser-Traut & Turner, 2020; Langer-Osuna et al. 2020) in other studies Weber's influence can be traced indirectly through citing Pace and Hemmings (2007) for a definition of authority (Herbel-Eisenmann, 2010; Wagner & Herbel-Eisenmann, 2014a, 2014b).

Pace and Hemmings' (2007) study is a review of literature and the direct citation from their paper is in reference to Metz (1978), who summarizes Weber (1947). In Weber's (1925/1947) types of authority, legal authority refers to the rules or laws established through some sort of bureaucratic system. Traditional authority is characterized by social foundations people occupy and from which authority can issue commands (i.e., parent, teacher, mentor). Charismatic authority, in contrast, rests on followers' devotion to the exceptional sanctity, heroism, or exemplary character of an individual leader, granting legitimacy through personal magnetism rather than established rules or tradition (Weber, 1925/1947). Except for Amit and Fried (2005) and Gerson and Bateman (2010), the majority of studies on authority in mathematics education do not explicitly reference a particular type of Weber's authority upon which they are building, such as legal, traditional, or charismatic. Through tracing the definitions, we found that authors in the field of mathematics education most commonly build on Weber's concept of traditional authority.

While only seven studies were explicitly traced to the work of Weber, utilizing the intertextual tracing provided further implications of note. Specifically, the web of citations is vast within this taxonomy, most notably being influenced by Amit and Fried (2005). Only a few studies can be noted as not being influenced by the work of Weber on authority (Hamm & Perry, 2005; Inglis & Mejia-Ramos, 2009; Wilson & Lloyd, 2000). With most studies in this taxonomy being traced back to the work of Weber, understanding how these conceptualizations have adopted and extended Weber's concepts of authority is important and also addressed in the next section. Later, we return to the influence of Weber in the implications of this paper.

Figure 2

Authority Taxonomy Mapping



Mathematical Authority Domain

The domain of mathematical authority encompasses 8 studies of the 21 in the Authority taxonomy. This domain represents studies which have conceptualized authority as a constant in mathematics education; in other words, authority is viewed as something that is delegated (Dunleavy, 2015), shared (Kinser-Traut & Turner, 2020; Wilson & Lloyd, 2000), granted (Depaepe et al., 2012; Hamm & Perry, 2002; Inglis & Mejia-Ramos, 2009), or devolved (Solomon et al., 2021). Within this conceptualization, authority is viewed as a pedagogical tool that teachers utilize as part of their daily mathematical instruction. Authority is a unilateral exchange between teacher and students. In this domain, the teacher is perceived as a constant source of authority, and the focus of these studies trace how authority is distributed to students as a singular entity.

Beginning with Wilson and Lloyd (2000), the focus on mathematical authority centralizes around the process of distributing authority from teacher to students. The concept of distributing authority dictates that authority is ultimately held by the teacher, as both a position and a content expert (Wagner & Herbel-Eisenmann, 2014b). Dunleavy (2015), citing Gresalfi and Cobb (2006) (which was not included in the dataset due to being non-empirical), further articulates this process of distributing authority to focus on the degree to which students are given opportunities to make mathematical contributions within the learning of mathematics. A key distinction in this domain is that students are referred to and operationalized as a group that is viewed as subservient in their relationship to the teacher. Routinely cited in this domain, students occupy a position of receivers of knowledge (e.g., Depaepe et al., 2012). Because of this distinct relationship and conceptualization of authority, the underlying goal of studies in this domain is to examine how authority moves from the teacher to the students. For example, Hamm and Perry (2002) focus on how teachers often hold students "accountable for their mathematical ideas" (p. 135) through the process of distributing authority by inviting students to explain their ideas or thinking during lessons. Similarly, Kinser-Traut and Turner (2020) examine how one teacher began to distribute authority to students by including student-based instructional practices and approaches more frequently than teacher derived ones in whole-class discussions.

Authority Structures Domain

This literature review encompasses five studies for the domain of authority structures. This domain represents studies which have conceptualized authority as structures present in the mathematics classroom with established rules and norms for determining authority between teachers and students. these studies use positioning theory (Davies & Harré, 1990) to determine how people in mathematics classes are positioned as *in authority* or as *an authority* (Skemp, 1979). Wagner and Herbel-Eisenmann (2014a) articulate the distinctions between "being *an* authority because of one's content knowledge and being *in* authority because of one's position" (p. 872). In this domain, authority is again viewed as a constant in classroom-based mathematics, but the goal is to understand how it is structured (Herbel-Eisenmann & Wagner, 2010) and the ways specific authority structures are made visible and influenced by the discursive patterns of the teacher.

This domain builds from the work of Herbel-Eisenmann and Wagner (2010) and their study of lexical bundles in the classroom-based discourse of secondary mathematics educators. Herbel-Eisenmann and Wagner examined what they called stance bundles, or three or more words that frequently occur together in a similar register (e.g., I want you to, I'm going to do, you are going to do) that teachers discursively use to communicate feelings, attitudes, directions or judgments, to their learners. Based upon their analysis, they categorized four types of authority structures in classroom-based mathematics: *personal authority*, *demands of the discourse as authority, more subtle discursive authority*, and *personal latitude*. In their 2010 study, which they subsequently elaborated upon (Wagner & Herbel-Eisenmann, 2014a, Wagner & Herbel-Eisenmann, 2018), the different structures were defined through the lens of the positioning theory and the linguistic cues that illuminate the different structures in the classroom. *Personal authority* describes the ways teachers used personal pronouns (building from Fairclough, 2001) to position students to follow a specific perceived obligation or act in the classroom. Teachers relied on some sort of *personal authority* (as *in* or *an* authority) to provide directives for students to follow with no further justification offered. Indicators of this personal authority structure are evidenced when people follow directives of another without explicit reasoning (Herbel-Eisenmann & Wagner, 2010).

Demands of the *discourse as authority* are marked by the times that an external authority (other than the teacher) is referenced in the exchange between teachers and students. Some examples are visible when a teacher uses the personal pronoun, *we,* in statements such as *we are going to have to.* In later work (Andersson & Wagner, 2019; Tatsis et al., 2018; Wagner & Herbel-Eisenmann 2014a), this kind of authority was referred as *discourse as authority*, to note the explicit strong obligations for students within mathematics classrooms. In the *more subtle discursive authority*, stance bundles were marked as the times teachers were "thinking ahead, but this was a special kind of forward thinking, giving the sense that the speaker knows what will happen" (Herbel-Eisenmann & Wagner, 2010, p. 56). For example, a teacher might reference a test that will happen or reference a future event that a specific mathematics skill might be needed. Later, Wagner and Herbel-Eisenmann (2014a) updated this structure to *discursive inevitability,* to capture language that suggests an inevitable outcome despite the speaker being unaware of the probability of its occurrence. There is no underlying obligation; instead, this structure highlights that the upcoming actions are simply bound to occur. In a sense, there are no decisions to be made. The authority in this structure rests outside of the singular interaction between teacher and student. Finally, Tatsis and colleagues (2018) built upon *personal latitude* which refers to the situations in mathematics classrooms wherein people recognize they and others can make decisions about their actions. These situations are marked by open-ended questions or invitations for additional mathematics ideas or choices.

Authority Relationships Domain

The domain of *authority relationships* encompasses eight studies. This domain represents studies that conceptualized authority as a socially constructed relationship between people in mathematics classrooms. Within this domain, authority is viewed less as a constant; but instead, is examined through the different relations that develop among teachers and students as well as among students during collaborative learning endeavors. The focus of these studies remains within the interactions of people in mathematics classrooms; thus, much of this work involves analysis of particular social positionings. In essence, the *authority relationships* domain represents studies that examine who possesses authority in interactions and the ways authority influences different opportunities for learning mathematics.

Much of this domain stems from the work of Amit and Fried (2005) and their investigation of an eighth-grade mathematics classroom. Their analysis represents the first time in mathematics education research that authority was referred to as a social relationship constructed within the classroom settings. They also provide the most in-depth discussion of authority in educational settings of any of the included studies in this taxonomy. Because of their early work, studies in this domain shift from studying authority as "domination and obedience to negotiation and consent" (Amit & Fried, 2005, p. 164). This shift reconceptualized students from simple receivers of mathematical information to being co-participants in a community of learners who shape and develop different relationships of authority. Gerson and Bateman (2010) build upon this conceptualization from Amit and Fried (2005) to further denote that authority relationships encapsulate three axioms. First, authority is made visible through a relationship between two or more people. Second, authority relationships are illuminated by a change in behavior of one person based upon the actions of another. Third, the person with authority must maintain some sort of legitimacy that is recognized in the interaction.

Lai and Baldinger (2021) also build from Amit and Fried (2005) and their assertion of modeling authority relationships as *expert* or *shared*. According to Lai and Baldinger (2021), expert authority can take the form of teachers who expect to be treated by students as the final arbitrator of what work is produced and whether it is correctly done. Expert authority can also take the form of students who look to teachers to be told what to believe (Lai & Baldinger, 2021). Lai and Baldinger (2021) further note that "in contrast, shared authority leaves open the possibility that students can learn to be effective and legitimate arbiters of what mathematical work to take up and whether the reasoning holds" (p. 26). Much of the focus here, and the work that has built upon Amit and Fried (2005) is the relation between the students and the teacher during mathematics instruction. Lai and Baldinger (2021) even state that "authority relationships become visible in the ways students and teachers talk with one another" (p. 27).

Of the eight studies in this *authority relationships* domain, four are the work of Langer-Osuna and colleagues, which explicitly focus on the authority relationships among student peer interactions. Three studies (Langer-Osuna, 2016, 2018; Langer-Osuna et al., 2020) directly build from the *influence framework* (Engle et al., 2014) wherein the conceptualization of authority is further articulated to describe two specific types of authority: social and intellectual. Langer-Osuna (2016) first defines social authority as "the authority to issue directives to peers in the management of group dynamics" (p. 109) and later refines the definition in terms of relations between people. *Social authority* relations are enacted through interactions that position students as having the right to issue directives to their peers" (Langer-Osuna et al., 2020, p. 337). Langer-Osuna (2016) defines intellectual authority, through the lens of positioning theory (Davies & Harré, 1990), as "the positioning of students as credible sources of information pertinent to the particular task at hand" (p. 109). They further conceptualize this type of authority to again focus on relations between and among people by articulating that "intellectual authority relations are enacted through interactions that position students as credible sources of mathematical information" (Langer-Osuna et al., 2020, p. 337).

Clearly, much of the focus of authority relationships examines human interactions within classroom spaces. However, de Freitas and colleagues (2012) also assert that specific classroom-based objects might also exhibit authority in classrooms (e.g., the textbook, whiteboard, or anchor charts). While this addition of inanimate objects to the conceptualization of authority is briefly mentioned here, the inclusion of objects as authority do not reappear in other studies within this domain, revealing a present gap in understanding.

Taxonomy 2: Autonomy in Mathematics Education

Of the 36 studies included in this analysis, 7 explicitly highlight autonomy as it relates to mathematics education. Based upon a domain and a taxonomic analysis, two kinds of autonomy domains were made visible: autonomy as choice and sociomathematical autonomy. Table 2 outlines the included studies within each of the domains. In the following section, each domain is described, and the characteristics of the findings are articulated.

Table 2

Autonomy Studies Taxonomy

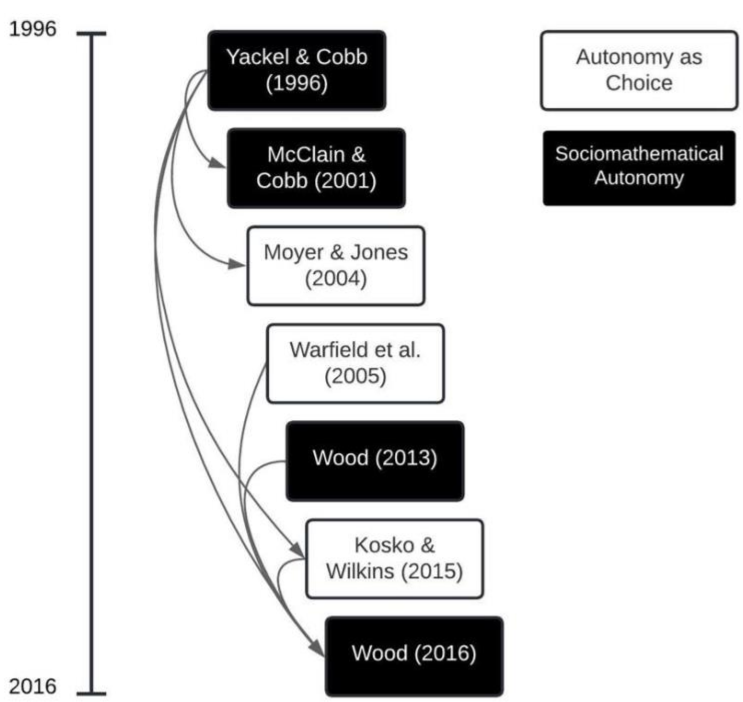
Taxonomy	*Domain*	*Studies*
Autonomy	Autonomy as choice	Moyer & Jones (2004)
		Warfield et al. (2005)
		Kosko & Wilkins (2015)
	Sociomathematical autonomy	Yackel & Cobb (1996)
		McClain & Cobb (2001)
		Wood (2013)
		Wood (2016)

Intertextual Mapping of Autonomy

Through the intertextual mapping of the included studies in the autonomy taxonomy, we uncovered that this subfield of mathematics education research is relatively small and not recently explicitly studied. In fact, with the exception of Wood (2016), studies in this autonomy taxonomy branch from one study that explicitly researched autonomy: Yackel and Cobb (1996). Figure 3 displays the intertextual citations in this taxonomy.

Figure 3

Autonomy Taxonomy Mapping



Autonomy as Choice

The domain of *Autonomy as Choice* encompasses three studies (Kosko & Wilkins, 2015; Moyer & Jones, 2004; Warfield et al., 2005). This domain represents authors whose studies conceptualized autonomy as giving students choice in their pursuits to do mathematics. These studies approach autonomy through the lens of freedom for the students in the classroom. For example, Kosko and Wilkins (2015) defined autonomy through students' individual "sense of control in the manner one engages in doing mathematics, while maintaining a sense of freedom in their engagement with mathematics" (p. 371). These studies focused on creating opportunities for students to freely engage in mathematics content, specifically that of whole class discussions. Warfield et al. (2005) focused on acts that were determined to be *autonomous,* or acts wherein a person senses that choices are free of outside influences. Meanwhile, Moyer and Jones (2004) focused more on shifting control from mathematics teachers to offer opportunities for students to self-select or choose preferred mathematics manipulatives during learning activities. In essence, the studies in this domain maintained that autonomy is creating opportunities for students or teachers through choices in their mathematical learning endeavors.

Sociomathematical Autonomy Domain

The domain of sociomathematical autonomy encompasses four studies (McCain & Cobb, 2001; Wood, 2013; Wood, 2016; Yackel & Cobb, 1996). This domain represents and builds on studies wherein authors conceptualize autonomy as being co-constructed through the practices of students and their teacher through mathematical learning opportunities. Here, autonomy is conceptualized as more than simply providing students choice in their use of manipulatives, representation procedures, or even correct answers. Instead sociomathematical autonomy maintains that students must also possess the freedom to decide and construct what counts as mathematics (Yackel & Cobb, 1996) and what it means to do mathematics (Wood, 2013, 2016).

Beginning with Yackel and Cobb (1996), autonomy is characterized in two ways: *social* and *intellectual*. Yackel and Cobb (1996) do not fully define social autonomy; instead, social autonomy is briefly mentioned as a benefit of inquiry-based approaches to teaching mathematics. By tracing cited studies (i.e., Cobb et al., 1991), we depended on Cobb and colleagues' chapter on radical constructivism, where we were able to define social autonomy. Here, social autonomy is conceptualized through Piaget's (1948/1973) notions of autonomous actions of children, namely the freedom to explore and experiment with the world around them. Thus, Yackel and Cobb (1996) conceptualize social autonomy as the freedom to interact with mathematics, peers, and mathematical tools.

Likewise, Yackel and Cobb's (1996) study is devoted to the development of what they refer to as intellectual autonomy. Yackel and Cobb (1996) cite Kamii (1985) to define and conceptualize intellectual autonomy as:

The conception of autonomy as a context-free characteristic of the individual is rejected. Instead, autonomy is defined with respect to students' participation in the practices of the classroom community. In particular, students who are intellectually autonomous in mathematics are aware of, and draw on, their own intellectual capabilities when making mathematical decisions and judgements as they participate in these practices (Kamii, p. 473).

Building on conception of intellectual autonomy from Yackel and Cobb (1996), Wood (2016) asserts that the definition of autonomy "reemphasizes the need for autonomous activity to include a decision about truth and untruth" (p. 331). They further ground the conceptualization of autonomy in the work of Piaget (1948/1957), in asserting that "intellectual autonomy is more than having a choice and more than having an answer. It is the student's process of reasoning about mathematical ideas by herself" (Wood, 2016, p. 331). Wood (2016) also presents autonomy through a communication lens as a students' intellectual autonomy in how they "wrestle with truth and untruth" (p. 332) of mathematical narratives in the classroom. Wood (2013) asserts that intellectual autonomy is crucial in students being able to communicate is more than the simple revoicing of their peers' or classroom teachers' thinking. Simply stated, intellectual autonomy is focused on students making decisions and communicating about what it means to do mathematics and for what purposes.

Agency in Mathematics Education

Of the 36 studies included in this review, eight explicitly focus on agency as it relates to mathematics education. Based upon our analysis, we made visible two kinds of autonomy: agency of the self and agency and racial identity. Table 3 outlines the included studies within the two autonomy domains. In the following sections, we outline the intertextual mapping, describe each domain, and articulate the characteristics of the findings.

Table 3

Agency Studies Taxonomy

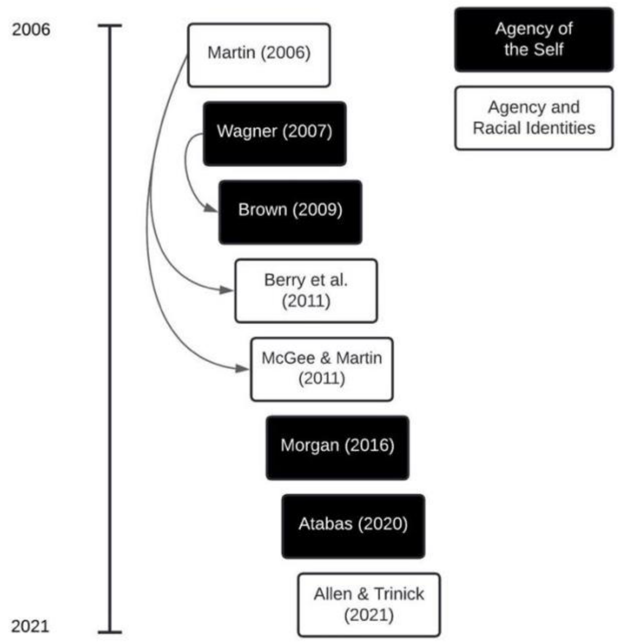
Taxonomy	*Domain*	*Studies*
Agency	Agency of the self	Wagner (2007)
		Brown (2009)
		Morgan (2016)
		Atabas et al. (2020)
	Agency and racial identities	Martin (2006)
		Berry et al. (2011)
		McGee & Martin (2011)
		Allen & Trinick (2021)

Intertextual Mapping Agency in Mathematics Education

The intertextual mapping of studies included in this agency taxonomy makes clear how disconnected the field is in terms of citations and connecting to prior work that explicitly researches agency in mathematics education. Figure 4 displays the studies included in this domain. Of the eight included studies in this literature review, only three studies explicitly cite prior work in this domain, furthering evidence of the isolated nature of studies that explicitly research agency. The domains of agency of the self and agency and racial identities are discussed below.

Figure 4

Agency Taxonomy Mapping



Agency of the Self

The domain of *Agency of the self* encompasses four studies (Atabas et al., 2020; Brown, 2009; Morgan, 2016; Wagner, 2007). Studies were included in this domain because they conceptualized agency through philosophical investigations of the *self* within mathematics classrooms. These studies view mathematics as a social institution that is made visible and constructed through discourse; thus, constructing the *self* is as much an individual as it is a discursive process within the larger mathematical culture for learning. Here, *self* refers to the individual within the larger space of the classroom. However, the individual is not understood as a singular person, but instead as a person who interacts and constantly develops through and with their peers.

Wagner (2007), Brown (2009) and Morgan (2016) view the *self* as individual awareness of how a singular student begins to understand their role within larger classroom discourses. Wagner (2007) and Morgan (2016) refer to this kind of agency as *human agency*. While both Wagner (2007) and Morgan (2016) use the term *human agency,* they each build from different theoretical groundings. Wagner (2007) uses the work of Pickering (1995). Wagner (2007) asserts the guiding question for conceptualizing agency, "who is said to be making things happen?" (p. 37). Morgan (2016) instead focuses on the "philosophical debates on the nature of mathematical discovery" (p. 123), and if mathematics is viewed as a human act. Here, Morgan referred to the understanding of *human agency* as it relates to how mathematical discoveries are conceptualized in school mathematics. Morgan (2016) questioned, where the *self* is in discussions of the origins of mathematical discoveries in school mathematics. While, Wagner (2007) and Morgan (2016) use the term *human agency,* Brown (2009) builds on a different concept of the *self*. Brown (2009) draws on the work of Cobb and Hodge (2002) when discussing agency. Brown (2009) focuses on how students begin to understand their abilities to be aware of the "social positions" (p. 182) students navigate while doing mathematics. Brown (2009) directly builds from Wagner (2007) when describing the importance of pronouns (i.e., I, me, you, her, we) to identify shifts in students' views of themselves in relation to the classroom community. Brown (2009) denoted how students view themselves as both mathematicians and as individuals who contribute and use a community of learners to know and do mathematics.

Similarly, Atabas et al.'s (2020) article is included in this domain, because the research focuses on the *self* through middle school students in classroom-based mathematics and the ways they understand their role within larger mathematical discourses. While still included in this domain, Atabas et al.'s (2020) study remains separate from the other three studies, because it approaches agency differently from the other studies. Atabas et al. (2020) researched agency through the concept of authority and autonomy, which is particularly troublesome due to a lack of clear theoretical grounding of either. Below, we include Atabas and colleagues' (2020) definitions of agency to illuminate their conceptualization as defined through other researchers.

Disciplinary agency (in the context of mathematics), involves the use of established procedural skills for computing the solution to a problem (Cobb et al., 2009; Grootenboer & Zevenbergen, 2007; Hull & Greeno, 2006). When the teacher is the authority, students may be provided few opportunities to reason mathematically to make sense of problems. Students instead must rely on the methodologies. Provided by the teacher—they are engaging only in disciplinary agency. (Atabas et al., 2020, p. 3)

Note that Atabas and colleagues (2020) use all three terms, authority, agency, and autonomy, when defining conceptual or disciplinary types of agency, which will be discussed in more detail in the Discussions and Implications section. In intertextually tracing their definition, we note that Cobb et al. (2009) categorize agency as *conceptua*l or *disciplinary* in nature. *Conceptual agency* involves student autonomy in which students are responsible for developing their own understanding of relationships between concepts (Cobb et al., 2009; Grootenboer & Zevenbergen, 2007). When authority is shared with students, students are positioned to understand when and for what purposes to use disciplinary tools to solve problems (Boaler & Greeno, 2000). In the following section, the second domain findings from the agency domain are described.

Agency and Racial Identity Domain

This second domain of agency encompasses four studies (Allen & Trinick, 2021; Berry et al., 2011; Martin, 2006; McGee & Martin, 2011). Studies were included in this domain because they have conceptualized agency through the lens of racial identities in mathematics education. In this domain of agency and racial identity, three of the studies explicitly focus on African American (Martin, 2006) and/or Black students (Berry et al., 2011; McGee & Martin, 2011). The fourth study focuses on the Indigenous Māori people of Aotearoa, New Zealand (Allen & Trinick, 2021). Except for Allen and Trinick (2021), the authors included in this domain did not formally define or conceptualize agency; instead, they used the term "agency" to describe specific actions of African American parents (Martin, 2006) and Black students (Berry et al., 2011; McGee & Martin, 2011). Beginning with Martin (2006), from which Berry et al. (2011) and McGee and Martin (2011) directly build, agency is used as way to articulate behaviors associated with individual actions to promote positive African American/ Black identities in mathematics education. To better understand this kind of agency, we traced the intertextual references to Martin (2000). All studies referenced Martin's (2000) book when describing agency in their studies (Berry et al., 2011; Martin, 2006; McGee & Martin, 2011). Martin's (2000) book, *Mathematics success and failure among African-American youth,* was not included in our review based on our inclusion criteria of only scholarly articles.

In Martin's (2000) book, agency is referenced in relation to the work of Bandura's (1986) Social Cognitive Theory. According to Bandura (1986), agency is the ability to influence the course of events of which one is a part. In this domain, agency is used to document those times that individuals influenced the course of events as they relate to their learning of mathematics. In conceptualizing agency through racial identities, agency is presented through the larger social, cultural, and racial contexts that affect historically excluded populations under investigation. For example, Martin (2006) researched African American parents' ability to demonstrate agency and the ways they negated particular racial influences (e.g., dominant white narratives in mathematics education) to maintain positive identity development for their African American children and community.

After tracing the intertextual references of agency cited in the studies within the domain, we noted that Allen and Trinick draw on Barker (2005) to state "concepts of agency involve an individual's capacity to act of their own free will to make autonomous choices" (p. 334). Barker's (2005) book *Cultural Studies: Theory and Practice* outlines several conceptions of agency from multiple sources (e.g., Bandura, Bourdieu, Foucault, Marx), making it difficult to determine which view of agency Allen and Trinick (2021) most rely upon in their study. They, Allen and Trinick (2021), explicitly build on the work of Bourdieu (1986) and the view of structure-agency, or the way groups of people take on and construct specific ways of interacting and being as part of belonging to a specific group. Thus, Allen and Trinick (2021) define agency through the lens of the entire Māori population and the ways the group develops their own free will and autonomous choices in opposition to largely white and western views of mathematics education.

Discussion and Implications

In this systematic review of literature, we used ethnographic perspectives to explore how authors have conceptualized and studied authority, autonomy, and agency in the published literature in the field of mathematics education. In the following section, we outline several points for the field to consider in reflecting upon existing mathematics education research and in moving forward with future studies.

Kinds of Authority, Autonomy, and Agency

One of the major contributions of this study is the illumination of the different domains of authority, autonomy, and agency concepts used within the field of mathematics education. Most studies reviewed for the current study did not clearly articulate their conceptualization of authority, agency, or autonomy, which contributes to a lack of understanding and clarity surrounding these terms and sometimes misuse amongst studies. Through our review of literature, we categorized a multitude of ways or domains that authority, autonomy, and agency have been used and could be used for future studies. Overall, theses thin conceptualization have led to multiple, differing definitions of the same term or even interchangeable, but fuzzy synonyms, without full understanding of the term used or its historical foundations. The current study demonstrates how each term has a rich history with particular denotations from prior work and connotations within the sociocultural contexts the studies take place. Because the field of mathematics education holds so many varying conceptualizations, a single definition for authority, for autonomy, or for agency cannot capture the essence of what each of these terms has come to mean in mathematics education. Therefore, Tables 4a and 4b are presented to capture the various conceptualizations of Authority, Autonomy, and Agency in mathematics education research gleaned from the current study.

Table 4a

Conceptualizations of Authority

Term	*Domain*	*Definition*	*Attributes*
Authority	**Mathematical authority**	Authority as a constant in mathematics education and viewed as something that is delegated, shared, granted or devolved between the teacher and students, as a collective.	Findings focus solely on the ways authority flows from the teacher to the students. Findings point to the importance of including students in the process of learning mathematics. Teachers employ strategies to distribute authority to students within every day mathematical learning.
	Authority structures	Authority as structures present in the mathematics classroom with established rules and norms for determining authority between teachers and students.	These studies use positioning theory (Davies & Harré, 1990) to determine how people in mathematics classes are positioned as *in authority* or as *an authority*. Characterized by the use of Herbel-Eisenmann and Wagner's (2010) authority structures. Studies focus on linguistic cues found within mathematics classrooms and the implications of authority structures between teachers and students.
	Authority relationships	Authority as a socially constructed relationship between people in mathematics classrooms.	Authority is viewed less as a constant but is examined through the different relations that develop among teachers and students as well as among students during collaborative learning endeavors. Findings are focused on how specific classrooms come to share authority, or the ways legitimacy is gained in different constructions of authority.

Table 4b

Conceptualizations of Autonomy, and Agency

Term	*Domain*	*Definition*	*Attributes*
Autonomy	**Autonomy as choice**	Autonomy as giving students choice in their pursuits to do mathematics.	Autonomy through the lens of freedom for the students in the classroom. These studies focused on creating opportunities for students to freely engage in mathematics content, specifically that of whole class discussions. Findings point to a constant thread of teachers giving choice for students to make decisions during mathematical learning opportunities is present.
	Sociomathematical autonomy	Autonomy as being co-constructed through the practices of students and their teacher through mathematical learning opportunities	Autonomy is more than choice but maintains that students must also have freedom to decide and construct what counts as mathematics. Autonomy is co-constructed between students and teachers in classrooms. It is focused on ways students develop autonomous actions and ways teachers can hinder autonomous actions. Yackel and Cobb (1996) can be credited as the first study in this domain.
Agency	**Agency of the self**	Agency through philosophical investigations of the *self* within the mathematics classrooms.	Mathematics as a social institution that is constructed through discourse; thus, constructing the *self* is as much an individual as a discursive process within the larger mathematical culture for learning. Findings show ways students view themselves in mathematics or ways teachers can support agency in their classrooms.
	Agency and racial identities	Agency through the lens of racial identities in mathematics education.	Findings show agency as both an individual endeavor and a collective stance in the face of oppressive mathematics education practices. Findings focus on the ways individuals and groups navigate oppressive educational systems. Studies offer supports to radically reconceptualize mathematics education.

Attending to Groundings of Authority, Autonomy, and Agency

This systematic review has illuminated the complexities inherent in conceptualizing authority, autonomy, and agency in mathematics education. Our findings reveal that while these constructs are widely discussed, their historical, epistemological, and ontological groundings are often insufficiently examined. This lack of sustained engagement with their theoretical roots has resulted in fragmented and occasionally ambiguous conceptualizations. For instance, while Weber's (1947) framework of authority, Piaget's (1948/1973) developmental theories on autonomy, and Bandura's (1986) work on agency are frequently cited, these foundational contributions are often engaged with only at a surface level, without deeper interrogation of their implications for contemporary educational contexts.
Furthermore, understanding how the constructs of authority, autonomy, and agency have been conceptualized in mathematics education research requires attention to the intellectual histories that have shaped the field. Foundational theories such as Weber's (1947) typology of authority, Piaget's (1948) developmental framing of autonomy, and Bandura's (1986) social cognitive theory of agency have provided essential starting points for examining classroom dynamics, learner identity, and participation. These frameworks have long guided efforts to make sense of students' roles and relationships in mathematics learning, offering conceptual language for describing influence, independence, and action within educational settings.

For instance, Weber's (1947) account of traditional authority has framed analyses of hierarchical structures in classrooms, while Piaget's (1948) focus on autonomy as a developmental milestone has supported understandings of individual mathematical reasoning and independence. Bandura's (1986) emphasis on agency as the capacity to act with intentionality has served as a foundation for identifying agentic moments within instruction.

While these frameworks remain influential, as our findings demonstrate, it is equally as important to note, they emerged from specific sociopolitical and historical contexts that differ from contemporary, dialogic, and culturally diverse views of classrooms. As we, as researchers, increasingly attend to the complex, situated, and relational nature of learning, it becomes necessary to explore theoretical perspectives that build upon, and also critically expand upon these early foundations. For example, Weberian accounts of authority may not fully capture the distributed or negotiated power structures that characterize many student-centered or collaborative learning environments (see Edelen et al., 2023; Edelen et al., 2024; Edelen et al., 2025). Similarly, Piagetian (1948) views of autonomy may overlook the ways that students co-construct classroom norms or collectively define what counts as legitimate mathematical reasoning. And individualist framings of agency may obscure how structural, cultural, and institutional constraints shape students' opportunities to act with power, particularly for those from historically marginalized communities.

Other frameworks offer complementary tools for reimagining these constructs. Foucault's (1977) theorization of power as relational and enacted through discourse enables researchers to trace how authority is constantly negotiated through language and interaction. Fairclough (2001) similarly shows how discourse both reflects and reproduces social power, making visible how language might legitimize or constrain different forms of mathematical reasoning. We also note that Bourdieu's (1986) concept of habitus and social fields directs attention to how authority, autonomy, and agency are structured by social positioning and access to cultural capital. These additional perspectives invite us, as mathematics education researchers to foreground the institutional, cultural, and discursive dimensions of classroom life.

In particular, we find distinctive value in ethnographic epistemological approaches to understanding authority, autonomy, and agency. Ethnographic approaches prioritize emic, insider perspectives and illuminate how learners themselves make sense of their roles, relationships, and learning experiences (Skukauskaitė, 2023). Ethnographic orientations offer a generative stance for tracing how children enact and contest power within everyday activity systems, allowing for a layered analysis of how constructs like authority, autonomy, or agency unfold over time, across settings, and within communities (Edelen & Skukauskaitė, 2025; Skukauskaitė & Green, 2023). By placing foundational and contemporary theories in dialogue mathematics education research can better account for the sociocultural, historical, and interactional complexities of classroom life. This kind of theoretical layering supports conceptual clarity and moves to advancing equity by attuning research to the diverse ways students come to exist, participate, and learn in mathematics classrooms for whom and for what purposes.

When theoretical roots are insufficiently explored, it constrains the development of new insights, limiting how these constructs can be operationalized and studied. For example, autonomy and agency have traditionally been framed as individual endeavors, reflecting the developmentalist focus of their origins. Such framings often fail to capture collective dimensions of these constructs, where groups or communities might act autonomously or agentically within mathematics classrooms. Although emerging work, such as Allen and Trinick's (2021) study of collective agency, begins to challenge this individualist paradigm, much remains unexplored about how shared agency and autonomy function in mathematics education.

Engaging deeply with the histories of these constructs is a necessary step in advancing their utility. By revisiting the foundational definitions and examining their evolution, researchers can better understand the assumptions that underlie current studies. For example, authority is frequently conceptualized as a unidirectional flow from teacher to student. This perspective often neglects how authority can be co-constructed or contested within classroom interactions. Similarly, the emphasis on autonomy as choice overlooks the sociomathematical dimensions of autonomy, where students make choices as well as negotiate the very definitions of what counts as mathematics.

The field must also critically evaluate how these constructs intersect. Authority, autonomy, and agency are not isolated phenomena as they are interrelated dimensions of classroom life. For instance, shifts in authority structures, such as when teachers share authority with children, may simultaneously impact how autonomy and agency are experienced and enacted. Understanding these intersections requires researchers to articulate the specific kinds of authority, autonomy, or agency they are studying and to consider how these constructs influence and shape one another.

Future Directions

To advance the study of authority, autonomy, and agency, we propose several critical directions for future research. First, researchers must engage more explicitly with the histories of these constructs, building on or contesting their foundational theories to generate new insights. This includes exploring underexamined dimensions such as collective agency and sociomathematical autonomy, as well as questioning how these constructs operate within evolving pedagogical and sociopolitical contexts. Deepening theoretical engagement across time and traditions will support more nuanced understandings of how power, participation, and identity are structured in mathematics education.

Second, the field must strive for greater conceptual clarity. Our review demonstrates that ambiguous or interchangeable uses of these terms have led to a lack of coherence across studies. Researchers must define these constructs with precision, specifying their attributes, boundaries, and implications for classroom practice. Such clarity strengthens individual studies, which in turn fosters cumulative knowledge-building that can guide both research and reflective practice. While the goal of this review was to clarify the conceptual terrain, the implications of this work extend beyond definitional precision. The three taxonomies developed here offer researchers a foundation for future empirical studies to investigate how these constructs are enacted in classroom teaching, teacher professional learning, and educational leadership. For example, researchers might examine how teachers navigate tensions between authority structures and student autonomy, or how sociomathematical autonomy is fostered through particular instructional practices. Similarly, teacher educators could use these conceptualizations to support preservice and in-service teachers in reflecting on their roles in constructing equitable mathematics learning environments. At the policy level, future research might explore how institutionalized definitions of authority and agency influence curriculum design, teacher evaluation, and broader accountability structures.

Finally, we urge researchers to continue investigating how authority, autonomy, and agency contribute to more equitable learning environments. Future research should include examining how these constructs operate within diverse cultural and social contexts, particularly for historically marginalized groups. By attending to the ways that authority, autonomy, and agency intersect with issues of identity, power, and access, the field can better address the complexities of teaching and learning mathematics.

Limitations

Our study is limited by its focus on mathematics education, a deliberate choice to allow for depth of analysis within a specific field. While this narrow focus provided rich insights into the conceptualizations of authority, autonomy, and agency, it also limits the generalizability of our findings to other disciplines. Future research could extend this work by examining these constructs in broader educational contexts, such as science education or interdisciplinary studies.

Additionally, we note the concentrated influence of a small group of scholars in shaping these constructs within mathematics education. While their work provides a strong foundation, it may also narrow the diversity of perspectives represented in the literature. Additionally, this review was limited to studies published in English, which constrains the cultural and linguistic diversity of perspectives represented in our analysis. Future research should also attend to how authority, autonomy, and agency are conceptualized across global and multilingual educational contexts, where cultural norms, policy, and institutional logistics may shape these constructs in distinct ways. Expanding the field to include voices from global contexts, interdisciplinary approaches, and underrepresented groups, especially childrens' voices, is essential for fostering a more inclusive and comprehensive understanding of these constructs.

Conclusions and Call to Action

As mathematics education continues to grapple with calls for more equitable and inclusive pedagogical practices, the field must respond with precision and intentionality. Authority, autonomy, and agency are powerful constructs with the potential to transform teaching and learning, but their utility depends on how clearly these constructs are defined and operationalized. This review offers a starting point for this critical work, providing taxonomies that map existing knowledge and highlight gaps for future exploration.

Moving forward, researchers must critically engage with the histories, theoretical underpinnings, and intersections of these constructs. By doing so, the field can move beyond surface-level engagement to generate deeper, more meaningful insights. This work advances scholarship by equipping educators with the tools they need to navigate and transform the power dynamics of mathematics classrooms.

Ultimately, the goal is to foster learning environments that equitable and empowering, where children and teachers alike can exercise authority, autonomy, and agency in ways that enrich their mathematical experiences. With greater transparency, clarity, and intentionality, the field of mathematics education can rise to meet this challenge, ensuring that these constructs move beyond theoretical ideals to practical realities in classrooms around the world.

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

Daniel Edelen (dedelen@gsu.edu) is an assistant professor in the Department of Early Childhood and Elementary Education at Georgia State University. His research uses interactional and microethnographic approaches to examine how authority, autonomy, and agency are socially constructed in elementary mathematics classrooms. He collaborates with children to explore how they co-create cultures for learning and participation.

Sarah B. Bush (Sarah.Bush@ucf.edu) is a Professor of K–12 STEM Education and the Lockheed Martin Eminent Scholar Chair at the University of Central Florida. She is the Director of the Lockheed Martin/UCF Mathematics and Science Academy and program coordinator of the Mathematics Education PhD track. Her research focuses on mathematics teacher education, transdisciplinary STE(A)M education, and mathematics and STE(A)M teacher professional learning and leadership.

Karen S. Karp (karen@louisville.edu) is a mathematics educator who focuses on the intersection of mathematics education and special education. She was previously a professor at Johns Hopkins University and a distinguished teaching professor of elementary mathematics education at the University of Louisville in Kentucky where she is professor emerita. She is the author or co-author numerous book chapters, articles, and books,

Audra Skukauskaitė (audra@ucf.edu) is Professor of qualitative research in the department of Learning Sciences and Educational Research at the University of Central Florida. Her research and teaching interests focus on qualitative and ethnographic methodologies, including transparency, quality, creativity, and utilization of qualitative and ethnographic logic across disciplines. She has co-edited volumes on Interactional Ethnography and Engaging Students in Qualitative Research Pedagogies and has published numerous articles on varied aspects of qualitative research, ethnography, and invention education.

Farshid Safi (Farshid.Safi@ucf.edu) is the Associate Director for Teaching and Service, and an Associate Professor of K-12 Mathematics Education at the University of Central Florida. Over the last 20+ years throughout the U.S and Canada, his research has focused on supporting teachers conceptual understanding of elementary and secondary mathematics, connecting mathematical topics through the intentional integration of technology in professional development efforts, as well as supporting and facilitating equitable teaching practices. He is an active contributor to national policy efforts and position statements related to inclusive teaching practices and working with multiple organizations to support the teaching and learning of mathematics. Dr. Safi is currently serving as the President of the Association of Mathematics Teacher Educators (AMTE).

Sherron Killingsworth Roberts (sherron.roberts@ucf.edu) served as the Elementary Education Ph.D. Coordinator and currently serves as the Heintzelman Literature Scholar at UCF. Her research examines literacy as social practice, writing, innovative pedagogy, and children's literature content analyses. Formerly co-editor of Literacy Research and Instruction, she currently serves as Associate Editor for Early Childhood Education Journal and on seven other editorial review boards. She has published in Reading Teacher, Journal of Teacher Education, Journal of Research in Childhood Education, Teaching and Teacher Education, Reading Horizons, and Journal of Reading Education among others.

References

*References marked with an asterisk indicate studies included in the literature review.

Agar, M. (2006). Culture: Can you take it anywhere*? International Journal of Qualitative Methods, 5*(2), 1–12. https://doi.org/10.1177/160940690600500201

Alexander, P. A. (2020). Methodological guidance paper: The art of and science of quality systematic reviews. *Review of Educational Research, 90*(1), 6-23. https://doi.org/10.3102/0034654319854352

*Allen, P., & Trinick, T. (2021). Agency–structure dynamics in an indigenous mathematics education community in times of an existential crisis in education. *Educational Studies in Mathematics, 108*(13-14)**,** 351–368. https://doi.org/10.1007/s10649-021-10098-1

*Amit, M., & Fried, M. (2005). Authority and authority relations in mathematics education: A view from an 8th grade classroom. *Educational Studies in Mathematics, 58*, 145–68. https://doi.org/10.1007/s10649-005-3618-2

*Andersson, A., & Wagner, D. (2019) Identities available in intertwined discourses: Mathematics student interaction. *ZDM: The International Journal on Mathematics Education, 51*(3), 529-540. https://doi.org/10.1007/s11858-019-01036-w

*Atabas, S., Schellinger, J., Whitacre, I., Findley, K., & Hensberry, K. (2020). A tale of two sets of norms: Comparing opportunities for student agency in mathematics lessons with and without interactive simulations. *Journal of Mathematical Behavior, 58*, 1-23. https://doi.org/10.1016/j.jmathb.2020.100761

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Prentice-Hall.

Barker, C. (2005). *Cultural studies: Theory and practice.* Sage.

Baron, S. (2020). *The birth of intertextuality: The riddle of creativity*. Routledge.

*Berry III, R. Q., Thunder, K., & McClain, O. L. (2011). Counter narratives: Examining the mathematics and racial identities of black boys who are successful with school mathematics. *Journal of African American Males in Education, 2*(1), 10-23.

Bloome, D., & Egan-Robertson, A. (1993). The social construction of intertextuality in classroom reading and writing lessons. *Reading Research Quarterly, 28*(4), 305–333. https://doi.org/10.2307/747928

Boaler, J., & Greeno, J. J. (2000). Identity, agency, and knowing in mathematics worlds. In J. Boaler (Ed.), *Multiple perspectives on mathematics teaching and learning* (pp. 171–200). Ablex.

Bourdieu, P. (1986). The forms of capital. In J. Richardson (Ed.), *Handbook of theory and research for the sociology of education* (pp. 241–258). Greenwood Press.

*Brown, R. (2009). Teaching for social justice: Exploring the development of student agency through participation in the literacy practices of a mathematics classroom. *Journal of Mathematics Teacher Education, 12*(3), 171–185. https://doi.org/10.1007/s10857-009-9110-7

Cobb, P., Gresalfi, M., & Hodge, L. L. (2009). An interpretive scheme for analyzing the identities that students develop in mathematics classrooms. *Journal for Research in Mathematics Education, 40*(1), 40–68. https://www.jstor.org/stable/40539320

Cobb, P., & Hodge, L. L. (2002). A relational perspective on issues of cultural diversity and equity as they play out in the mathematics classroom. *Mathematical Thinking and Learning, 4*(2 & 3), 249–284. https://doi.org/10.1207/S15327833MTL04023_7

Cobb, P., Yackel, E., Wood, T., Nicholls, J., Wheatley, G., Trigatti, B., & Perlwitz, M. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*(1), 3–29. https://doi.org/10.2307/749551

Davies, B., & Harré, R. (1990). Positioning: The discursive production of selves. *Journal for the Theory of Social Behaviour, 20*(1), 43–63. https://doi.org/10.1111/j.1468-5914.1990.tb00174.x *de Freitas, E., Esmonde, I., Wagner, D., Knipping, C., Lunney Borden, L. & Reid, D. (2012). Discursive authority and socio-cultural positioning in the mathematics classroom: New directions for teacher professional development. *Canadian Journal for Science, Mathematics and Technology Education*, *12*(2), 137-159. https://doi.org/10.1080/14926156.2012.679994

*Depaepe, F. De Corte, E., & Verschaffel, L. (2012). Who is granted authority in the mathematics classroom? An analysis of the observed and perceived distribution of authority. *Educational Studies, 38*(2), 223-234. https://doi.org/10.1080/03055698.2011.598676

*Dunleavy, T. K. (2015). Delegating mathematical authority as a means to strive toward equity. *Journal of Urban Mathematics Education, 8*(1), 62-82.

Edelen, D., Bush, S. B., & Andreasen, J. (2024). Authority and positionings in elementary mathematics: An interactional ethnographic approach. *Learning, Culture and Social Interaction, 49*, 100866. https://doi.org/10.1016/j.lcsi.2024.100866

Edelen, D., Bush, S. B., Skukauskaitė, A., Karp, K. S., Roberts, S. K., & Safi, F. (2023). The social construction of authorities: An interactional ethnographic examination of positional legitimacy. *Linguistics and Education, 75,* 101177. https://doi.org/10.1016/j.linged.2023.101177

Edelen, D., & Cox, R. (2025). "They" as a discursive marker of curricular authority: Co-constructing chronotopes in early mathematics education. *Investigations in Mathematics Learning*. Advance online publication. https://doi.org/10.1080/19477503.2025.2584925

Edelen, D., & Skukauskaitė, A. (2025). Collective memorying of kindergarten through the logic of children. *Linguistics and Education, 87*, 101409. https://doi.org/10.1016/j.linged.2025.101409

Engle, R. A., Langer-Osuna, J. M., & McKinney de Royston, M. (2014). Toward a model of influence in persuasive discussions: Negotiating quality, authority, privilege, and access within a student led argument. *Journal of the Learning Sciences, 23*(2), 245–268. https://doi.org/10.1080/10508406.2014.883979

Esmonde, I. (2009). Ideas and identities: Supporting equity in cooperative mathematics learning. *Review of Educational Research, 79*(2), 1008-1043. https://doi.org/10.3102/0034654309332562

Fairclough, N. (2001). *Language and power* (2nd ed.). Longman.

Foucault, M. (1977). *Discipline and punish: The birth of the prison* (A. Sheridan, Trans.). Pantheon Books.

*Gerson, H., & Bateman, E. (2010). Authority in an agency-centered, inquiry-based university calculus classroom. *The Journal of Mathematical Behavior, 29*(4), 195–206. https://doi.org/10.1016/j.jmathb.2010.10.003

Green, J. L. (1983). Research on teaching as a linguistic process: A start of the art. *Review of Research in Education, 10*(1), 151-252. https://doi.org/10.3102/0091732X010001151

Green, J. L., Castanheira, M., Skukauskaitė, A., & Hammond, J. (2015). Developing a multifaceted research process: An ethnographic perspective for reading across traditions. In N. Markee (Ed.), *Handbook of classroom discourse and interaction* (pp. 26–43). Wiley Blackwell.

Gresalfi, M. S., & Cobb, P. (2006). Cultivating students' discipline specific dispositions as a critical goal for pedagogy and equity. *Pedagogies: An International Journal, 1*(1), 49-57. https://doi.org/10.1207/s15544818ped0101_8

Gresalfi, M., Martin, T. Hand, V., & Greeno, J. (2009). Constructing competence: An analysis of student participation in the activity systems of mathematics classrooms. *Educational Studies in Mathematics, 70*(1), 49-70. https://doi.org/10.1007/s10649-008-9141-5

Grootenboer, P., & Zevenbergen, R. (2007). Identity and mathematics: Towards a theory of agency in coming to learn mathematics. *Mathematics: Essential Research, Essential Practice, 1,* 335–344.

*Hamm, J. V., & Perry, M. (2002). Learning mathematics in first-grade classrooms: on whose authority? *Journal of Educational Psychology*, *94*(1), 126- 137. https://doi.org/10.1037/0022-0663.94.1.126

*Herbel-Eisenmann, B., & Wagner, D. (2007). A framework for uncovering the way a textbook may position the mathematics learner. *For the Learning of Mathematics, 27*(2). 8-14.

*Herbel-Eisenmann, B., & Wagner, D. (2010). Appraising lexical bundles in mathematics classroom discourse: obligation and choice. *Educational Studies in Mathematics*, *75*(1), 43–63. https://doi.org/10.1007/s10649-010-9240-y

Hull, G. A., & Greeno, J. G. (2006). Identity and agency in nonschool and school worlds. *Counterpoints, 249*, 77–97. https://www.jstor.org/stable/42979590

*Inglis, M., & Mejia-Ramos, J. P. (2009). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction, 27*(1), 25-50. https://doi.org/10.1080/07370000802584513

Kamii, C. (1985). *Young children reinvent arithmetic: Implications of Piaget's theory.* Teachers College Press.

*Kinser-Traut, J. Y., & Turner, E. E. (2020). Shared authority in the mathematics classroom: successes and challenges throughout one teacher's trajectory implementing ambitious practices. *Journal of Mathematics Teacher Education*, *23*(1), 5-34. https://doi.org/10.1007/s10857-018-9410-x

*Kosko, K. W., & Wilkins, J. L. M. (2015). Does time matter in improving mathematical discussions? The influence of mathematical autonomy. *The Journal of Experimental Education, 83*(3), 368-385. https://doi.org/10.1080/00220973.2014.907225

*Lai, Y., & Baldinger, E. E. (2021). Authority relations and the tertiary-to-secondary (dis)continuity. *For the Learning of Mathematics, 41*(2), 26-31.

*Langer-Osuna, J. M. (2016). The social construction of authority among peers and its implications for collaborative mathematics problem solving. *Mathematical Thinking and Learning, 18*(2),107–124. https://doi.org/10.1080/10986065.2016.1148529

Langer-Osuna, J. M. (2017). Authority, identity, and collaborative mathematics. *Journal for Research in Mathematics Education*, *48*(3), 237–247. https://doi.org/10.5951/jresematheduc.48.3.0237

*Langer-Osuna, J. M. (2018). Exploring the central role of student authority relations in collaborative mathematics. *ZDM: The International Journal on Mathematics Education, 50,* 1077-1087. https://doi.org/10.1007/s11858-018-0965-x

*Langer-Osuna, J. M., Chavez, R., Kwon, F., Malamut, J., Gargroetzi, E., Lange, K., & Ramirez, J. (2021). "I'm telling!": Exploring sources of peer authority during a K-2 collaborative mathematics activity. *Studia Paedagogica, 26*(2), 97–111. https://doi.org/10.5817/SP2021-2-5

Langer-Osuna, J. M., & Esmonde, I. (2017). Insights and advances on research on identity in mathematics education. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 637–648). National Council of Teachers of Mathematics.

*Langer-Osuna, J. M., Gargroetzi, E., Munson, J., Chavez, R., Wallace, T. L., & Kuo, E. (2020). Exploring the role of off-task activity on students' collaborative dynamics. *Journal of Educational Psychology*, *112*(3), 514–532. https://doi.org/10.1037/edu0000464

Martin, D. (2000). *Mathematics success and failure among African-American youth.* Lawrence Erlbaum Associates.

*Martin, D. B. (2006). Mathematics learning and participation as racialized forms of experience: African American parents speak on the struggle for mathematics literacy. *Mathematical Thinking and Learning, 8(*3), 197-229. https://doi.org/10.1207/s15327833mtl0803_2

*McClain, K., & Cobb, P. (2001). An analysis of development of sociomathematical norms in one first-grade classroom. *Journal for Research in Mathematics Education*, *32*(3), 236-266. https://www.jstor.org/stable/749827

*McGee, E. O., & Martin, D. B. (2011). "You would not believe what I have to go through to prove my intellectual value!" Stereotype management among academically successful black mathematics and engineering students. *American Educational Research Journal, 48*(6), 1347-1389. https://doi.org/10.3102/0002831211423972

Metz, M. H. (1978). *Classrooms and corridors: The crisis of authority in desegregated secondary schools*. University of California Press.

*Morgan, C. (2016). Studying the role of human agency in school mathematics. *Research in Mathematics Education, 18*(2), 120–141. https://doi.org/10.1080/14794802.2016.1176595

*Moyer, P. S., & Jones, M. G. (2004). Controlling choice: Teachers, students, and manipulatives in mathematics classrooms. *School Science and Mathematics, 104*(1), 16-31. https://doi.org/10.1111/j.1949-8594.2004.tb17978.x

National Council of Teachers of Mathematics. (2018). *Catalyzing change in high school mathematics: Initiating critical conversations.*

National Council of Teachers of Mathematics. (2020a). *Catalyzing change in early childhood and elementary mathematics: Initiating critical conversations.*

National Council of Teachers of Mathematics. (2020b). *Catalyzing change in middle school mathematics: Initiating critical conversations.*

*Ng, O. L., Cheng, W. K., Ni, Y., & Shi, L. (2021). How linguistic features and patterns of discourse moves influence authority structures in the mathematics classroom. *Journal of Mathematics Teacher Education*, *24*(6), 587–612. https://doi.org/10.1007/s10857-020-09475-z

Pace, J. L., & Hemmings, A. (2007). Understanding authority in classrooms: A review of theory, ideology, and research. *Review of Educational Research, 77*, 4–27. https://doi.org/10.3102/003465430298489

Piaget, J. (1948/1973). *To understand is to invent: The future of education*. Grossman Publishers.

Pickering, A. (1995). *The mangle of practice: Time, agency, and science.* University of Chicago Press.

Skemp, R. (1979). *Intelligence, learning and action*. Wiley.

Skukauskaitė, A. (2023). Ethnography: Foundations, challenges, and spaces of possibilities. In R. J. Tierney, F. Rizvi, & K. Ercikan (Eds.), *International encyclopedia of education* (4th ed., pp. 92–101). Elsevier.

Skukauskaitė, A., & Green, J. (2023). Ethnographic spaces of possibility: Interactional ethnography in focus. In Skukauskaitė, A., & Green, J. (Eds.), *Interactional ethnography: Designing and conducting discourse-based ethnographic research* (pp. 1-19). Routledge.

*Solomon, Y., Hough, S., & Gough, S. (2021). The role of appropriation in guided reinvention: establishing and preserving devolved authority with low-attaining students. *Educational Studies in Mathematics*, *106*(2), 171–188. https://doi.org/10.1007/s10649-020-09998-5

Spradley, J. P. (2016). *Participant observation.* Waveland Press. (Original work published 1979).

*Tatsis, K., Wagner, D., & Maj-Tatsis, B. (2018). Authority and politeness: Conflict and alignment in mathematics group work. *ZDM: The International Journal of Mathematics Education, 50* (8), 1029-39. https://doi.org/10.1007/s11858-018-0990-9

*Wagner, D. (2007). Students' critical awareness of voice and agency in mathematics classroom discourse. *Mathematical Thinking and Learning, 9*(1), 31–50. https://doi.org/10.1080/10986060709336604

*Wagner, D., & Herbel-Eisenmann, B. (2014a). Identifying authority structures in mathematics classroom discourse: A case of a teacher's early experience in a new context. *ZDM—The International Journal on Mathematics Education, 46*(6), 871–882. https://doi.org/10.1007/s11858-014-0587-x

*Wagner, D., & Herbel-Eisenmann, B. (2014b). Mathematics teachers' representations of authority. *Journal of Mathematics Teacher Education*, *17*(3), 201–225. https://doi.org/10.1007/s10857-013-9252-5

Wagner, D., & Herbel-Eisenmann, B. (2018). A discourse-based framework for identifying authority structures in mathematics classrooms. In U. Gellert, C. Knipping, & H. Straehler-Pohl (Eds.), *Inside the mathematics class: Sociological perspectives on participation, inclusion, and enhancement* (pp. 291–313). Springer

*Warfield, J., Wood, T., & Lehman, J. D. (2005). Autonomy, beliefs and the learning of elementary mathematics teachers. *Teaching and Teacher Education, 21*(4), 439–456. https://doi.org/10.1016/j.tate.2005.01.011

Weber, M. (1947). *The theory of social and economic organization* (A. M. Henderson & T. Parsons, Trans.). Free Press. (Original work published 1925).

*Wilson, M. R., & Lloyd, G. M. (2000). Sharing mathematical authority with students: the challenge for high school teachers. *Journal of Curriculum & Supervision*, *15*(2), 146–169.

*Wood, M. (2013). Mathematical micro-identities: Moment-to-moment positioning and learning in a fourth-grade classroom. *Journal for Research in Mathematics Education*, *44*(5), 775-808. https://www.jstor.org/stable/10.5951/jresematheduc.44.5.0775

*Wood, M. (2016). Rituals and right answers: Barriers and supports to autonomous activity. *Educational Studies in Mathematics, 91*(3), 327-348. https://doi.org/10.1007/s10649-015-9653-8

*Yackel, E., & Cobb, P. (1996). Sociomathematical norms, argumentation, and autonomy in mathematics. *Journal for Research in Mathematics Education, 27*(4), 458–47. https://doi.org/10.5951/jresematheduc.27.4.0458

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 29, NO. 4, 29-62

Keeping Records While Solving Problems: A Study of Multilingual Learner and First-Language English Speaking Students

Daniel J. Heck
Horizon Research, Inc.

Anthony Fernandes
UNC Charlotte

Johannah Nikula
Education Development Center

Evelyn M. Gordon
Horizon Research, Inc.

ABSTRACT

Creating written records while working on mathematics tasks may help students make sense of tasks and free cognitive resources for reasoning as they offload elements of the problem-solving process to paper. We investigated the extent of cognitive processes of multilingual learner (ML) and first-language English speaking (non-ML) students' record keeping (RK) on tasks designed with and without supports for RK and the association between evidence of students' cognitive process in RK (EC-RK) and the correctness of their solutions. Grades 7-9 (aged 12 to 16) students worked on RK-Supported or RK-Unsupported versions of three tasks, and we rubric-scored their solutions for both EC-RK and correctness. Overall, higher EC-RK scores were associated with greater correctness, confirming the utility of EC-RK for solving mathematics tasks. The presence of supports, though, did not increase the extent to which students' RK reflects their cognitive processes, yet correctness of ML students' solutions was associated with solving the RK-Supported versions of the tasks. This result suggests benefits of these supports for ML students apart from encouraging EC-RK.

Keywords: student record keeping, geometry and measurement; problem solving; multilingual students; task design; cognitive load

Introduction

Motivation and Research Questions

Engaging in problem solving is an essential part of mathematical learning (Lindquist et al., 2017; National Council of Teachers of Mathematics [NCTM], 2000, 2014). Understanding what fosters successful engagement in problem solving is vital for supporting students. Previous studies provide evidence that keeping records in various forms supports successful engagement in mathematical problem solving (Murata, 2008; Stylianou & Silver, 2004). Accordingly, the Mathematical Record Keeping Supports Cognition and Communication study investigated the role that Grades 7 to 9 (aged 12 to 16) students' record keeping (RK) plays during their mathematical problem solving. We sought to understand how task design can support successful record keeping, guided by foundational ideas about problem solving (Polya, 1957; Schoenfeld, 1980, 1992) and Cognitive Load Theory (Sweller, 1988, 1994, 2003). Students who are multilingual learners (MLs) are of special interest in the study because the increased cognitive load they face with language demands (Barbu & Beal, 2010; Campbell et al., 2007) suggests that they may be particularly poised to benefit from RK.

Supporting Students' Problem Solving

The importance of problem solving is emphasized in standards that guide mathematics teaching, learning, and assessment in the United States and internationally. Examples include the National Council of Teachers of Mathematics' (NCTM) problem-solving process standards and effective teaching practices (NCTM, 2000, 2014), the applying and reasoning domains of the *TIMSS 2019 Assessment Frameworks (Lindquist et al., 2017),* and the *Standards for Mathematical Practice (SMP)—including SMP1: "*Make sense of problems and persevere in solving them"—articulated in the *Common Core State Standards for Mathematics* (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). These policies stem from decades of research on the centrality of problem solving in mathematics teaching and learning (e.g., Liljedahl et al. 2016, Nunokawa, 2005; Schoenfeld, 1980, 2014). Both policy and research support an emphasis on learning and using strategies for solving mathematical problems as essential elements of a strong and successful mathematics education for all students.

Record Keeping

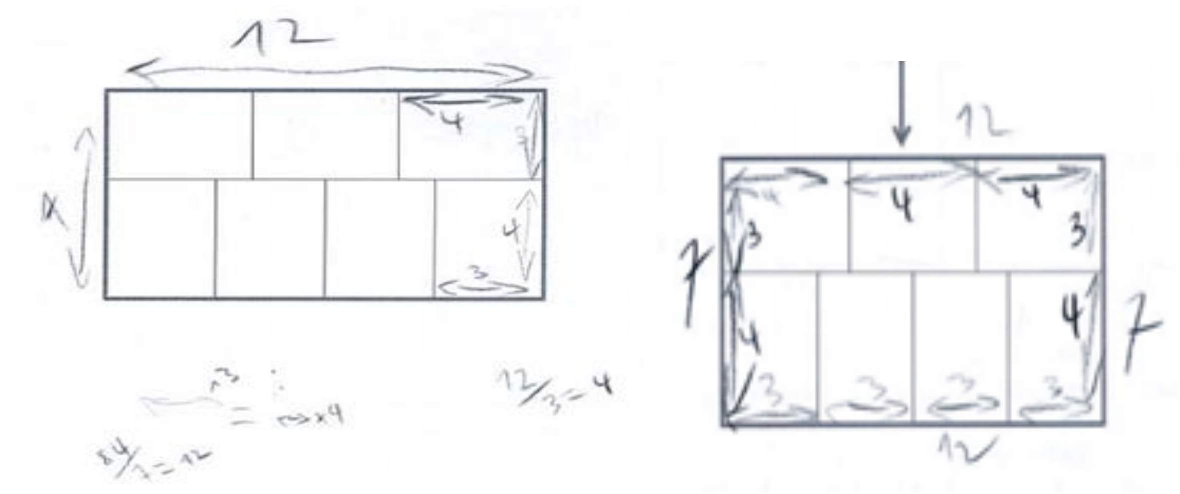
Previous research points to numerous cognitive processes that affect students' problem-solving performance, including executive function (Swanson, 2011) and working memory (e.g., LeBlanc & Weber-Russell, 1996; Ng & Lee, 2009; Paas & van Merriënboer, 2020). Other research (e.g., Cellucci, 2019; De Toffoli, 2018; Murata, 2008; Stylianou & Silver, 2004; Sunzuma et al., 2020) suggests that RK can support students' effective management of these processes and use of cognitive resources by allowing them to offload some demands of problem solving into external records.

For this study, RK refers to the act of capturing pieces of information on paper (or electronically) during work on a mathematical problem. Records may include words, quantities, symbols, and equations (e.g., Rexigel et al., 2024), as well as drawings and diagrams (e.g., Stylianou & Silver, 2004). Problem solvers' records may serve a range of purposes: highlighting information that is provided in the problem or ideas and predictions related to solving the problem, creating various representations, and documenting problem-solving steps or partial solutions (Gordon et al., 2015). Problem solvers can act on the information captured in records or retrieve it later, as needed, without having to rely on memory (Paas & van Merriënboer, 2020). The act of creating records can therefore support mental focus on other aspects of problem solving (Gordon et al., 2015).

Consider, for example, the work shown in Figure 1 that a student produced to determine the perimeter of the large rectangle, given that its total area is 84 square units and it is composed of seven congruent smaller rectangles arranged as shown. The records that this Grade 8 student (a ML) wrote and drew on two different copies of the given picture while working on the task eventually led them to correctly find the perimeter of the large rectangle and the side lengths of the small rectangles.

Figure 1

Student Work on the Seven Rectangles Task



Using two copies of the diagram allowed the student to offload calculations and information about relationships between side lengths, and this in turn afforded them the opportunity to attend more fully to the problem-solving process and to note the needed connections to correctly solve the problem. The student labeled some of the side lengths with numerals, tracking the relationship between the parts of the geometric figure. The arrows accounted for side lengths of the smaller rectangles constituting the sides lengths of the larger rectangle.

Mathematics education literature offers diverse terminology which intersects with parts of our definition of RK, including literature examining problem solvers' creation of diagrams (e.g., Cellucci, 2019; Diezmann & English, 2001; Murata, 2008; Nunokawa, 1994; Purchase, 2014; Sunzuma et al., 2020; Willis & Fuson, 1988), external representations (Zhang, 1997), and inscriptions (Moschkovich, 2008). Some of these terms are laden with meanings that we do not intend. For example, diagrams and representations normally refer to records with some mathematical meaning that is relevant to students' conceptual development of ideas. The scope of records in which we are interested includes conceptually meaningful records, although we argue that less mathematically substantial records can also further students' problem solving and are, therefore, worthy of study (Fernandes, et al., 2015; Neumayer-DePiper, et al., 2015). For example, simply placing dots or hash marks on a diagram to keep track of parts that have already been counted or managed can allow a student to focus on other information needed to solve a problem. In addition, such RK can be a precursor to more meaningful RK. We have observed students returning to and changing records after thinking about another aspect of a task. For example, students have replaced dots that initially signaled that they had accounted for parts with numbers that supported them in enumerating or totaling measures of those parts.

Investigating Supports for Record Keeping

Prior to the study reported here, we had investigated features of mathematics task presentation intended to promote RK due to the evidence that RK plays a role in problem solving. For the current study, we used these features to design and modify mathematics tasks to create two parallel versions, one that incorporated features intended to support RK and one that did not include these features. The aim of the current study was to better understand students' success in problem solving when working with tasks specifically designed to support RK. The study's overarching research questions were: (1) What is the relationship between students' use of record keeping and performance on tasks? (2) How does student performance differ on tasks designed with supports for record keeping versus tasks without these supports? Given the role that visual representations and other records may play in the mathematics classroom in supporting multilingual learners, the third research question was: (3) What differences are evident in the impacts of students' record keeping, and task-embedded record-keeping supports, for multilingual learners compared to first-language English speaking learners?

Theoretical Framework

Record Keeping and Cognition

Our theoretical framework is guided by Cognitive Load Theory (CLT), a learning and instructional theory based on the temporary and limited nature of working memory and the comparative permanence and unlimited capacity of long-term memory (Sweller, 1988, 1994, 2003). Working memory draws on long-term memory but can only store about seven chunks of information and process only two or three chunks of information at a time. If these limits are exceeded, working memory becomes overloaded. CLT considers three types of cognitive load a learner needs to manage for successful learning and performance: *intrinsic load, extraneous load,* and *germane load* (Paas & van Merriënboer, 2020; Renkl & Atkinson, 2003; Sweller et al., 1998). *Intrinsic load* is the cognitive load generated by the nature of the problem and the elements that need to be considered in working memory simultaneously to understand the problem. A problem solver who has already learned what is needed to solve a problem deals mainly with intrinsic load. *Extraneous load* is generated by processes that are not necessary for performance, which may include distractions, anxiety, or expectations for organization or presentation that do not aid learning. Finally, *germane load* is the cognitive load generated in the process of learning. Germane load is particularly relevant for problem solvers who are developing their understanding of the ideas needed to solve the problem. During problem solving, students need to effectively manage the intrinsic and extraneous load within a problem to progress towards a solution. RK may help students focus on intrinsic load and ignore extraneous load, in part by offloading their thinking (i.e., onto paper) to free up working memory to manage germane load.

Researchers working with students from early elementary school through college have found that successful problem solvers are able to develop representations of problems rather than working directly with the text as given (De Corte et al., 1985; Diezmann & English, 2001; Fischbein, 1977; Larkin et al., 1980; Nunokawa 1994; Rexigel et al., 2024) and that experts construct many more visual representations than novices do during problem solving (Bodner & Domin, 2000; Stylianou & Silver, 2004). Developing a representation not only records information about a problem for storage and retrieval but also shapes the problem-solver's thinking (Chu et al., 2017; Meira, 1995). An effective representation makes evident important relationships and constraints in a problem, allowing the solver to determine actions that will lead to a solution to the problem (De Toffoli, 2018). Therefore, appropriately capturing the structure of the problem can be a key step in determining a solution (Bodner & Domin, 2000; Diezmann & English, 2001; Sunzuma et al., 2020), and doing so may involve multiple steps of RK (Nunokawa, 1994). The first diagram that a student draws may not capture the inherent structure of the problem; instead, it represents the elements that the student immediately notices in the situation and the relationships among those elements. Such a step manages some of the intrinsic and extraneous load. As the student continues interacting with the problem, they may modify initial records to capture the inherent structure of the problem, enabling their working memory to focus on the germane load.

RK's Potential as Support for Multilingual Learners

Current research on students who are multilingual learners (MLs) emphasizes the importance of translanguaging, meaning students' use of their full linguistic repertoire to engage in communication and meaning-making (Garza & Arreguín-Anderson, 2018; Grapin et al., 2025). Translanguaging acknowledges that students do not compartmentalize their languages in rigid ways; instead, they fluidly navigate between them to communicate and construct meaning (Elshafie & Zhang, 2024; García, 2023; García & Solorza, 2021). An expansive view of translanguaging highlights that students' repertoire can also encompass non-verbal modes, such as gestures, drawings, or manipulating concrete materials (González-Howard et al., 2023). MLs, who are learning both content and language simultaneously, often face challenges with the language of a problem (Abedi & Lord, 2001; Martiniello, 2008). The simultaneous demands of both challenging content and complex language can lead to cognitive overload (Campbell et al., 2007). In terms of mathematical learning, this overload is largely driven by the extraneous cognitive load imposed by language. For MLs, expansive translanguaging, which includes resources like RK, is a critical asset for addressing language challenges and more fully engaging with mathematical problem solving.

Any student can reduce extraneous load generated in a problem statement by using RK to isolate the mathematical characteristics of the problem, for example by marking up the problem statement, making notes, or creating a diagram. For MLs, these tools are particularly effective when paired with translanguaging strategies, allowing students to describe, question, and analyze mathematical concepts using all of their linguistic resources. By using diagrams, for instance, students can bolster the capacity of working memory by offloading part of their thinking onto the environment (Paas & van Merriënboer, 2020; Tabachneck-Schijf et al., 1997) in order to access tasks in ways that emphasize patterns, relationships, and spatial reasoning, fostering a deeper understanding of mathematical structures (Echevarría et al., 2017). When these external records are created, the student can focus working memory on a few key quantities and relationships at a time as they progress in solving the problem (Paas & van Merriënboer, 2020; Zhang, 1997). Prior research into the use of nonverbal resources, such as drawings, gestures, and manipulation of concrete objects, along with writing or speech, suggests that such resources provide further opportunities for MLs (and others) to develop proficiency with mathematics and mathematical language (Driscoll et al., 2012; Fernandes et al., 2017; Fernandes & McLeman, 2012; Moschkovich, 1999, 2002, 2010; Paas & van Merriënboer, 2020).

Methods

Participants

Fifty-six students participated in this study. Students were identified as MLs (n=20) or non-MLs (n=36) based on their teachers' reports of current receipt of their school's ESL services. Each student self-reported their gender, age, grade, and current mathematics class. (Students also responded to survey questions about their current and previous participation in English as a Second Language instruction, but anomalies in the data suggested that a number of students misinterpreted our intent in these questions.)

Table 1

Participant Information

	Grade 7 (16)		Grade 8 (27)		Grade 9 (13)		All Students (56)	
Math Class	ML	Non-ML	ML	Non-ML	ML	Non-ML	ML	Non-ML
Accelerated	1	1	0	3	0	0	1	4
Regular	7	7	6	12	0	0	13	19
Remedial	0	0	1	0	0	0	1	0
Algebra I	0	0	0	1	3	2	3	3
Algebra II	0	0	0	0	0	5	0	5
Other	0	0	2	2	0	3	2	5
Gender								
Female	6	3	2	9	2	5	10	17
Male	2	5	7	9	1	5	10	19
Total	8	8	9	18	3	10	20	36

Table 1 displays information about the sample of students. The distribution of females and males was about the same for the participating ML and non-ML students, except that the participating MLs in Grade 7 were disproportionately female while the participating MLs in Grade 8 were disproportionately male. The sample included 27 females and 29 males aged 12 to 16, of whom 16 were in Grade 7, 27 in Grade 8, and 13 in Grade 9. Twenty students (36%) were identified by their teachers as current ML students, while the remaining 36 students (64%) were not designated as MLs at the time of participation.

Data Collection Instruments

The instruments used in this study were developed and refined in earlier phases of work. We first selected 11 geometry and measurement tasks that could be completed in 10 to 15 minutes, had multiple entry points or solution strategies, were of high cognitive demand, and offered opportunities to use RK for conceptualizing and solving the task. We revised the tasks to make them clearer, removed unnecessary language that might be difficult for MLs, and provided space and/or prompts for student RK. During cycles of administering, analyzing, and revising we had students solve tasks and interviewed them about their work, reviewed the written work and video-recordings of the sessions, and revised the tasks for subsequent rounds of administration and interviews. Thirty-six pilot students participated during this phase (10 MLs, 26 non-MLs). Most completed 3 tasks, resulting in 96 task interviews. We analyzed the dataset of written work and interviews to identify task features that supported students' RK (Heck et al., 2015).

We then developed RK-Supported (RK-S) and RK-Unsupported (RK-U) versions for five of these eleven tasks that elicited a variety of RK used in solutions and for two additional tasks similar to those five. The RK-S versions included several features we had identified as supporting students' RK, such as including an early prompt to write or draw something, having extra copies of diagrams, formatting the task to include space for making records, designating specific answer spaces, and providing an active audience or "real world" context for the solution. The RK-U versions were designed to present the same task with the same cognitive load, but without the features to support RK. These tasks were again refined through an iterative process of interviewing students (21 students, 8 MLs, 13 non-MLs for 74 total task interviews) about their work on the tasks, making modifications to the tasks, and testing the modifications in further interviews. We also solicited feedback from three mathematics educators who reviewed the RK-S and RK-U versions of these seven tasks, specifically to judge the comparability of cognitive demands of the mathematics and the language in the two versions, and recommended ways to improve comparability.

This process led to selection of three tasks for the current study: Floor Plan, Painted Shapes, and Seven Rectangles (see Appendix A). We selected these tasks because they had RK-S and RK-U versions that appeared to provide differing levels of support for RK without altering the cognitive demands of the task. The tasks were accessible to many students, meaning that most of the students interviewed were able to make some progress even if they were not able to complete the task. At the same time, the tasks were complex enough that students were not able to complete them mentally. Multiple strategies could be used to successfully solve each task, and for one task (the Floor Plan task) there are many different correct solutions. In addition, we selected tasks that address different geometry and measurement standards and use different skills and knowledge.

Along with the three tasks, task booklets included a background survey for students to self-report age, gender, and mathematics class[1] (see Appendix A). We varied the order of the three tasks and the task versions (RK-S versus RK-U) across 12 forms of the booklet (Forms A-H) to ensure an even distribution of data among the RK-S and RK-U versions of the three tasks. The order of RK-S and RK-U tasks within each booklet was chosen to accommodate two goals. First, we wanted to delay students' exposure to the RK supports because we believed students' RK for a subsequent task could be affected by exposure to RK supports in a first task. Therefore, every task booklet started with a RK-U version of one task followed by a RK-S version of a second task. We also varied whether the third task was RK-S or RK-U. The 12 forms of the task booklet were randomly assigned to students, blocking by grade level and by ML status to ensure comparable distribution on these factors.

Data Collection and Preparation Process

From February 2016 through June 2016, three researchers, including authors 2 and 3, collected data from 56 students in two school districts in Massachusetts and one in North Carolina. Data collection took place in students' schools via one-on-one sessions between a researcher and a student.

At the beginning of the one-hour session with each student, we gave the student a task booklet and followed a script to instruct students about how to work in the booklet. Students could clarify any word's meaning at this time. They were instructed to work in order and let us know before they moved to the next task. Students could use colored pencils/pens and were asked not to erase any work (they could cross out work). The students were moved to the next task after 15 minutes to ensure that they worked on all three tasks. When students started the second task, which was designed to support RK, we made additional scrap paper available and pointed out extra copies of figures provided with the task. When students started the third task, if it was an RK-U task version we collected the extra paper and set it aside. If the third task was an RK-S task, we again pointed out the additional supplies, including any extra copies of figures that were associated with that task.

During data collection, we used two video cameras to capture each student's working process. One camera was focused on the task booklet to capture the student's RK, and the other camera was positioned to record the student as they worked on the task. In addition to collecting this video footage, we documented the student's work using a researcher note-taking version of the task booklet. Our intention was to document as thoroughly as possible the student's use of RK on each task so that the note-taking booklets, when combined with the student's actual work booklet, could serve as the primary artifacts for analysis. The video recording of the session served as additional evidence for any instances in which it was unclear when or how a student made and used particular records.

[1] Students were also asked to self-report race/ethnicity and present or past engagement with English as a Second Language services at school. Numerous anomalies in these data suggested that students' responses were likely not valid for reporting.

Scoring Rubrics and Scoring Process

We analyzed each task that students completed using two different rubrics – one focused on the extent to which student's RK on the task provided evidence of the student's cognitive processes for solving the task, regardless of whether they ultimately achieved a correct response. Scores on this rubric indicated the extent to which students' RK as a whole provided evidence of their cognitive processes while solving the task; such evidence might be found in individual records (e.g., the placement of an auxiliary line) or the evidence might appear in connections among records (e.g., counting dots connected to quantities they measure). The second rubric focused on the correctness of the student response to that task (see Appendix A). These rubrics went through several rounds of revision, each informed by members of the research team applying the rubrics to student work products and discussing the scoring. The Evidence of Cognitive Processes in Record Keeping rubric was the same for all tasks, although specific anchoring examples were provided in relation to the three different tasks. The Correctness rubric was specific to each of the three tasks. Possible scores on each rubric were 0, 1, 2, 3, or 4, with a higher score indicating greater evidence of cognitive processes in RK (EC-RK) or a more complete and correct response (Correctness).

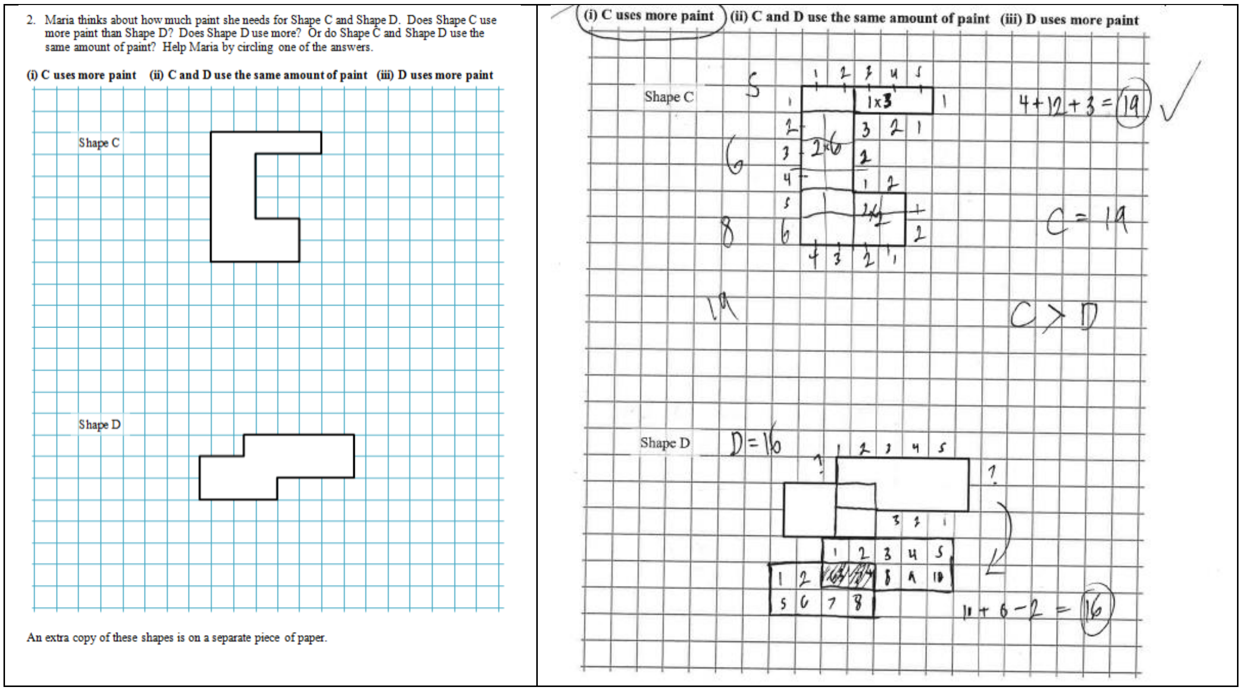
For example, the work shown in Figure 2 on the Painted Shapes task by a Grade 9 student (a non-ML) received a rating of 4 for EC-RK and 3 for Correctness. The rating of 4 for EC-RK indicates that:

- the RK provided evidence for how the student conceptualized and worked through the task, because it shows the decomposition of the figures and means of counting square units;
- the RK appeared to have a problem-solving purpose, because it documents how total areas were determined; and
- connections could be identified among individual instances of RK, because the decompositions of the figures and the enumeration of square units corresponded to the equations used to find total areas.

Although the student correctly answered that Shape C would require more paint, they did not find the correct area for both figures. The Correctness rating of 3 accordingly indicates that the student's work was mostly correct, with the incorrect area for Shape D apparently the result of a minor error in translating the figure, including an extra half square unit in the bottom row of the figure, rather than evidence of a conceptual misunderstanding.

Figure 2

Student work on the Painted Shapes Task (RK-S)



Three members of the research team, including authors 1 and 4, trained to use the rubrics before scoring the tasks. First, we reviewed a set of responses and discussed them together in relation to each rubric, making edits to the rubrics until a consensus was reached. Next, the scorers independently scored a small set of responses, and then we discussed and resolved discrepancies in the scores, leading to further editing and additional examples provided on the rubrics to improve consistency. Finally, two members of the research team independently scored each of the task responses. The two scorers discussed all discrepancies to come to a resolution, sending responses to a third scorer if they could not resolve a discrepancy through discussion. To improve consistency, we scored and reconciled all responses to one task at a time. Initial inter-rater reliability was good for Correctness (independent scores were the same for 83 percent of responses) and marginal for EC-RK (54 percent of responses were assigned the same score independently). For each of the rubrics, over 97 percent of the independent scores were within 1 point of each other across the set of responses, and scorers resolved over 99 percent of the discrepancies through discussion.

Results

Analysis

After rubric scores were determined, we performed a series of within- and across-student quantitative analyses. We first examined the relationship between students' EC-RK and correctness of solutions. We then compared students' work on tasks with and without RK supports and examined whether the impact of RK supports was different for ML and non-ML students.

We employed two overarching multi-level models, one with EC-RK as an outcome variable and the other with Correctness as the outcome variable. We built the models progressively to examine our two factors of interest—tasks designed with and without RK supports, and students' EC-RK in problem solving—first separately and then in combination. For both factors, we also investigated differences between ML and non-ML students in interaction with these two factors. For each model, we nested the three tasks that students completed (level 1) within each student (level 2). To account for differences among the tasks and students' grade levels, we accounted for task type and student grade in all models. The variables included in each model, one progression with EC-RK as the outcome variable and the second progression with Correctness as the outcome variable, are outlined in Tables 2 and 3, respectively (equations for the models can be found in Appendix B).

Table 2

Analytic Models for EC-RK Outcome

	Level 1 (Task)		Level 2 (Student)		Interaction
Model	Task type	RK Support	Grade	ML status	ML status x RK Support
RK-0	Y		Y	Y	
RK-1	Y	Y	Y	Y	
RK-2	Y	Y	Y	Y	Y

Table 3

Analytic Models for Correctness Outcome

	Level 1 (Task)		Level 2 (Student)			Interactions	
Model	Task type	RK Support	Grade	ML status	EC-RK	ML status x RK Support	ML status x EC-RK
C-0	Y		Y	Y			
C-1	Y		Y	Y	Y		
C-2	Y		Y	Y	Y		Y
C-3	Y	Y	Y	Y			
C-4	Y	Y	Y	Y		Y	
C-5	Y	Y	Y	Y	Y		
C-6	Y	Y	Y	Y	Y	Y	Y

Note: The task type dummy variables at level 1 excluded the Painted Shapes task. At level 2, Grade 7 and non-ML students were the excluded categories. All models used grand mean centering for all variables to aid in the interpretation of coefficients.

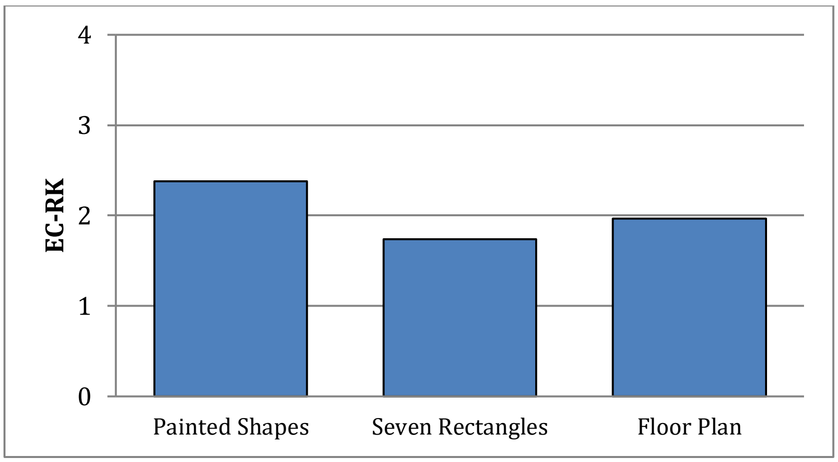
Results for Use of EC-RK in Responses

Our original intent in creating the RK supports was to encourage students to make records as a part of their solving processes. We designed the RK supports in pilot studies to identify and test features that students indicated either encouraged or discouraged their creating of records (Fernandes et al., 2015; Heck et al., 2015). For this study, our purpose was to examine whether the inclusion of RK supports in the task design was associated not with the presence or quantity of records, but more pointedly with the extent to which records reflected students' cognitive processes, that is to say, with students' EC-RK. Model RK-0 included variables for task type at level 1 and students' grade level and

ML status at level 2, establishing the foundation for this set of analyses. EC-RK scores did not differ by students' grade level or ML status. Evidence of cognition in RK was significantly lower for both the Seven Rectangles (t (df) = -4.24 (110), $p < .05$) and Floor Plan (t (df) = -2.71 (110), $p < .05$) tasks compared to the Painted Shapes task. Figure 3 shows the expected scores in Model RK-0 for each of the tasks (full results in Appendix B). Inclusion of the task type variables accounted for these differences in all further analyses.

Figure 3

Model RK-0 Expected Scores for EC-RK, by Task



In Model RK-1, we examined whether inclusion of RK supports in the task design had an effect on the level of students' EC-RK. No significant association was detected. Finally, in Model RK-2, the effect of including RK supports was considered in interaction with students' ML status. Here again, no significant association was found. Neither model resulted in an appreciable reduction in variance at either the task or student levels (full results in Appendix B).

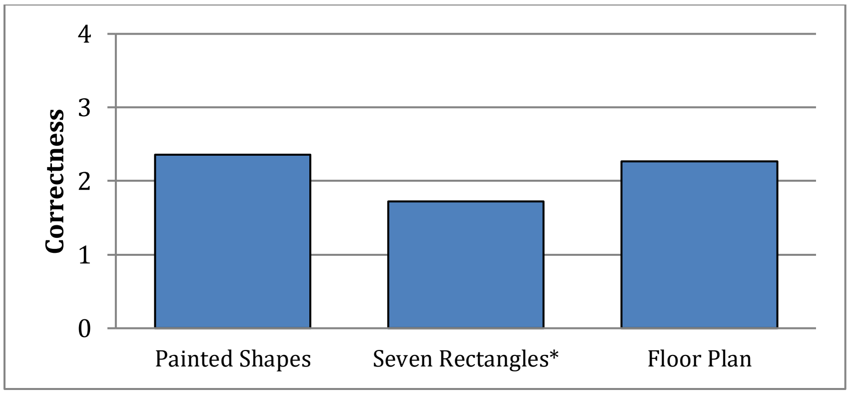
Results for Correctness of Responses

To investigate the impacts of RK supports on students' problem solving, and particularly on ML students' problem solving, we proceeded in three stages. First, we examined whether students' EC-RK was associated with greater progress toward a correct response, both as a main effect and in interaction with ML status. Next, we examined whether the inclusion of RK supports in the task design was associated with correctness, regardless of observed EC-RK, overall and by ML status. Finally, we examined the combined effect of students' EC-RK and the inclusion of RK supports in the task design, again as both a main effect and in interaction with ML status.

We began with a model (C-0) that included task type, grade level, and ML status to provide a foundation against which other models could be compared. The expected scores for Model C-0 shown in Figure 4 indicate that scores on the Seven Rectangles task were lower, on the whole, than scores on the other two tasks (t (df) = -3.95 (110), $p < .05$). Neither grade level nor ML status were significant predictors of correctness scores. The remaining variance in model C-0 was 74 percent for level 1 and 86 percent for level 2 (full results in Appendix B). The three tasks were not designed to be of equal difficulty, so overall differences in correctness scores by task was acceptable. Including task type in all models accounted for these differences analytically. The results indicating no overall differences by grade level or ML status suggested that, on average, the collection of tasks did not favor students according to these factors, which was the intended result of the task selection, review, and design work completed in the early phases of the study.

Figure 4

Model C-0 Expected Scores for Correctness, by Task



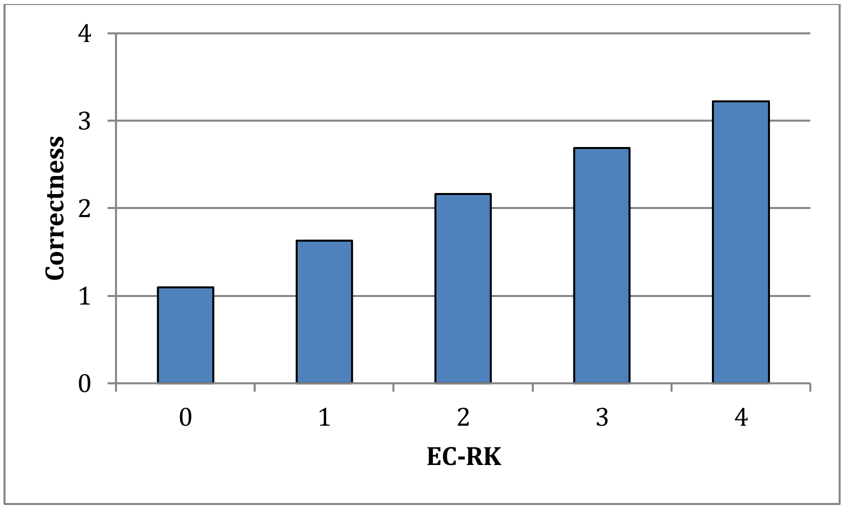
Note: Scores on Seven Rectangles lower than scores on Painted Shapes and Floor Plan

(t (df) = 3.95 (110), $p < .05$).

Models C-1 and C-2 analyzed the effect of students' EC-RK during problem solving, first as a main effect only, and then in interaction with ML status. As illustrated in Figure 5, the results for Model C-1 indicated a strong, positive effect of EC-RK on correctness of the responses (t (df) = 6.98 (109), $p < .05$). A one point gain on the EC-RK rubric was associated with slightly more than a half point gain on the correctness rubric. It is interesting to note that accounting for EC-RK eliminated the significant difference in scores between the Seven Rectangles task and the other tasks. Including the EC-RK predictor variable reduced variance at both levels of the model compared to Model C-0. In Model C-1, the reduction was a modest 8% of variance at the task level (from 0.74 to 0.68), indicating some differences in the effect of students' EC-RK on correctness across the three tasks. At the student level, there was a substantial reduction in variance of 52% (from 0.86 to 0.41), suggesting that differences in EC-RK across students have a considerable effect on correctness. Model C-2 added an interaction effect between ML status and EC-RK which was non-significant, indicating there was no detectable difference in the effect of EC-RK on correctness between ML and non-ML students. Accordingly, no additional reduction in model variance was evident.

Figure 5

Model C-1 Expected Scores for Correctness, by EC-RK Score

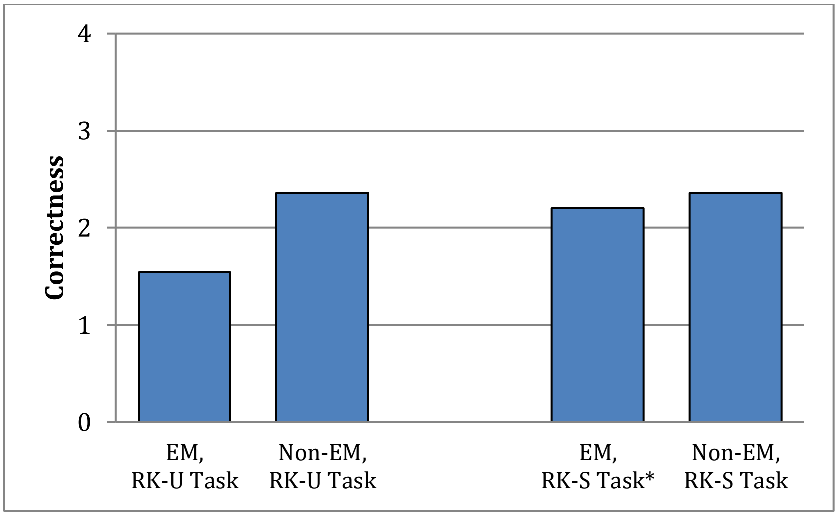


Note: Expected Scores are adjusted for Task, Grade Level, and ML status.

Models C-3 and C-4 analyzed the effect of including RK supports in the task design, first as a main effect and then in interaction with ML status. According to the results of Model C-3, the inclusion of RK supports made no overall difference in correctness scores—the coefficient for RK Supports was 0 and coefficients for all other variables, along with the intercept, were essentially the same as in Model C-0. However, the results of Model C-4 reveal an important difference among students. The interaction between RK supports in task design and ML status was significant and positive (t (df) = 2.27 (108), $p < .05$). That is, ML students' correctness scores were higher on task versions that provided RK supports than on the versions that did not. This result is evident in a positive coefficient for this interaction and a negative coefficient for ML status (t (df) = -2.49 (52), $p < .05$). For these two models, remaining variance at both levels remained essentially unchanged from Model C-0; in addition, correctness scores on Seven Rectangles were again significantly lower than for the other two tasks. In the model, the inclusion of RK supports appears to explain the similar performance of ML and non-ML students on the tasks overall. These supports appear to have been helpful for ML students while having no detectable effect for non-ML students, as illustrated in Figure 6.

Figure 6

Model C-4 Expected Scores for Correctness, by Students' ML Status and Task RK Support



Note: Expected Scores are adjusted for Task and Grade Level

* Significant interaction of ML status and Task RK support (t (df) = 2.27 (108), $p < .05$)

Models C-5 and C-6 examined the combined effects of students' EC-RK and the inclusion of RK supports in task design. In Model C-5 these two factors were included as main effects only, with the results again indicating that students' EC-RK was a strong, positive predictor of correctness (t (df) = 7.01 (108), $p < .05$), while the inclusion of RK supports did not itself predict correctness scores. Also, adjustments for the inclusion of these two main effects did not result in a detectable difference in correctness scores by students' ML status.

In Model C-6, students' EC-RK and the inclusion of RK supports were examined in interaction with students' ML status. In this model, only the main effect of students' EC-RK was significant, and it was positive (t (df) = 6.62 (106), $p < .05$). This result suggests that once the positive association of students' EC-RK is accounted for, neither students' ML status nor the inclusion of RK supports in the task design help to explain correctness of scores. The reductions in variance at both levels, compared to Model C-0, are similar to Models C-1 and C-2, also suggesting that accounting for students' EC-RK is responsible for these results.

Summary of Findings

There were statistically significant differences in both EC-RK and Correctness scores across the three tasks; these differences were accounted for by including task type in the analyses. Overall, ML students and non-ML students tended to give similarly correct responses to the tasks; that is, when all responses to the tasks were considered, regardless of task version, there was no significant difference in students' Correctness scores related to ML status (Model C-0). Neither students' EC-RK nor the correctness of their responses differed by grade level.

Findings for the three research questions are summarized in Table 4. The specific evidence from analytic results to support each finding is presented in the sections that follow.

Table 4

Summary of Findings by Research Question

Research Question	Findings
1. What is the relationship between evidence of cognition in students' record keeping and performance on tasks?	Higher EC-RK scores were associated with higher Correctness scores, regardless of ML status or provision of RK supports.
2. How does student performance differ on tasks designed with supports for record keeping versus tasks without these supports?	Across all students, the inclusion of RK supports on tasks did not account for differences in students' EC-RK or Correctness scores.
3. What differences are evident in the impacts of RK supports for multilingual learner students compared to first-language English speaking students?	Non-ML students' Correctness and EC-RK scores did not differ for tasks with and without the RK supports. ML students' Correctness scores were higher on tasks with the RK supports, even though ML students' EC-RK scores did not differ on tasks with and without the RK supports.

What is the Relationship Between Evidence of Cognition in Students' RK and Performance?

Students whose RK provided more evidence of their cognitive process in problem solving tended to have higher scores for correctness than those whose RK provided less of this evidence, regardless of ML status, as indicated by the positive association between EC-RK and Correctness (Models C-1 and C-2) and the non-significant interaction between ML status and EC-RK (Model C-2).

How Does Student Performance Differ on Tasks Designed With Supports for RK Versus Tasks Without These Supports?

There was no difference in overall student performance on tasks that included the RK supports and tasks that did not; as a whole, students' RK did not provide greater evidence of their cognitive processes, nor did they respond more correctly on the task versions with RK supports. That is, there was no significant difference between the expected scores on tasks with RK supports and those without RK supports for either EC-RK (Model RK-1) or Correctness (Model C-3). In addition, the positive association between students' EC-RK and the correctness of their responses was similar on the RK-S and RK-U versions of tasks, as indicated by the similarity between Models C-1, which did not include the RK Supports variable, and C-5, in which RK Supports had a non-significant effect.

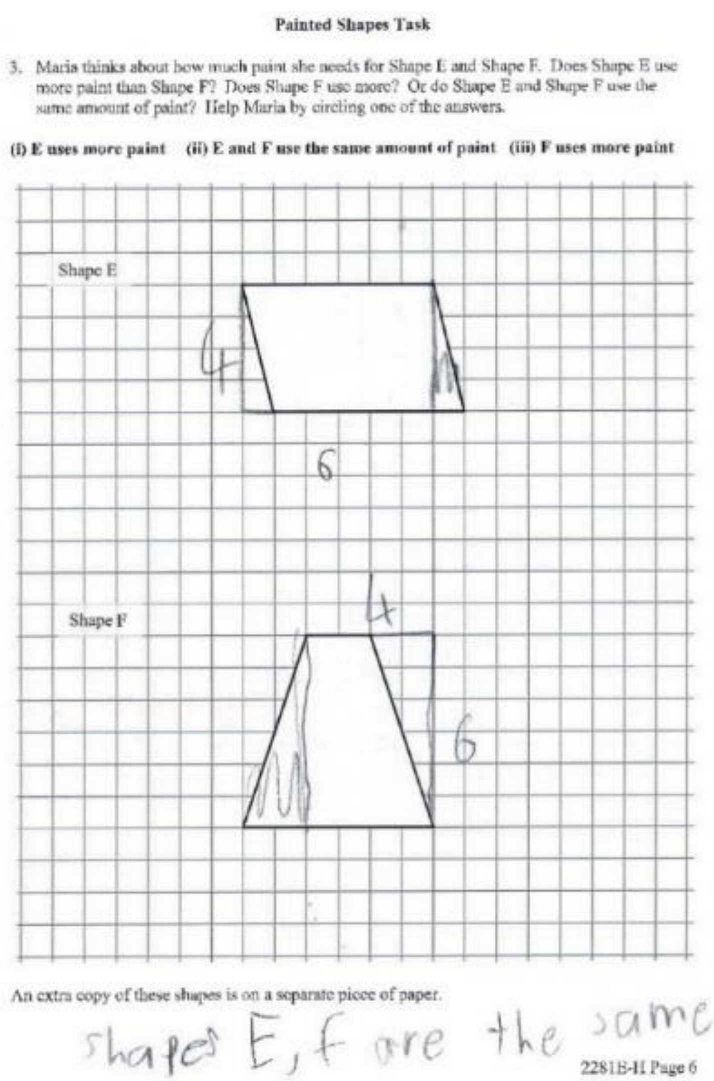
What Differences are Evident in the Impacts of RK Supports for Emergent Multilingual Students Compared to First-language English Speaking Students?

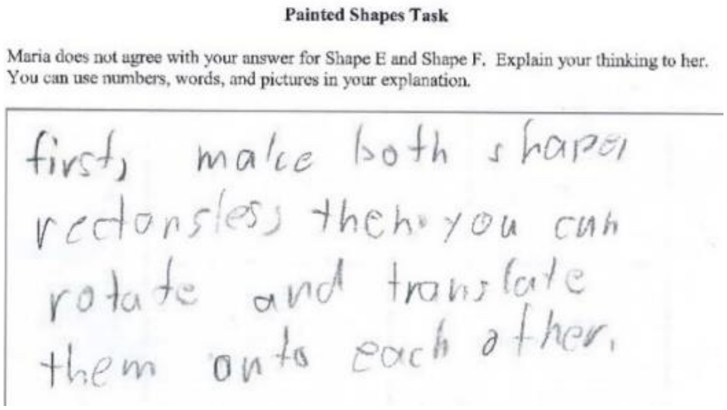
ML students' responses to tasks that included RK supports were more correct than their responses to versions without the RK supports, unlike non-ML students, for whom no difference was detected. No other significant differences were detected between ML students and non-ML students on the RK-S and RK-U versions of the tasks. The different impact on Correctness for MLs and non-MLs is evident in the significant positive association between the RK Supports*ML status interaction term and Correctness (Model C-4). The RK supports appear to explain the similar overall correctness of MLs' and non-MLs' responses, as suggested by the presence of a significant negative association between ML status and Correctness in the model that includes the significant positive association for the RK Supports*ML status interaction (Model C-4) but not otherwise.

These results were somewhat contradictory to the full set of hypotheses originally driving our investigation, namely that RK supports would lead to more RK in general, yielding greater EC-RK that would in turn lead to greater correctness. Our results indicate that ML students developed more correct responses to the RK-S versions of the tasks *even though* their RK on those versions did not offer greater evidence of their cognitive processes in problem solving. To illustrate this finding and offer an example for further investigation, Figures 7 and 8 show two ML students' work on RK-S and RK-U versions, respectively, of the Painted Shapes task.

Figure 7

Student Work on the Painted Shapes Task (RK-S)

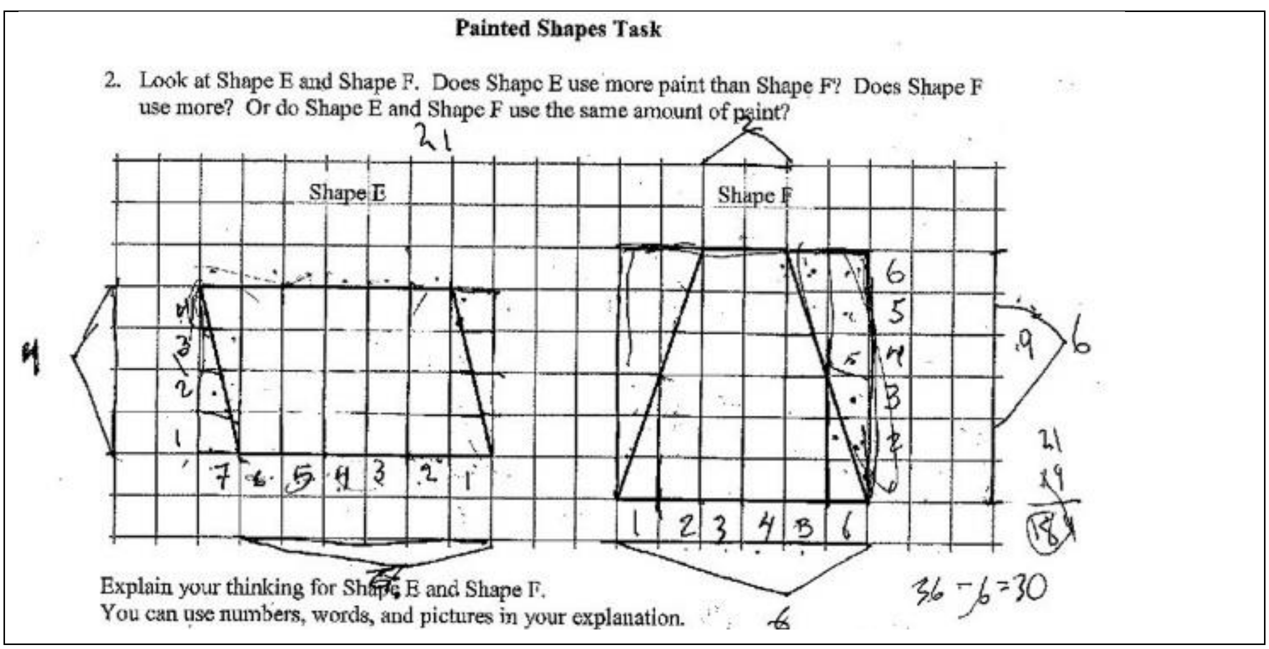


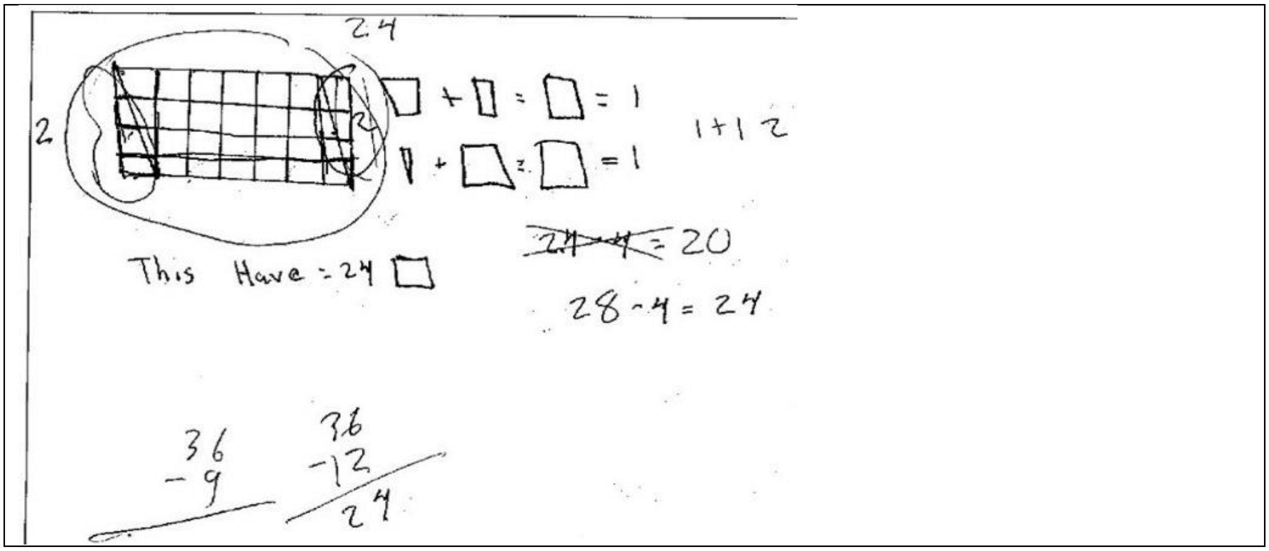


Note: In Figure 7, the student generated rectangles that had the same areas as the shapes and recognized that both shapes have areas equal to the same rectangle (though it was rotated in one case). In Figure 8, the student working with the RK-U version of the task created extensive records, however, no comparison was made between the areas of the shapes. The multiple calculations for one figure's area were provided without clear indication of which was final.

Figure 8

Student Work on the Painted Shapes Task (RK-U)





Both responses were rated at level 3 on the EC-RK rubric, which indicated evidence of RK supporting conceptualization of and solution to the problem, but without all elements of the solution process represented in RK. The two responses, however, were judged differently for correctness, with the first rated a 4 for its correct comparison without errors related to concepts or calculations. This student did not actually calculate the areas of the two figures, but rather decomposed and recomposed them into figures that could be determined as congruent via rotation and translation, leading to the conclusion that they require the same amount of paint.

The second response was rated a 2 for correctness because the strategy of partitioning the figures into portions made up of whole unit squares and partial unit squares is viable as is the approach of enveloping the parallelogram (Shape E) and the trapezoid (Shape F) within a rectangle and subtracting enclosed areas that are not part of the target figure's area. However, the student's work does not apply this approach consistently, leading to a correct determination of the area of Shape E but not Shape F because the subtraction of area outside Shape F was incomplete. In fact, the records in the explanation portion of the student's work show the full and correct subtraction for Shape E, but not for Shape F. One of the unconnected calculations within this portion of the response (36-12=24) may actually represent the correct subtraction for Shape F, but the response does not include an explanation for calculating the area of Shape F comparable to the drawing used for Shape E. Such an explanation for Shape F may have led the student to notice the original error in subtraction of areas and then identify the correct conclusion.

Examples such as these allow us to hypothesize how the RK-S versions may have supported ML students in a few ways. First, the RK supports potentially improved ML students' understanding of what solving the task entailed. The RK-S version of Painted Shapes had the student first practice drawing a figure that required the same amount of paint as a given figure, so the notion of comparing areas of figures was elicited in this support that specifically prompted creation of a record. Also, the RK-S version provided response options below the question (see Figures A3 and A4 in Appendix A) to reinforce what the question asked and what sorts of results a solution could lead to. It is notable that the work shown in Figure 7 provides an answer to the task's question, but the work shown in Figure 8 does not explicitly do so.

Our exploratory examination of such examples from ML students, with similar EC-RK but varying correctness, suggests that these two features may have enhanced ML students' ability to interpret correctly what was being asked in the task, leading to more complete solutions to what the task required. Other features of the RK-S version, such as additional space and the audience (a fictitious student named in the task instructions in Figure 7) for the solution, may have enabled students to organize and make use of their RK to manage the cognitive load needed to solve the problem in ways the EC-RK rubric did not indicate. The records in Figure 7 are organized and succinct in how targeted they are to the cognitive process the student has used in their solution. Many more records arose in the solution shown in Figure 8, with the lack of space making it crowded. The diagrams and calculations, while demonstrating much of the student's geometric and numeric cognitive processes, are not organized in a way that makes clear what result the student's solution to the task supports.

Discussion

We found within our sample that students whose RK provided greater evidence of their cognitive processes when completing measurement/geometry tasks were more successful in solving the tasks. The tasks were, by design, complex enough to make it difficult for students to do all needed work mentally, so offloading through RK provided a way to manage intrinsic load and avoid extraneous load. The association between evidence of cognitive processes in RK and correctness of responses also suggests that the nature of students' records matters. All students engaged in some RK for almost all of their responses. However, higher scores on our EC-RK rubric required RK that appeared to help students conceptualize the task and that exhibited connections among different records. In many high-scoring responses, RK that was not inherently meaningful, such as counting dots, was present, and it was connected to more meaningful records, such as numeric labels. These characteristics suggest that the records may serve purposes beyond offloading some of the cognitive demand for storage and retrieval. It appears that students' creation of records contributed to their thinking process, as others have posited (e.g., Chu et al., 2017; Meira, 1995; Paas & van Merriënboer, 2020).

The RK supports we included in the design of the tasks appeared to help ML students solve the tasks correctly, *even though* the supports did not result in significantly greater evidence of cognitive processes in their RK. As explained in the results, this finding did not fully reflect the chain of hypotheses for the study yet suggests the RK supports were useful to students in ways apart from generating RK that reflected their cognitive processes. MLs may experience heightened intrinsic and extraneous load associated with solving tasks due to language demands in the tasks (Barbu & Beal, 2010). These tasks with and without RK supports were designed to have the same intrinsic demand, and both versions of each task were designed to minimize extraneous demand. However, the features designed to support RK may have promoted ML students' understanding of the requirements of the task, and organizational features and additional space to support RK may have led to greater utility of RK for MLs. That is, although we did not observe greater evidence of cognitive processes in RK on tasks designed to support it, the supports nonetheless appear to have aided MLs in utilizing their problem-solving assets more effectively to solve tasks correctly. RK supports should be studied further, as our results suggest that they may have strengthened students' ability to effectively use RK or related assets (e.g., translanguaging) for processes such as interpreting the language of the task, offloading and retrieving information, or making connections, and thus contributed to ML students' success in solving the tasks.

There were several limitations to this exploratory study that have implications for future research. We found a correlation between evidence of cognitive processes in RK and correct work in a small sample study of Grades 7 to 9 students' work on three mathematical tasks, and it will be important to examine students' RK on other tasks, at other grade levels, and in content areas other than geometry and measurement to establish the extent to which these results might generalize.

The rubrics we developed for this study, particularly EC-RK, focused our work but also narrowed our view. Our definition of RK requires that records be externalized, and therefore observable, but applying the details of the rubric required interpretations about the purposes and connections among records that may not have been clearly observable. Think-aloud or stimulated recall studies would reveal more than we were able to understand. In addition, more interpretive studies may provide insights into the mechanisms by which RK supports students' success in solving problems. Our follow-up interpretations of students' RK provide clues to how the extent and quality of RK may support correctness, but further studies are needed to investigate whether RK is an explanatory factor in increased correctness or if there is some other underlying factor that explains both RK and correctness.

Although we attempted to collect students' self-report of current or past receipt of school ESL services, we necessarily relied on teachers' reports for grouping students in the study. In either case, we acknowledge that we have characterized the multi-faceted identity of multilingual learner with a very simple designation of ML and non-ML. We were not able, in this study, to examine the influence of varying first languages, language proficiencies, or experiences apart from receipt of school-provided services. Research knowledge around these ideas is rich and rapidly developing. We hope to inspire more nuanced research on the intersection of RK and other facets of MLs' language and mathematics experiences.

Finally, we intentionally limited our study to RK in written form, because the spontaneous and variable use of written records we had observed inspired our research. Digital platforms that would permit RK when solving tasks such as these were not readily available and familiar to students at the time of the study. The more common use now of digital platforms for students to conduct and document their mathematics work is structurally different from writing alone, certainly influencing the potential for designing RK supports and potentially influencing how students will use RK and to what effect.

In practice, our findings have implications for mathematics teaching, teacher preparation, and task design. The correlation between evidence of cognitive processes in students' RK and their success in solving tasks implies that encouraging and supporting RK can aid students in successful problem solving. Task design alone did not provide sufficient encouragement and support to result in increased evidence of cognitive processes in RK of a form that supports success. However, task design to support RK does appear to aid students, especially MLs, in interpreting and accessing problems and in using RK effectively to solve problems. Curriculum material designers and teachers can incorporate RK supports into tasks and assignments. In our experience, student materials often do not provide structure for students to explore a problem through organized RK. Materials designers could add features we identified that offer this structure. Teachers can be prepared to take advantage of task design features that include these supports by explicitly helping students use RK for identifying and organizing relevant information as well as offloading cognitive demand during problem solving. Student materials also seldom provide space for exploring and solving problems. Teachers can format handouts to provide ample blank space, ensure that diagrams are large enough for students to write or draw in, and make extra copies of the task or parts of it available. Students may also be hesitant to make use of blank space in this way. During this study, in fact, many students specifically asked the interviewers if they could write on diagrams or in the blank space provided, suggesting that teachers should explicitly permit or encourage students to use available space and resources for RK.

The relationship between evidence of cognitive processes in RK and correctness of solutions for all students, but the failure of the RK supports within the tasks to generate such records, suggests that additional work is needed to understand how to promote and support effective RK. Other efforts in teaching may be required to engage students in showing evidence of cognitive processes in their RK, such as teacher modeling, highlighting effective RK in presentations of student work, and questioning techniques that press for students to record their cognitive processes while problem solving. Most importantly, alongside support for ML students' assets such as translanguaging, supporting RK in task design and encouraging RK in teaching may strengthen MLs' mathematical engagement in getting started, persisting, and succeeding in solving problems.

Funding

This material is based upon work supported by the U.S. National Science Foundation under Grant No. DRL- 1348810. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Daniel J. Heck (dheck@horizon-research.com), President of Horizon Research, Inc. (HRI), holds a Ph.D. in Education from the University of Illinois at Urbana-Champaign, with a specialization in Educational Psychology, Quantitative and Evaluative Research Methodologies. In more than 25 years at HRI, Dr. Heck has been Principal Investigator of multiple research grants for studies in mathematics learning, learning environments, teaching, and professional development, and studies of computing in K-12 education.

Anthony Fernandes (anthony.fernandes@charlotte.edu) is Professor of Mathematics Education in the Department of Mathematics and Statistics at the University of North Carolina at Charlotte. His research lies at the intersection of language and mathematics, with a focus on how Multilingual learners utilize multimodal resources to engage with mathematical concepts. Recently, Dr. Fernandes has turned his attention to critical statistical literacy, designing statistical investigations that foster meaningful dialogue around issues of institutional racism. By integrating these investigations into preservice teacher education, he aims to cultivate critical consciousness, normalize discussions of race and racism in mathematics and statistics classes, and equip future educators to understand and address the systemic factors that shape people's lives.

Johannah Nikula (jnikula@edc.org), a mathematics education expert and Senior Project Director at Education Development Center, leads a body of research focused on making engaging, rigorous mathematics accessible to all learners. She partners with teachers, administrators, and state departments of education while developing resources and engaging in research. She has particular interest in mathematics teaching and learning for students who are multilingual learners.

Evelyn M. Gordon (egordon@horizon-research.com) is a Senior Researcher at Horizon Research, Inc. She conducts STEM education research and is an external evaluator for STEM education research projects. Her work focuses on mathematics education, teacher preparation and professional development, and multilingual learners.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219-234. https://doi.org/10.1207/s15324818ame1403_2

Barbu, O. C., & Beal, C. R. (2010). Effects of linguistic complexity and math difficulty on word problem solving by English learners. *International Journal of Education, 2*(2), 1-19. https://doi.org/10.5296/ije.v2i2.508

Bodner, G. M., & Domin, D. S. (2000). Mental models: The role of representations in problem solving in Chemistry. *University Chemistry Education*, *4*(1), 24-30.

Campbell, A. E., Adams, V. M., & Davis, G. E. (2007). Cognitive Demands and Second-Language Learners: A Framework for Analyzing Mathematics Instructional Contexts. *Mathematical*

Thinking & Learning: An International Journal, 9(1), 3-30. https://doi.org/10.1080/10986060709336603

Cellucci, C. (2019). Diagrams in mathematics. *Foundations of Science*, *24*(3), 583–604. https://doi.org/10.1007/s10699-019-09583-x

Chu, J., Rittle-Johnson, B., & Fyfe, E. R. (2017). Diagrams benefit symbolic problem-solving. *British Journal of Educational Psychology*, *87*(1), 273–287. https://doi.org/10.1111/bjep.12149

De Corte, E., Verschaffel, L., & De Win, L. (1985). Influence of rewording verbal problems on children's problem representations and solutions. *Journal of Educational Psychology, 77*(4), 460-470. http://dx.doi.org/10.1037/0022-0663.77.4.460

De Toffoli, S. (2018). *Epistemic roles of mathematical diagrams* (Publication No. 2509013480) [Doctoral dissertation, Stanford University], ProQuest Dissertations & Theses Global.

Diezmann, C. M., & English, L. D. (2001). Promoting the use of diagrams as tools for thinking. In A. Cuoco (Ed.), *2001 National Council of Teachers of Mathematics Yearbook: The Role of Representation in School Mathematics* (pp. 77-89). National Council of Teachers of Mathematics.

Driscoll, M., Heck, D. J., & Malzahn, K. A. (2012). Knowledge for teaching English language learners mathematics: A dilemma. In S. Celedon-Pattichis & N. Ramirez (Eds.), *Beyond good teaching: Advancing mathematics education for ELLs* (pp. 163–181). National Council of Teachers of Mathematics.

Echevarría, J., Vogt, M., & Short, D. J. (2017). *Making content comprehensible for English learners: The SIOP model.* Pearson.

Elshafie, M., & Zhang, J. (2024). Pedagogical translanguaging in content areas: Exploring preservice teachers' lesson plans for emergent bilinguals. *Education Sciences*, *14*(7), NA. http://dx.doi.org/10.3390/educsci14070702

Fernandes, A., Heck, D., & Nikula, J. (2015). Student record keeping for cognition and communication. In T. G. Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez, (Eds.). *Proceedings of the 37th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (pp. 399-402). Michigan State University.

Fernandes, A., Kahn, L., & Civil, C. (2017). A closer look at bilingual students' use of multimodality in the context of an area comparison problem from a large-scale assessment. *Educational Studies in Mathematics*, *95*(3), 263-282. https://doi.org/10.1007/s10649-017-9748-5

Fernandes, A. & McLeman, L. (2012). Interpreting and using gestures of English language learners in mathematics teaching. *Teaching Equity and Excellence in Mathematics, 4*(1),16-23. https://doi.org/10.63966/teem.v4i1.1725

Fischbein, E. (1977). Image and concept in learning mathematics. *Educational Studies in Mathematics*, *8*(2), 153-165. https://doi.org/10.1007/bf00241022

García, O. (2023). Translanguaged TESOL in transit. *NYS TESOL Journal*, *10*(1), 5–18.

García, O., & Solorza, C. R. (2021). Academic language and the minoritization of U.S. bilingual Latinx students. *Language and Education*, *35*(6), 505–521. https://doi.org/10.1080/09500782.2020.1825476

Garza, E., & Arreguín-Anderson, M. G. (2018). Translanguaging: Developing scientific inquiry in a dual language classroom. *Bilingual Research Journal*, *41*(2), 101–116. https://doi.org/10.1080/15235882.2018.1451790

González-Howard, M., Andersen, S., Méndez Pérez, K., & Suárez, E. (2023). Language views for scientific sensemaking matter: A synthesis of research on multilingual students' experiences with science practices through a translanguaging lens. *Educational Researcher*, *52*(9), 570-579. https://doi.org/10.3102/0013189X231206172

Gordon, E. M., Heck, D. J., Fernandes, A., Smith, A. A., & Moffett, G. E. (2015). How students' record keeping during problem solving can support cognition and communication. *For the Learning of Mathematics, 35*(2), 22-25.

Grapin, S. E., Ramos Borrego, M., & Navarro, V. G. (2025). Translanguaging in US K–12 science and engineering education: A review of the literature through the lens of equity. *Journal of Research in Science Teaching*, *62*(1), 15–48. https://doi.org/10.1002/tea.22012

Heck, D. J., Gordon, E. M., & Lyons, K. M. (2015, October 10). *Promoting middle grade students' access to mathematics through tasks designed to support record keeping* [Paper presentation]. University of North Carolina-Chapel Hill School of Education Research Symposium: Honoring the Legacy of Carol E. Malloy, Chapel Hill, NC, United States.

Larkin, J., McDermott, J., Simon, D. P., & Simon, H. A. (1980). Expert and novice performance in solving physics problem. *Science, 208*(4450), 1335–1134.

LeBlanc, M. D., & Weber-Russell, S. (1996). Text integration and mathematical connections: A computer model of arithmetic word problem solving. *Cognitive Science, 20*(3), 357– 407. https://doi.org/10.1016/s0364-0213(99)80010-x

Liljedahl, P., Santos-Trigo, M., Malaspina, U., & Bruder, R. (2016). *Problem solving in mathematics education*. Springer Nature. https://doi.org/10.1007/978-3-319-40730-2

Lindquist, M., Philpot, R., Mullis, I. V. S., & Cotter, K. E. (2017). TIMSS 2019 Mathematics Framework. In I. V. S. Mullis & M. O. Martin (Eds.), *TIMSS 2019 Assessment Frameworks* (pp. 11-25). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timssandpirls.bc.edu/timss2019/frameworks/

Martiniello, M. (2008). Language and the performance of English-language learners in math word problems. *Harvard Educational Review, 78*(2), 333-368. https://doi.org/10.17763/haer.78.2.70783570r1111t32

Meira, L. (1995). The microevolution of mathematical representations in children's activity. *Cognition and Instruction*, *13*(2), 269-313. https://doi.org/10.1207/s1532690xci1302_5

Moschkovich, J. (1999). Understanding the needs of Latino students in reform-oriented mathematics classrooms. In W. Secada (Ed.), *Changing the faces of mathematics: Perspectives on Latinos* (pp. 5-12). Reston, VA: National Council of Teachers of Mathematics.

Moschkovich, J. (2002). A situated and sociocultural perspective on bilingual mathematics learners. *Mathematical Thinking & Learning, 4*(2-3), 189-212. https://doi.org/10.1207/s15327833mtl04023_5

Moschkovich, J. N. (2008). "I went by twos, he went by one": Multiple interpretations of inscriptions as resources for mathematical discussions. *The Journal of the Learning Sciences*, *17*(4), 551-587. https://doi.org/10.1080/10508400802395077

Moschkovich, J. N. (Ed.). (2010). *Language and mathematics education: Multiple perspectives and directions for research*. Information Age Publishing.

Murata, A. (2008). Mathematics teaching and learning as a mediating process: The case of tape diagrams. *Mathematical Thinking and Learning, 10*(4), 374–406. https://doi.org/10.1080/10986060802291642

National Council of Teachers of Mathematics (NCTM). (2000). *Principles and standards for school mathematics.* NCTM.

National Council of Teachers of Mathematics (NCTM). (2014). *Principles to actions: Ensuring mathematical success for all.* NCTM.

National Governors Association Center for Best Practices and Council of Chief State School Officers. (2010). *Common core state standards for mathematics.* Author.

Neumayer-DePiper, J., Heck, D. J., Fernandes, A., & Nikula, J. (2015, April 13). Task design to support English learners' geometric record keeping. [Poster presentation]. National Council of Teachers of Mathematics research conference, Boston, MA, United States.

Ng, S. F., & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education, 40*(3), 282–313. https://doi.org/10.5951/jresematheduc.40.3.0282

Nunokawa, K. (1994). Improving diagrams gradually: One approach to using diagrams in problem solving. *For the Learning of Mathematics, 14*(1), 34–38.

Nunokawa, K. (2005). Mathematical problem solving and learning mathematics: What we expect students to obtain. *The Journal of Mathematical Behavior*, *24*(3-4), 325-340. https://doi.org/10.1016/j.jmathb.2005.09.002

Paas, F., & van Merriënboer, J. J. G. (2020). Cognitive-Load Theory: Methods to manage working memory load in the learning of complex tasks. *Current Directions in Psychological Science*, *29*(4), 394-398. https://doi.org/10.1177/0963721420922183

Polya, G. (1957). *How to solve it: A new aspect of mathematical method*, 2nd edition. Doubleday Anchor Books.

Purchase, H. C. (2014). Twelve years of diagrams research. *Journal of Visual Languages & Computing*, *25*(2), 57–75. https://doi.org/10.1016/j.jvlc.2013.11.004

Renkl, A., & Atkinson, R. K. (2003). Structuring the transition from example study to problem solving in cognitive skill acquisition: A cognitive load perspective. *Educational Psychologist, 38*(1), 15-22. https://doi.org/10.1207/s15326985ep3801_3

Rexigel, E., Kuhn, J., Becker, S., & Malone, S. (2024). The more the better? A systematic review and meta-analysis of the benefits of more than two external representations in STEM Education. *Educational Psychological Review*, *36*(4), 124. https://doi.org/10.1007/s10648-024-09958-y

Schoenfeld, A. H. (1980). Teaching problem-solving skills. *The American Mathematical Monthly*, *87*(10), 794-805. https://doi.org/10.1080/00029890.1980.11995155

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 334-371). Macmillan.

Schoenfeld, A. H. (2014). *Mathematical problem solving*. Elsevier.

Stylianou, D. A., & Silver, E. A. (2004). The role of visual representations in advanced mathematical problem solving: An examination of expert-novice similarities and differences. *Mathematical Thinking & Learning, 6*(4), 353-387. https://doi.org/10.1207/s15327833mtl0604_1

Sunzuma, G., Chando, C., Gwizangwe, I., Zezekwa, N., & Zinyeka, G. (2020). In-service Zimbabwean teachers' views on the utility value of diagrams in the teaching and learning of geometry. *LUMAT: International Journal on Math, Science and Technology Education*, *8*(1), 1–18. https://doi.org/10.31129/LUMAT.8.1.1316

Swanson, H. L. (2011). Working memory, attention, and mathematical problem solving: A longitudinal study of elementary school children. *Journal of Educational Psychology*, *103*(4), 821. https://doi.org/10.1037/a0025114

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(2), 257–285. https://doi.org/10.1016/0364-0213(88)90023-7

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). Academic. https://doi.org/10.1016/s0079-7421(03)01015-6

Sweller, J., van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*(3), 251–296. https://doi.org/10.1023/a:1022193728205

Tabachneck-Schijf, H. J. M., Leonardo, A. M., & Simon, H. A. (1997). CaMeRa: A computational model of multiple representations. *Cognitive Science, 21*(3), 305-350. https://doi.org/10.1207/s15516709cog2103_3

Willis, G. B., & Fuson, K. C. (1988). Teaching children to use schematic drawings to solve addition and subtraction word problems. *Journal of Educational Psychology, 80*(2), 192-201. https://doi.org/10.1037//0022-0663.80.2.192

Zhang, J. (1997). The nature of external representations in problem solving. *Cognitive Science, 21*(2), 179-217. https://doi.org/10.1016/s0364-0213(99)80022-6

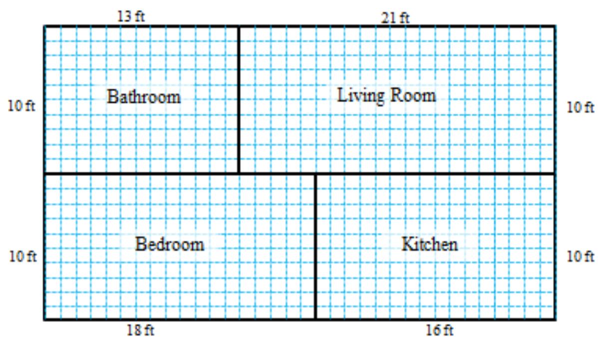
Appendix A: Study Instruments

Figure A1

The Floor Plan Task (RK-U Version)

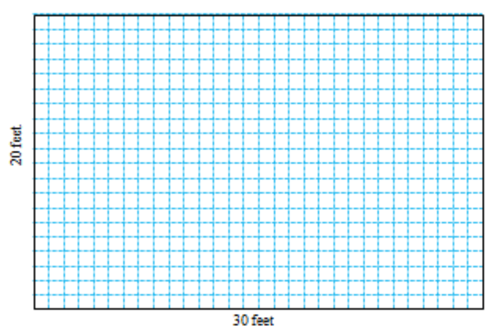
Floor Plan Task

This floor plan shows the four rooms in an apartment. The floor plan shows the measurements for each room.



Design a floor plan for a new apartment with **five rooms**. The apartment will be rectangular with a length of 30 feet and a width of 20 feet.

1. Draw a floor plan for the family that shows the five rooms on the rectangle below. Include a living room, a bathroom, a kitchen, and two bedrooms.
   Label the length of the walls of each room.
   - The five rooms will take up all the space in the apartment. The apartment does not have any halls or other rooms.
   - Make the area of the living room 120 feet squared (ft^2) or bigger.
   - Make the area of the bathroom 50 feet squared.
   - The kitchen and the two bedrooms will be 10 feet long and 10 feet wide, or bigger.



2. What is the area of each room?

«ID» Page 3

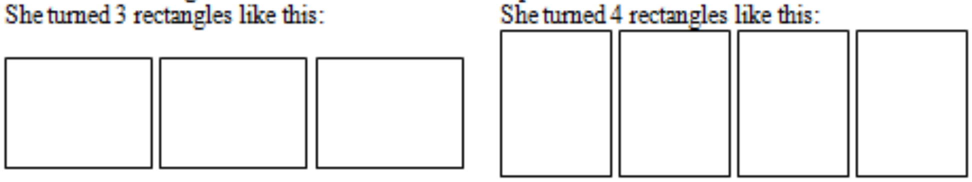
Figure A2

Seven Rectangles Task (RK-S Version; Extra Diagrams Omitted)

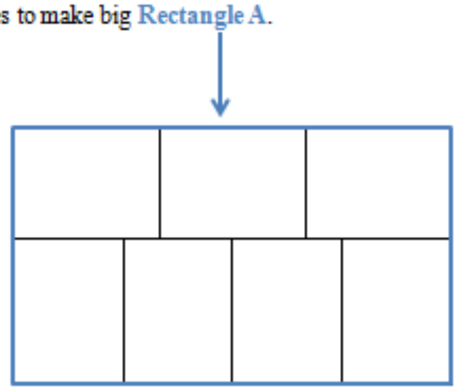
Seven Rectangles Task

Maria has 7 rectangles that are the same size and shape.

She turned 3 rectangles like this: She turned 4 rectangles like this:



She moves the rectangles to make big Rectangle A.



The total area of big Rectangle A is 84 square inches.

1. Find the length and width of one of the small rectangles.

   Length = _______ inches Width = _________ inches

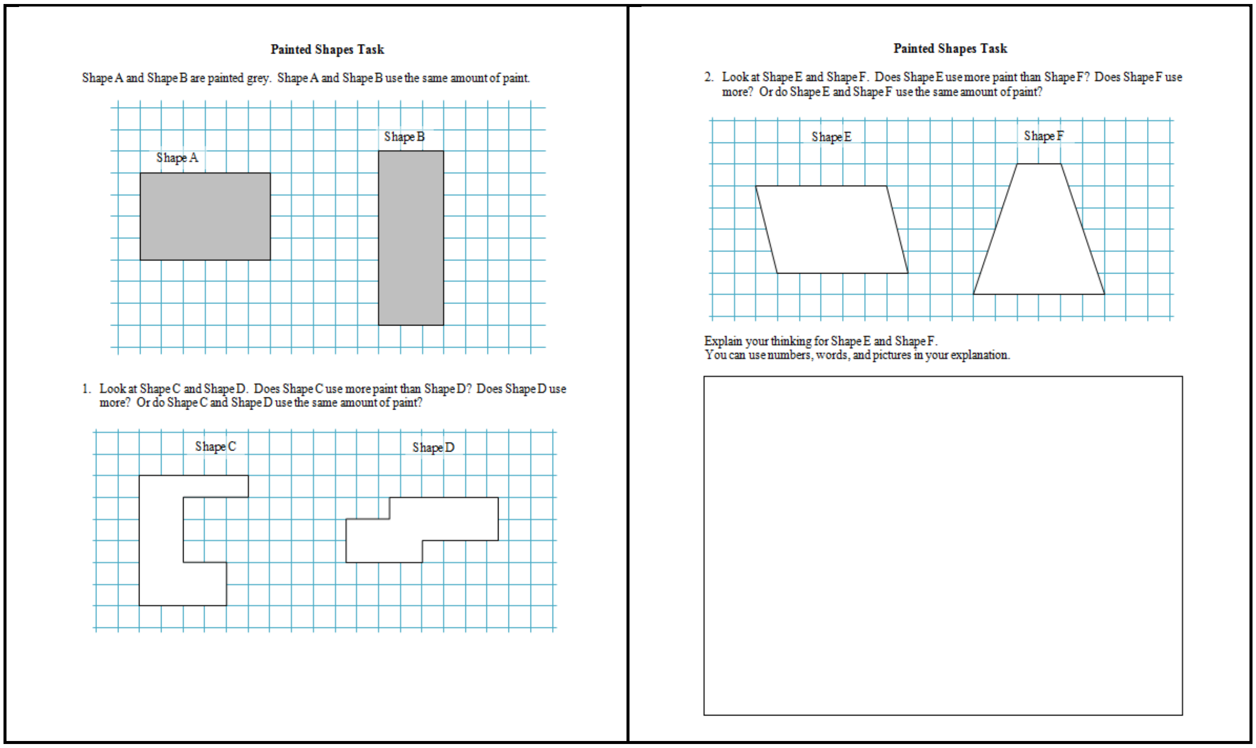
2. What is the perimeter of big Rectangle A?

   The perimeter of big Rectangle A is ________ inches.

«ID» Page 4

Figure A3

Painted Shapes Task (RK-U Version)


Painted Shapes Task

Shape A and Shape B are painted grey. Shape A and Shape B use the same amount of paint.

1. Look at Shape C and Shape D. Does Shape C use more paint than Shape D? Does Shape D use more? Or do Shape C and Shape D use the same amount of paint?

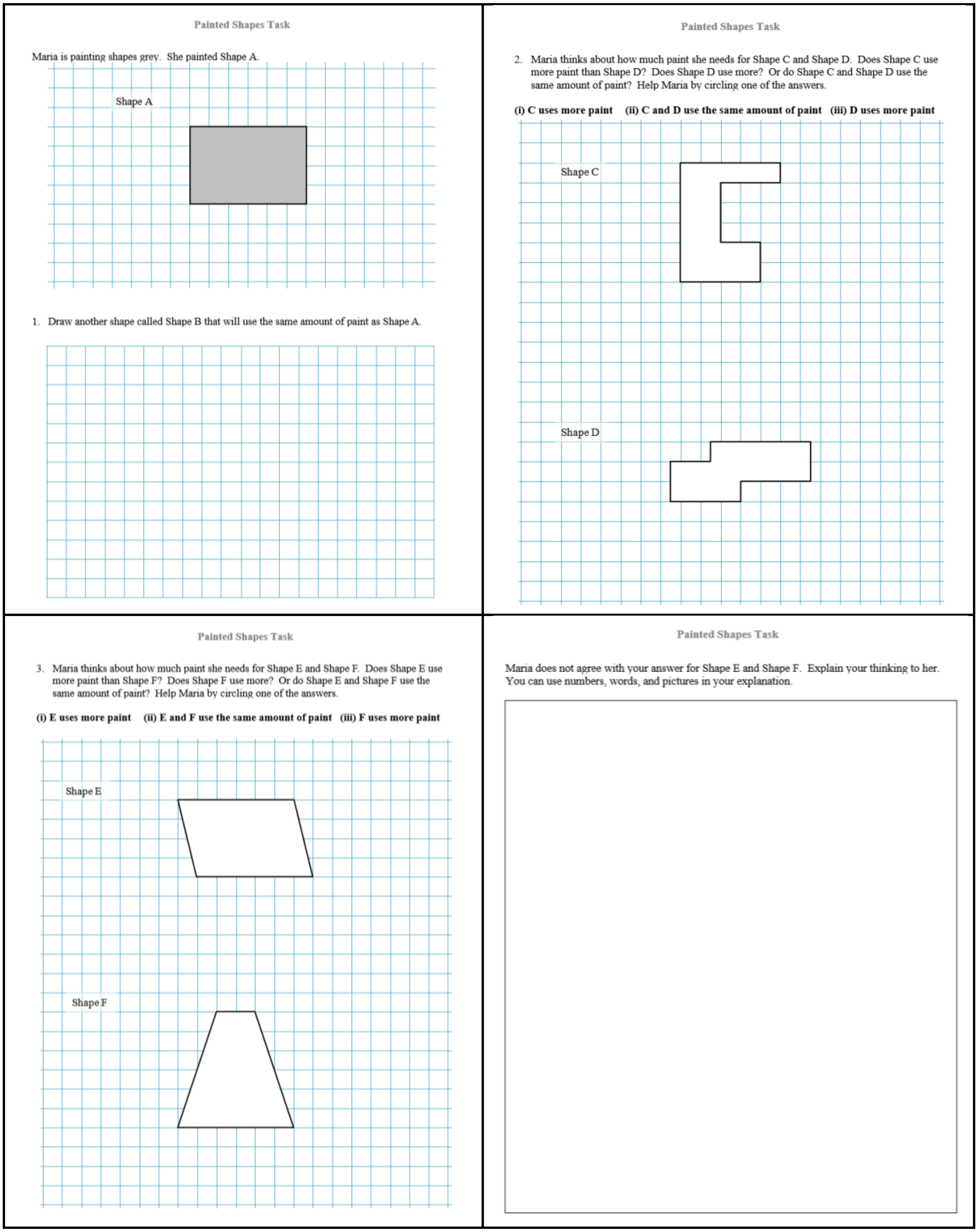
Painted Shapes Task

2. Look at Shape E and Shape F. Does Shape E use more paint than Shape F? Does Shape F use more? Or do Shape E and Shape F use the same amount of paint?

Explain your thinking for Shape E and Shape F.
You can use numbers, words, and pictures in your explanation.

Figure A4

Painted Shapes Task (RK-S Version, Extra Diagrams Omitted)


Painted Shapes Task

Maria is painting shapes grey. She painted Shape A.

Shape A

1. Draw another shape called Shape B that will use the same amount of paint as Shape A.

Painted Shapes Task

2. Maria thinks about how much paint she needs for Shape C and Shape D. Does Shape C use more paint than Shape D? Does Shape D use more? Or do Shape C and Shape D use the same amount of paint? Help Maria by circling one of the answers.

(i) C uses more paint (ii) C and D use the same amount of paint (iii) D uses more paint

Shape C

Shape D

Painted Shapes Task

3. Maria thinks about how much paint she needs for Shape E and Shape F. Does Shape E use more paint than Shape F? Does Shape F use more? Or do Shape E and Shape F use the same amount of paint? Help Maria by circling one of the answers.

(i) E uses more paint (ii) E and F use the same amount of paint (iii) F uses more paint

Shape E

Shape F

Painted Shapes Task

Maria does not agree with your answer for Shape E and Shape F. Explain your thinking to her. You can use numbers, words, and pictures in your explanation.

Figure A5

Sample Correctness Rubric (Painted Shapes Task) and the Evidence of Cognitive Processes in Record Keeping Rubric

MaRKS Task Scoring Rubric for Correct Answer/Workable Approach – Painted shapes Task

Score	Description	Specific Indicators[1]
4	A mathematically sound response that is all correct	• Provides correct answer(s) to both of the comparisons: o Indicates (i) Shape C uses more paint (area of C=19, area of D=15.5) o Indicates (ii) E and F use the same amount of paint (24) AND • There is no evidence[2] of flawed approaches to finding or comparing areas nor small errors in area determination/comparison. AND • Provides an explanation that supports the student's answer for Shapes E and F by comparing the areas using a viable strategy (which may be evident elsewhere in the student's work, not necessarily the "explanation" box) such as: o counting squares or shading parts of the figures o decomposing and recomposing the figures o matching parts of the two figures to compare areas o encasing the figures in larger shapes and subtracting out extra area o multiplicative calculations that appear to use linear dimensions.
3	A response that is mathematically sound, but has some small errors or omissions	• Provides an explanation that supports the student's answer for Shapes E and F that compares the areas using a viable strategy. AND • For the C-D and E-F comparisons, any incorrect comparisons or incorrect area determinations are only due to small errors in calculation, counting, or matching.
2	A response that has some mathematically sound and relevant ideas, but has incorrect answers due to some error in understanding	• There is evidence the student counted, calculated, or compared areas (possibly with errors or omissions) using viable strategies, but arrived at incorrect answers due to a misunderstanding of how to use the strategies. (This includes using a viable strategy that involves different area formulas for different shapes but where some formulas used are incorrect.) OR • There is evidence the student used a viable general strategy to compare areas, but the student consistently used inappropriate strategies for counting partial squares in one or both comparisons. OR • The student correctly answered the comparisons of Shapes C and D and Shapes E and F BUT there is no evidence (including in the explanation for the last question) to indicate how the Shape E and F comparison was made.
1	A response that has fundamental mathematical flaws and leads to incorrect answers (or possibly correct answers, but not for the correct reasons – like using perimeter instead of area but it happens to work out)	• May or may not provide a correct answer(s) to one or both of the comparisons, and there is evidence the student used a non-viable strategy for both comparisons (e.g., comparing with something other than area; comparing areas of encasing rectangles). • There is no evidence of how the student made comparisons and only one comparison is answered correctly.
0	No response or, if answers only, nothing is correct	• The student did not record any answers at all OR • There is no evidence of how the student made comparisons and neither comparison is answered correctly.

[1] DO NOT use student work related to Shape A or Shape B in judging correctness of the solution or a workable approach. Shapes A and B are different in the supported and unsupported versions and are not part of the task to be scored.
[2] "Evidence" can come from student's record keeping or from researcher notes.

MaRKS Scoring Rubric for Record Keeping

Score	Record Keeping	Description
4	The student's cognitive process on the task is evident from the RK.	A set of steps for arriving at the answer is evident from the RK (why a false start was abandoned may not be evident). Records are purposeful[1] and connected.[2] Some simple calculations/counts may have been done mentally.
3	It is evident how RK supported the student's cognition on the task but some aspects are not evident from the RK.	It is clear how RK supported most, but not all, of the student's cognition and conceptualization of the task. The purpose for most records is evident, but some connections are not apparent.
2	It is clear how the student was using RK to support cognition on isolated parts of the task and provides evidence about the student's conceptualization of part(s) of the task.	Few, if any, connections are made between records. The purpose for at least some records is evident from the records, not from scorer's understanding of the task. **All tasks**: Knowing where numbers in a calculation came from based on diagram labeling provides sufficient evidence of conceptualization to rate at least a 2. **Painted Shapes**: Drawn grid lines and evidence of counting provide sufficient evidence of conceptualization to rate at least a 2. Auxiliary lines suggesting decomposition on at least 2 figures provide sufficient evidence of conceptualization to rate at least a 2. **Floor Plan**: Labels for dimensions and room names on the floor plan and area calculations or results for named rooms provide sufficient evidence of conceptualization to rate at least a 2. Check marks next to items on the list of floor plan criteria and evidence of attempts to address those criteria (maybe incorrectly) provide sufficient evidence of conceptualization to rate at least a 2.
1	It appears the student used RK primarily to offload isolated procedural or operational work.	The student did some RK, but the RK did not evidently support conceptualization of the task. There are no evident connections between records, and the records are not explicitly related to elements of the task.
0	The student did not use RK for cognition.	There is no RK to support cognition.

Note on student use of mental math: If the student's written calculations are missing, but their other records show enough information to bypass the writing down of the calculations their score has the potential of being a 4, assuming all of the other qualifications of a 4 are present.

[1] "Purposeful" means that there is evidence to show how records are related to the task.
[2] "Connected" means that there is evidence to show how records are related to one another.

Figure A6

Background Survey for Students

Brief Background Survey

Please respond to the following questions to help us better understand your background and math experiences.

1. Do you consider yourself:
   - □ Male
   - □ Female

2. Do you consider yourself to be of Hispanic or Latino origin?
   - □ Yes
   - □ No

3. Do you consider yourself: **[Select all that apply.]**
   - □ American Indian or Alaska Native
   - □ Asian
   - □ Black or African American
   - □ Native Hawaiian or Other Pacific Islander
   - □ White

4. How old are you?

5. What grade are you in?
   - □ Grade 6
   - □ Grade 7
   - □ Grade 8
   - □ Grade 9

6. What math class are you in? **[Select all that apply.]**
   - □ Grade 6 Remedial/Review Mathematics
   - □ Grade 6 Regular Mathematics
   - □ Grade 6 Accelerated Mathematics
   - □ Grade 7 Remedial/Review Mathematics
   - □ Grade 7 Regular Mathematics
   - □ Grade 7 Accelerated Mathematics
   - □ Grade 8 Remedial/Review Mathematics
   - □ Grade 8 Regular Mathematics
   - □ Grade 8 Accelerated Mathematics
   - □ Pre-Algebra
   - □ Algebra 1
   - □ Geometry
   - □ Algebra II
   - □ Other:

7. Are you in an English as a Second Language (ESL) program in school?
   - □ Yes
   - □ No

8. Have you been in an ESL program in the past?
   - □ Yes
   - □ No

THANK YOU.

Appendix B: Results Tables and Equations

Table B1

Summary of Scores

	N	0	1	2	3	4
Correctness Score	168	6	24	26	14	29
EC-RK Score	168	7	14	31	29	19

Equation B1

Equations for EC-RK Models

Level 1: EC-RK = Intercept + (N1*7 Rectangles + N2*Floor Plan)[1] + N3*RK Support[2] + e
Level 2: Intercept = (Grade 8 + Grade 9)[1] + ML status[1] + r
N1 = Int + r
N2 = Int + r
N3 = Int + ML status[3] + r

Note: Items at Level 1, Students at Level 2

Equation B2

Equations for Item Correctness Models

Level 1: Correctness = Intercept + (N1*7 Rectangles + N2*Floor Plan)[1] + N3*EC-RK[2] + N4*RK Support[2] + e
Level 2: Intercept = (Grade 8 + Grade 9)[1] + ML status[1] + r
N1 = Int + r
N2 = Int + r
N3 = Int + ML status[3] + r
N4 = Int + ML status[3] + r

Note: Items at Level 1, Students at Level 2

Table B2

Foundational Results for EC-RK as Outcome

	Model RK-0	
	Coeff.	*t* (df)
Level 1		
Intercept (G00)	2.38*	19.33 (52)
7 Rectangles (G10)	-0.64*	-4.24 (110)
Floor Plan (G20)	-0.41*	-2.71 (110)
Level 2		
ML status (G01)	-0.29	-1.10 (52)
Grade 8 (G02)	-0.46	-1.57 (52)
Grade 9 (G03)	-0.27	-0.76 (52)
	Remaining Variance	
	Level 1	Level 2
	0.64	0.64

Table B3

Inclusion of RK Supports Results for EC-RK as Outcome

	Model RK-1		Model RK-2	
	Coeff.	*t* (df)	Coeff.	*t* (df)
Level 1				
Intercept (B0)	3.04*	10.57 (52)	2.30*	16.64 (52)
7 Rectangles (B1)	-0.64*	-4.24 (109)	-0.62*	-4.06 (108)
Floor Plan (B2)	-0.41*	-2.72 (109)	-0.38*	-2.47 (108)
RK Supports (B4)	0.16	1.23 (109)	0.16	1.23 (108)
Level 2				
ML status (G01)	-0.29	-1.09 (52)	-0.47	-1.59 (52)
Grade 8 (G02)	-0.46	-1.57 (52)	-0.47	-1.59 (52)
Grade 9 (G03)	-0.26	-0.74 (52)	-0.26	-0.74 (52)
Interactions				
RK Supports*ML status (G41)			0.36	1.34 (108)
	Remaining Variance			
	Level 1	Level 2	Level 1	Level 2
	0.64	0.64	0.64	0.63

Table B4

Foundational Results for Correctness as Outcome

	Model C-0	
	Coeff.	*t* (df)
Level 1		
Intercept (G00)	2.36*	16.82 (52)
7 Rectangles (G10)	-0.64*	-3.95 (110)
Floor Plan (G20)	-0.09	-0.55 (110)
Level 2		
ML status (G01)	-0.49	-1.65 (52)
Grade 8 (G02)	-0.03	-0.10 (52)
Grade 9 (G03)	0.16	0.40 (52)
Remaining Variance		
	Level 1	Level 2
	0.74	0.86

Table B5

EC-RK Results for Correctness as Outcome

	Model C-1		Model C-2	
	Coeff.	*t* (df)	Coeff.	*t* (df)
Level 1				
Intercept (B0)	1.10*	5.23 (52)	1.12*	5.23 (52)
7 Rectangles (B1)	-0.30	-1.84 (109)	-0.31	-1.86 (108)
Floor Plan (B2)	0.13	0.81 (109)	0.12	0.78 (108)
EC-RK (B3)	0.53*	6.98 (109)	0.52*	6.73 (108)
Level 2				
ML status (G01)	-0.34	-1.49 (52)	-0.54	-1.34 (52)
Grade 8 (G02)	0.21	0.83 (52)	0.23	0.89 (52)
Grade 9 (G03)	0.30	0.99 (52)	0.34	1.09 (52)
Interactions				
EC-RK*ML status (G31)	-	-	0.09	0.60 (108)
	Remaining Variance			
	Level 1	Level 2	Level 1	Level 2
	0.68	0.41	0.69	0.41

Table B6

Inclusion of RK Supports Results for Correctness as Outcome

	Model C-3		Model C-4	
	Coeff.	*t* (df)	Coeff.	*t* (df)
Level 1				
Intercept (B0)	2.36*	15.10 (52)	2.36*	15.23 (52)
7 Rectangles (B1)	-0.64*	-3.93 (109)	-0.60*	-3.69 (108)
Floor Plan (B2)	-0.09	-0.55 (109)	-0.03	-0.19 (108)
RK Supports (B4)	0.00	-0.01 (109)	0.00	-0.01 (108)
Level 2				
ML status (G01)	-0.49	-1.65 (52)	-0.82*	-2.49 (52)
Grade 8 (G02)	-0.03	-0.10 (52)	-0.03	-0.10 (52)
Grade 9 (G03)	0.16	0.40 (52)	0.16	0.40 (52)
Interactions				
RK Supports*ML status (G41)			0.66*	2.27 (108)
	Remaining Variance			
	Level 1	Level 2	Level 1	Level 2
	0.75	0.86	0.72	0.84

Table B7

EC-RK and Inclusion of RK Supports Results for Correctness as Outcome

	Model C-5		Model C-6	
	Coeff.	*t* (df)	Coeff.	*t* (df)
Level 1				
Intercept (B0)	1.13*	5.22 (52)	1.19*	5.35 (52)
7 Rectangles (B1)	-0.30	-1.83 (108)	-0.60	-3.69 (106)
Floor Plan (B2)	0.13	0.82 (108)	-0.03	-0.19 (106)
EC-RK (B3)	0.54*	7.01 (108)	0.51*	6.62 (106)
RK Supports (B4)	-0.09	-0.65 (108)	-0.09	-0.68 (106)
Level 2				
ML status (G01)	-0.34	-1.49 (52)	-0.72	-1.73 (52)
Grade 8 (G02)	0.22	0.84 (52)	0.22	0.85 (52)
Grade 9 (G03)	0.30	0.98 (52)	0.32	1.03 (52)
Interactions				
EC-RK*ML status (G31)			0.06	0.43 (106)
RK Supports*ML status (G41)			0.47	1.65 (106)
	Remaining Variance			
	Level 1	Level 2	Level 1	Level 2
	0.69	0.41	0.68	0.41

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 29, NO. 4, 63-80

Exploring Teacher Participation and Engagement: Climate Change Professional Development and Collaboration

Molly Nation
Florida Gulf Coast University

Heather Skaza Acosta
Florida Gulf Coast University

ABSTRACT

Sustainable WATERS addressed the critical need to support Title I middle school teachers and students by creating a community of practice (CoP) around modeling and field exploration of climate impacts on Southwest Florida's watershed. The program integrated virtual and field environments to grow access to tools, technology, and expertise in STEM, allow for teacher and student asynchronous participation, and facilitate a long-term connection. The aim of the professional development program was to create a CoP and network of continued engagement and resource support to support climate change education opportunities for students and six middle school STEM teachers in underserved schools. The study outlines the participatory involvement and outcomes of each participant throughout the course of the study.

Keywords: community of practice, climate education, teachers' professional development

Introduction

Creating equitable access to STEM programs is complex due to economic disparities, geographic isolation, cultural bridges, language barriers, and socio-economic differences amongst districts (Munn et. al., 2018). Inequity in access to high-quality science education occurs especially within Title I schools (Jones & Stapleton, 2017). These schools are typically low-resourced; and situated in low-income communities with a high number of students underrepresented in STEM fields (Banilower et al., 2013; Chen & Weko, 2009; National Research Council, 2013).

Sustainable WATERS is an interdisciplinary program that provides teacher training, supplies, and digital resources to improve watershed literacy in Southwest Florida (SWFL). The overarching goal of the project is to improve educators and students' watershed literacy through the use and building of models, leading to a greater knowledge of and sense of agency in creating solutions to the impacts of climate change in SWFL. Each lesson within the program has a clear scientific focus, opportunities for field or lab work, data analysis, and model building all related to the learner's own backyard. As part of Sustainable WATERS' teacher training, a professional development (PD) Communities of Practice (CoP) program was developed that focused on teacher development and understanding the local impacts of climate change in Title I middle schools. The program transitioned from in-person PD to an online format as a result of the COVID-19 pandemic. It leveraged access to virtual tools to increase teacher and student access to modeling and climate change expertise relevant to SWFL communities. Sustainable WATERS' CoP engaged teachers and students in locally-focused climate education by integrating models and modeling.

Communities of Practice (CoP) in Education

A CoP is a is a social learning system where individuals come together to fulfill individual and group goals of a common interest (Cambridge et al., 2005). CoPs focus on sharing best practices and creating new knowledge to advance a professional practice. Ongoing interactions are an important part of CoPs, and many virtual CoPs (vCoPs) rely on face-to-face meetings as well as virtual collaborative environments to communicate, connect, and conduct community activities (Cambridge et al., 2005). CoPs are social learning systems, where members define competence around a discipline or practice by combining three elements: a sense of joint enterprise, mutually defined norms and relationships, and a shared repertoire of communal resources they create and can draw upon to further their competence (Wenger, 2010). We used a CoP approach as a PD partnership model to connect university researchers and K-12 teachers. In this study, CoP serve as the primary theoretical framework, to examine how teachers engage in professional development in climate education. CoP provides a structured approach to understanding teacher learning as a social process, where participants develop expertise through interaction, collaboration, and sustained engagement within a professional learning community.

School-university partnerships provide opportunities for collaboration with mutual benefits (Lynch & Smith, 2012; White et al., 2010). The benefits associated with these partnerships include "built-in support networks" for the teachers (Darling-Hammond, 2006, p.110). However, challenges and barriers exist when implementing school-university partnerships, including sharing space, time, and resources required (Green et al., 2020). Existing connections to the community and school district partners in watershed education allowed for us to grow the CoP to deepen teachers' skill, content knowledge, and participation in watershed education.

Within CoPs, the members can participate at different levels and can move between levels of engagement throughout their participation. *Core* members define CoP norms and create and share knowledge. *Active* members frequently participate in the CoP but may not be leaders or creators of knowledge and artifacts. *Peripheral* members participate less frequently but can move to be active or core when they develop their knowledge and contribute to the CoP. Core members can legitimize peripheral members as they develop (Borzillo et al., 2011). Participation in a CoP enhances teachers' self-efficacy by providing opportunities for mentorship, collaboration, and real-world application of new instructional strategies. According to Bandura (1997), self-efficacy is shaped by mastery experiences, social persuasion, and observational learning, all of which occur naturally within a CoP. As teachers progress from peripheral to core members, their confidence in teaching climate-related content increases, reinforcing their belief in their ability to facilitate student learning effectively.

Climate Change Education

Science education communities advocate for a climate-literate public equipped with the scientific knowledge and skills needed to make informed decisions about global climate change (GCC) (McNeal et al., 2014). For this study, we define *climate literacy* as the ability to apply scientific knowledge to advance understanding and engagement in climate science (McNeal et al., 2014). However, climate change is inherently complex; the global nature of the issue makes it challenging to observe climate change at the local level, limiting its relevance to students' daily lives and the need for long-term analysis and projections makes it challenging for science educators to fully understand and effectively communicate the processes behind GCC (Nation & Feldman, 2022). Science educators recognize that teaching climate change science is necessary to produce a citizenry that understands the causes of

GCC and ways to both mitigate it and prepare for its effects (ChewHung, 2022; Gutierrez et al., 2008) Teachers must integrate GCC into their curriculum, and students need to develop a deeper understanding of its causes and impacts. Adding climate change content to existing science curricula, however, is not enough. Teachers require preparation through PD in effective pedagogical strategies to teach climate-focused content meaningfully (Nation & Feldman, 2021).

The Next Generation Science Standards (NGSS Lead States, 2013) emphasize the importance of analyzing evidence for climate change (HS-ESS3-1) and using climate models (HS-ESS2-6, HS-ESS3-5). However, studies have shown that educators, both at universities and K-12 schools, often lack confidence in their subject knowledge and feel unprepared to adequately teach climate change (Oversby, 2015). Filho and Hemstock (2019) argued that educational institutions should actively pursue initiatives that promote awareness and encourage local solutions. The Sustainable WATERS program aims to bridge this gap by engaging teachers and students through locally focused models and simplified climate modeling, fostering connection to the material and enhancing comprehension of GCC's complexity for beginners. Models can be powerful tools to help educators and students describe, represent, and predict climate phenomena (Cartier et al., 2001), though these models must often be simplified to illustrate climate change effects on a local or regional scale. Climate models are critical for scientists studying global climate trends. The program seeks to make these models more accessible for students and secondary science teachers, helping them understand the complex interactions associated with GCC. Research by Holthuis et al. (2014) indicated that instructional approaches focused on modeling climate data can improve both teaching effectiveness and student understanding of climate change. Additionally, Bhattacharya et al. (2020) found that students' ability to analyze complex climate science and climate literacy can improve when they use multiple modeling methods. In later sections, we describe the types of models used within the Sustainable WATERS curriculum and their integration with existing science standards.

Professional Development

The impact of PD on efficacy and student learning is well-documented (Althauser, 2015; Fischer et al., 2018; Rutherford et. al, 2017). PD is vital to help teachers gain skills and knowledge to teach about current environmental and social issues (Borko, 2004; Guskey, 2002). While many PD opportunities are available to science teachers, most are not designed specifically for teaching and learning of climate science or to advance teacher understanding of this complex issue (Schneider & Plasman, 2011). For complex issues, such as climate change, research suggests that educators need PD that presents content paired with specific teaching strategies to build confidence and better incorporate the topic into their curriculum (Hestness et al., 2017; Kunkle & Monroe, 2018; Plutzer et al., 2016). This specific type of PD can increase teachers' subject matter knowledge (SMK), pedagogical content knowledge (PCK), and self-efficacy to enable them to teach climate change more effectively (Nilsson, 2014; Van Driel & Berry, 2012). For the purpose of this study, self-efficacy is defined as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (Bandura, 1997, p. 3). PD programs can increase educators' self-efficacy (Bandura, 1997; Chesnut & Burley, 2015; Holden et al., 2011; Morris et al., 2017) and teachers with greater self-efficacy tend to be more open to new ideas and willing to experiment with new methods to meet the needs of their students (Gavora, 2010).

Li et al. (2021) document a gap in the literature, in that while many climate change education PD programs are implemented, little empirical evidence of effective PD approaches specific to climate change education have been documented. Desimone (2009) offers five critical features for successful PD, of which the following were implemented in the Sustainable WATERS teacher training experiences including:

1. **Focus on content**: Sustainable WATERS incorporated the use and creation of models to communicate and represent their understanding of the problematic trends associated with the impacts of climate change in SWFL
2. **Opportunities to engage in active learning between and among the participants**: teachers engage with each other, experts in the field, (including Marine and Environmental Scientist and GIS specialists) and local ecosystems to learn about the impacts of climate change in local watersheds and their home environments.
3. **Coherence between new learning and teachers' knowledge and beliefs, collective participation:** data collected via Climate Literacy survey and virtual check-ins determined the progression of each learner.
4. **Extending the PD over an appropriate duration of time**: Sustainable WATERS took place over one Academic Year (AY) and weekly check-in

As PD models continue to evolve, vCoPs have emerged as powerful tools for supporting teacher learning, collaboration, and instructional confidence (Ghamrawi, 2022; Schwarzhaupt et al., 2021). Building on this foundation, our project applied a CoP model specifically rooted in climate-related watershed issues, focused on teachers' self-efficacy and the impacts of climate change in SWFL through participation in PD designed in the Meaningful Watershed Educational Experiences (MWEE) framework employed by the National Oceanic and Atmospheric Administration (NOAA, 2021). This model responds to the growing demand for equitable, high quality, climate change PD that supports content knowledge and instructional confidence across diverse educational settings and can be readily adapted to other geographic regions by contextualizing climate impacts to their local ecosystems.

Research Questions

1. How does participation in the professional development community of practice affect teachers' climate literacy and self-efficacy in climate education?

2. What elements of a hybrid professional development program foster the development of a community of practice and how?

Materials and Methods

A mixed methods design was used to examine relationships between program participation, CoP engagement, and climate literacy through the data collected via surveys, interviews, and meeting notes. A mixed methods approach was chosen to allow for a comprehensive exploration of both quantitative trends and deeper contextual insights vial qualitative data that would not be possible with either method alone. Given that this study aimed to assess both objective measures (e.g., CoP engagement levels, climate literacy growth) and subjective experiences (e.g., teachers' perceptions of their participation and self-efficacy), a mixed methods approach was the most suitable. Specifically, this study employed a sequential explanatory mixed methods design (Creswell & Plano Clark, 2018), where quantitative data were collected first, followed by qualitative data to provide deeper insight into initial findings and identify patterns in CoP engagement and climate literacy through surveys, interviews, and meeting notes to contextualize patterns within teachers' experiences.

To elicit teachers' perceptions of their program participation and CoP membership, we used a phenomenographical approach (Marton, 1986). By using a phenomenographical approach, we gained insights into the different ways the teacher participants perceived and engaged with the program and CoP. A phenomenographic approach was selected to explore the diverse ways teachers experienced and interpreted their participation in the PD and CoP and allowed for a deeper understanding of variations in teachers' experiences, beliefs, and levels of engagement, contributing to a richer analysis of the impact of the PD program on teachers' perceptions and practices related to climate change education within Sustainable WATERS. We examined teachers' understanding of the impacts of climate change in SWFL, perceptions of teaching confidence, perceptions of their CoP membership, and their actual participation.

Sustainable WATERS took place over two years, beginning in Fall of 2021. This study focuses on the experience of the first cohort. Sustainable WATERS supported teachers within a large district, both in student population and geographically. Over half the students identify as economically disadvantaged. 38% of the student population identifies as Hispanic, 15% Black, and 43% White.

Participants

Teachers were recruited through noncompetitive selection and the district partnership dissemination of applications. Six STEM teachers applied to participate in the first cohort, thus, all were selected as participants in the CoP Teacher PD. All participants worked in Title I middle schools within the district. Because the group was small and selected through an invitational process, findings may not be generalizable to all teacher populations or contexts.

Intervention

The program supported the key parts of a CoP: working together, sharing goals, and building resources through the following: teachers worked together through virtual check-ins and in-person sessions, exchanging ideas, offering feedback, and reflecting on classroom implementation. The program centered around a shared goal of improving climate literacy instruction using the NOAA MWEE framework, fostering a common sense of purpose and direction. Participants also contributed to a growing set of tools and resources, which were shared and refined throughout the PD.

Teachers were selected to participate as teams for an entire school year, between PD and classroom implementation to foster long-term engagement in the program. They had weekly communication with teachers from other schools through field experiences and synchronous weekly virtual check-ins for collective participation. Each week was designed to take approximately 10 hours of the teachers' time. The 32-hour hybrid program, included the following elements (see supplemental materials):

In-person Kick-off: Teachers were provided supply kits for curriculum training, introduced to the program's outdoor activities on local beaches (surveying local beaches for the impacts of erosion) and classroom activities (hurricane dynamics).

Virtual instruction: Modules contained videos, text instruction, and models to support teachers' engagement in curriculum activities in the classroom and schoolyards. Each module focused on one of four major impacts of climate change in SWFL- habitat shift, increased extreme weather events, sea level rise, and saltwater intrusion- through field studies, data collection and analysis, and using and creating models. Each was developed through inquiry-based activities aligned with the NOAA MWEE framework, facilitating four activities for students and teachers: Issue Definition and Background Research, Outdoor Field Activities, Synthesis and Conclusions, and the execution of Stewardship Action Project. For a detailed examination of NOAA's MWEE framework see: https://www.noaa.gov/education/explainers/noaa-meaningful-watershed-educational-experience

Synchronous, virtual check-ins**:** Weekly one-hour check-ins provided facetime with project partners, time for sharing climate change education resources, successes and challenges with other teachers, and a platform for collaboration. Table 1 describes each climate-related module and the MWEE elements included to support learner-centered practices in climate change education.

Table 1:

Sustainable WATERS content and MWEE alignment

CLIMATE IMPACT	ISSUE DEFINING QUESTION	OUTDOOR FIELD ACTIVITY	MODELS FOR SYNTHESIS AND CONCLUSIONS	COMMUNITY ACTION ACTIVITIES
Habitat Shift	How are organism populations changing as climate changes?	Schoolyard plant surveys and measurement, with data sharing, phenology surveys	Spatial models of plant location and characteristics	Defined by teacher
Increased Extreme Weather	How does increased storminess impact our watershed dynamics?	Schoolyard elevation surveys: Map Your Watershed	Schoolyard topography maps to identify vulnerabilities	Defined by teacher
Sea Level Rise	How does sea level rise impact our watershed dynamics?	Schoolyard elevation surveys: Map Your Watershed	Spatial map of sea level rise scenarios; NOAA Sea Level Rise Simulator	Defined by teacher
Saltwater Intrusion	Why are our mangroves "walking" inland?	Schoolyard plant surveys and measurement, with data sharing, phenology surveys; schoolyard surface and groundwater quality analysis	Spatial models of plant location and characteristics and potential change; combined water quality data portal with other participating schools	Defined by teacher

Data Collection

To measure the participation in the CoP affecting teachers' climate literacy, teachers were surveyed via pre- and post-Climate Literacy Survey (see Appendix A). The participants were surveyed on GCC content knowledge and perceptions, their experience teaching climate topics, using models in their instruction, and self-efficacy teaching climate topics and using models in their instruction. Post-PD, they were asked their perceptions of PD effectiveness, recommended changes, and resources needed for effective curriculum implementation.

Teachers completed feedback surveys (see Appendix B) after each module. They provided their implementation plan, recommendations for improvement and best classroom practices, and reported on their implementation experience. All responses were anonymous to ensure protection of identities.

Weekly virtual check-ins were recorded, during which the program coordinator probed teachers' perceptions of their participation, challenges to participation, and how it impacted their classroom practice. They were asked about their perceptions of the hybrid format, and how it supported or challenged their PD. These recordings were transcribed and coded for analysis using interrater reliability among three researchers on the team.

Analysis

Recordings of weekly check-ins, informal meeting observations, and surveys were analyzed through thematic coding and the development of learning progressions to understand how the teachers' climate literacy and perceptions, self-efficacy, and CoP engagement developed throughout the PD. To analyze the qualitative data collected, we used the concept of learning progressions as a framework to guide our thinking about how the teachers' knowledge progressed over time (Schneider & Plasman, 2011). Applying the framework from Feldman et al. (2021), to document the progress of each participant over the course of the PD. Learning progressions are used to describe the process of how learning becomes increasingly sophisticated about a topic over time (Duschl et al., 2007; Heritage, 2008; Smith & Wiser, 2015). The use of teacher learning progressions helps illustrate the development of pedagogical and content knowledge, and role within CoPs related to the climate-centered PD.

To construct each progression, we examined years of experience, what they hoped to gain from the PD, change in climate literacy and self-efficacy, type of engagement in activities, level of implementation of curriculum, impact on practice, perception of the PD, and their role for their future participation in the program. We assessed the change in understanding of concepts and skills over time to construct the progressions as opposed to making a single summative assessment upon completion (Wilson, 2009). We mapped the progressions, constructed the progressions as grouped instances, and then reformulated them into narratives. This informed our understanding of the teachers' progressions as a trajectory of development rather than a series of discrete events (Heritage, 2008). Each teacher was evaluated as an active or peripheral member of the CoP, based on their participation (Baker & Beames, 2016). Inter-researcher reliability was ensured through consensus of the research team of each progression. Member checking occurred throughout via check-ins, interviews, and opportunities for feedback.

Results

The following themes were identified through the analysis:

1. Confidence and a result of increased understanding: Participants showed varying levels of initial knowledge and confidence in teaching climate change topics. Post-program, there was a noticeable increase in their climate content knowledge and confidence in teaching these topics effectively.
2. Perceptions of anthropogenic-induced climate change: Participants' beliefs about climate change evolved throughout the program, with most shifting towards a stronger belief that climate change is happening, caused by humans, and supported by scientific consensus. This shift also included increased concern about the impacts of climate change.
3. Impact on Teaching Practice: The program had a positive impact on participants' teaching practices. They reported feeling more prepared, using new teaching strategies, and integrating climate change topics effectively into their curriculum. However, some participants faced challenges in implementation due to time constraints or other barriers.

4. Role of Mentorship: Mentorship played a role in supporting participants' engagement and learning. Mentorship contributed to increased engagement and confidence among mentees.

The PD program yielded varied outcomes for participating teachers, highlighting differences in engagement, growth in climate literacy, shifts in climate change perceptions, and contributions to the CoP (see table 2). Participants entered the program with diverse teaching experiences and confidence levels regarding climate change instruction. The case studies examined below provide insights into how teacher engagement, prior experience, and active collaboration can influence the effectiveness of climate change education initiatives in professional development settings.

Table 2
Overview of Teacher Participants

Participant	Gender Identity	Years of Teaching Experience	Same School as Another Participant	CoP Participation Status	Pre-Test Score	Post-Test Score
Monica	Woman	16	Yes (with Rachel & Mark)	Active	69	77
Rachel	Woman	2	Yes (with Monica & Mark)	Active	62	77
Mark	Man	2	Yes (with Monica & Rachel)	Peripheral	62	62
Francine	Woman	4	No	Active	92	85
Georgina	Woman	7	Yes (with Ashley)	Peripheral	62	62
Ashley	Woman	1	Yes (with Georgina)	Peripheral	54	85
					M: 66.83 SD: 13.21	M: 74.67 SD: 10.44

Learning Progressions for Participants

Monica and Rachel (Mentor/Mentee) - taught at the same school. Monica had 16 years of teaching experience. Prior to participation, she implemented climate change curriculum with her students frequently and felt somewhat comfortable teaching those topics. Monica believed climate change was happening and caused by humans, that there was scientific consensus to support it, and was very concerned about the impacts.

Monica's climate content knowledge increased from 69% to 77%. Post-PD, she felt completely comfortable in her climate content knowledge and ability to teach climate topics. Monica maintained climate change was happening and caused by humans, there was scientific consensus to support it, and was very concerned about the impacts.

Monica participated in all opportunities for engagement, she attended the in-person kick-off day, all virtual weekly check-ins, completed the pre- and post-PD assessment, and all four requests for feedback during the program. She implemented lessons within one month of completion. Monica was a key contributor to community dialogue, she shared plans to implement activities, suggested improvements, commented on content accuracy, coached the team on technology barriers, requested clarification and material supply provision. She perceived herself as connected to the CoP, stating she engaged most through the weekly virtual check-ins. Monica was satisfied with the program and felt it was a success for her. She felt the program was well-organized, relevant to her classroom practice, and developed her teaching skills. She was evaluated as an *active* CoP member.

Rachel had two years of teaching experience. Prior to participation, she implemented climate change curriculum frequently and felt confident in her teaching ability for climate topics. Initially, Rachel was somewhat sure climate change was happening and caused by humans, that scientists disagreed about the phenomenon, and was somewhat concerned about the impacts. She hoped to "gain more hands-on activities to increase student engagement".

Rachel's climate content knowledge increased from 62% 77%. Post-PD, she felt completely comfortable in her climate content knowledge and ability to teach climate topics. Rachel's climate perceptions shifted from pre- to post-PD. Post-PD, Rachel believed climate change was happening and caused by humans, there was scientific consensus to support it, and was very concerned about the impacts.

Rachel also implemented lessons within one month of program completion. Rachel was a key contributor to community dialogue, she shared implementation plans, experienced implementing activities in her classroom, suggested activities that would enhance curriculum and classroom strategies, and perceived herself as connected to the CoP; stating weekly check-ins were most useful for connecting with the rest of the cohort. Rachel felt the program was well-organized, relevant to her classroom practice, and developed her teaching skills. She specifically requested a field experience for the students as a way of making local climate issues meaningful for and memorable to them. She was evaluated as an *active* CoP member.

Mark - was at the same school as Monica and Rachel and taught for two years. Prior to participation, he implemented climate change curriculum frequently and felt confident in his teaching ability on the topic. At the beginning of the program Mark was very sure climate change was happening, was caused by humans, and there was scientific consensus about the phenomenon but was not at all concerned about the impacts. Through the PD, Mark hoped to integrate more environmental science projects into his curriculum and to deepen students' knowledge of environmental issues and stewardship.

Mark's climate content knowledge remained the same, at 62%. Post-PD, he felt completely comfortable in his climate content knowledge and his ability to teach climate topics. Mark's climate concern shifted from not at all concerned to very concerned.

Mark taught at the same school as Monica and Rachel. However, he did not collaborate, share supplies, or participate in a peer mentor relationship. He was reserved and would often "see how went with their implementation" before fully implementing the curriculum. Mark only went to two of the PD sessions and largely participated as an observer with limited contributions to the greater community. While being an enthusiastic member of the community, he didn't actually complete any of the modules and had limited responses to emails, check-ins, and has yet to implement any parts of the curriculum with his students. Mark found the timing of the program to be difficult for his students due to the testing schedule his students were participating in. That said, all participating members experienced the same testing period within the same school district.

Mark perceived himself as connected to the CoP but could not describe how he interacted with the community. He had no plans for implementation, but stated the other CoP teachers at his school were developing a plan and a timeline. Mark was evaluated as a *peripheral* CoP member.

Francine - was the third *active* member of the CoP, while not as central as Monica and Rachel to the community, she maintained active participation over the course of the semester. Francine had four years prior teaching experience. Prior to the PD, Francine frequently taught climate change and felt confident teaching concepts of GCC. She hoped to gain ways to incorporate the 5E model with climate change content from the PD. Francine participated in university-led PD two years prior.

Francine's climate content knowledge decreased from 92% correct on a content assessment to 85% correct. Post-PD, she felt completely comfortable in her climate content knowledge and her ability to teach climate topics. Francine maintained climate change was happening and caused by humans, there was scientific consensus to support it, and was very concerned about the impacts.

She actively participated in the PD but went to one less session than the other *active* members of the group. When she did participate, she was able to share new insights, and new resources she created related to the curriculum with the rest of the group. In one instance, during a discussion of how to incorporate mangroves with life science, she created her own photosynthesis game and sent a photo to the rest of the CoP. She was also the only member of the PD who was considering an Environmental Action Plan with her students for Earth Day. While she wasn't able to complete it, she did have initiative to use the PD to inform her practice in real-time.

Francine responded "I don't know" when asked if she was connected to the CoP, and also when asked to describe how she interacted with the community. Despite her engagement level in the community, Francine felt only somewhat prepared to implement the curriculum in class. Post-PD feedback revealed she only somewhat agreed the program supported her PD in this content. The number of activities and the limited amount of time were her primary barriers. Nonetheless, Francine was evaluated as an *active* CoP member.

Georgina and Ashley (Mentor/Mentee) **-** participated from the same school. Pre-PD, Georgina infrequently taught climate change in her class and felt "neutral" in her comfort level teaching GCC. She was somewhat sure climate change was happening and was caused by humans, felt there was disagreement among scientists about climate issues, and was somewhat worried about the phenomenon. She hoped to gain "useful classroom resources to engage students in real life experiences."

Georgina's climate content knowledge remained constant at 62%. Post-PD, she felt completely comfortable in her climate content knowledge and her ability to teach climate topics. Georgina's perceptions of climate changed from pre- to post-PD: she was sure climate change was happening and caused by humans, there was scientific consensus to support it, and was very concerned about the impacts.

While Georgina did have seven years teaching experience, her time was split between teaching science and technology, so she did not have the opportunity to practice the implementation of the curriculum as much as others. Georgina perceived herself as a CoP member and stated she participated by sending emails asking questions, discussing failures and successes, sharing information, attending meetings. She served as an informal mentor to Ashley. Based on her actual participation however, she was evaluated as a *peripheral* CoP member.

Ashley was a new teacher, pre-PD, Ashley never taught climate change subjects in her class and felt "neutral" in her comfort level teaching them. At the beginning of the program Ashley believed climate change was happening and caused by humans, there was scientific consensus to support it, and was very concerned about the impacts. She hoped to gain expanded knowledge on climate change and new ways to incorporate real life situations in the classroom.

Ashley's climate content knowledge increased from 54% 85%. She maintained her perceptions on climate change as happening, important, and human caused. Post-PD, she felt completely comfortable in her climate content knowledge and in her ability to teach climate topics.

one meeting, she was able to document her experience implementing the sea level rise module and give feedback to the rest of the community, particularly timing tips. However, her participation with the rest of the PD beyond that meeting was limited. She did not attend half of the virtual check-ins, and did not implement the rest of the curriculum beyond the sea level rise module. Ashley perceived herself as a CoP participant, stating teachers in the cohort "were all in the same boat with students." Based on her actual participation however, Ashley was evaluated as a *peripheral* CoP member.

Discussion

The results of the program reveal diverse outcomes among the six participating teachers, influenced by variations in experience, climate knowledge, and levels of engagement. There were distinct differences in actual CoP participation level that divided the group. Three teachers (Monica, Rachel, Francine) were *active* participants and three were *peripheral* (Mark, Georgina, Ashley), based on their PD completion, virtual meeting attendance and level of participation, survey responses, and program implementation. Monica and Rachel, highly active in the CoP, saw meaningful gains in climate content knowledge, increased confidence, and readiness to implement the curriculum. Both reported feeling deeply connected to the CoP, actively contributing ideas and resources. In contrast, Mark and Ashley, who engaged minimally, showed limited progress; they were evaluated as peripheral members and faced challenges in curriculum implementation. Francine and Georgina had moderate participation and displayed steady, though varied, impacts on their teaching. Francine maintained high engagement, but felt only somewhat prepared to teach the curriculum, Georgina balanced her role in the CoP with limited classroom implementation.

None of the participants were evaluated as *core* CoP members. Core members typically plan, coordinate, and lead other members to engage them in the CoP shared enterprise (Borzillo et al., 2011). Sustainable WATERS, was a new university-school partnership, *core* CoP membership was catalyzed through a university program coordinator, who maintained constant contact and support for new CoP members. This was essential to program success, however, for long-term impact, core participation from teachers is necessary. In the future, there should be a focus on how *active* members can move to the *core* of periphery as they gain experience with the curriculum.

Findings suggest teachers' actual participation with the program did not align with perceptions of CoP participation. All teachers, except Francine, felt connected to the CoP. It should be noted Francine was the only teacher participant who did not have a peer teacher at her school. While Mark felt connected, he could not describe what he did to participate, and while he did have peer teachers at his school, he did not collaborate with them as often as they collaborated with each other.

Monica and Rachel were active CoP members; both demonstrated an increase in their climate content knowledge over the course of the PD. Francine was an *active* member of the CoP, but did not perceive herself that way. She was the only participant to demonstrate a decrease in content knowledge over the course of the program. *Peripheral* members, Mark and Georgina, demonstrated no change in their climate content knowledge. Ashley, another *peripheral* member, had the largest increase of the CoP participants. However, it should be noted, she began the program with the lowest score. Examining the progression of individuals, and as a whole, we suggest both actual and perceived participation in a CoP can affect development of content knowledge over time. This can have future impact on the design of virtual environments and potential research questions - which are most likely to support actual CoP participation, and which are most likely to foster a *perception* of connectedness?

Active CoP members, Monica and Francine's climate perceptions were both considered alarmist and anthropogenic induced prior to the PD. Rachel did not begin the program as concerned, although completed it that way. Her mentor/mentee relationship with Monica may have contributed to the change (McCauley & Guthrie, 2007). The relationship within the program highlights the impact of school-based teacher teams participating. According to Vescio et al., (2008) participants are more likely to persist and contribute to CoPs through co-learning and collaboration when participating with other teachers from their home school. Our findings suggest that while vCoPs provided an essential platform for continuity, virtual meetings present challenges in forming peer connections. This aligns with Jocius et al. (2022), who found that face-to-face interactions create more opportunities for spontaneous collaboration and relationship-building. Future CoPs should prioritize hybrid models that blend the flexibility of virtual engagement with the relationship-building benefits of in-person collaboration that are particularly important to novice teachers.

There was no discernible pattern to changes in self-efficacy related to CoP membership, perceived or actual among the participants. Pre-PD, two teachers (one active, one peripheral) reported high levels of confidence teaching with models and teaching climate change topics. Four teachers (two active, two peripheral) reported medium levels of confidence. Post-participation, all participants reported high levels of confidence teaching GCC, aligning with previous studies indicating the use of vCoPs for in-service teacher PD can increase self-efficacy through increased opportunity for social networking, collaboration, and overcoming barriers typical to implementation of in-person PD (Boling & Martin, 2005; Kirschner & Lai, 2007; Moore & Barab, 2002). Sustainable WATERS virtual and in-person interactions supported these practices and positively affected self-efficacies for all participants.

Years of teaching experience played an important role in shaping CoP participation and its impact on climate literacy. More experienced teachers, such as Monica, demonstrated greater confidence in engaging with the CoP, likely due to her prior pedagogical expertise and familiarity with PD settings. Conversely, early-career teachers such as Ashley and Mark often remained in peripheral roles, citing uncertainty in both climate content knowledge and instructional strategies. These findings suggest scaffold PD, including mentorship or differentiated pathways, may help support early-career teachers fully integrating into CoPs.

Community Participation and Virtual Tools

All teachers participated in the in-person kick off day and completed pre- and post-PD surveys. Only Monica completed all module feedback surveys. Weekly virtual check-in participation matched overall CoP participation: *active* members attended most frequently and contributed most to the conversation. *Peripheral* members attended 50% of the meetings and were less engaged during their attendance. For successful CoP, members develop their own ways of contributing and mechanisms for CoP development outside of program coordination.

Participants described their participation in the CoP through discussions with other teachers, collaborative planning or implementation, and virtual check-ins, which was the mechanism for communication and collaboration. No one described CoP interaction beyond the university team set up. Therefore, none of the teachers were evaluated as *core* members based on the literature. Because core members are typically schedulers and coordinators, we situated the core position and associated responsibilities within the university program coordinator. The expectation is as *active* members participation deepens; they will become *core* members of the CoP. Future follow-up with participants is needed to determine if this occurred after the completion of the PD.

All interactions described by teachers were ones in which they received immediate feedback and acknowledged their contribution in real time. Ekici (2018) found online CoPs boosted self-efficacy as participants were able to compare their experiences to others and recognize their problems and struggles were similar to what others experienced; similar to in-person CoP development, participants typically report meetings, curriculum training, and social events as most impactful to their belonging (Fernández et al., 2003; Lee, 2008; Puchner & Taylor, 2006). This suggests leveraging virtual tools that mimic in-person interactions and provide immediate feedback will have the greatest positive impact on CoP development. To achieve this, when survey tools are used to shape resources, practices, and norms, contributors should see the result of their feedback in community resources immediately.

Previous research supports the importance of groups of teachers from the same schools participating together for increased persistence and incorporation of PD into practice. vCoPs can increase teachers' self-efficacy by connecting novice and veteran teachers who may not otherwise get a chance to collaborate (Ghamrawi, 2022; Lieberman et al., 2011; Schwarzhaupt et al., 2021). Our program aimed to do this by partnering teachers from the same schools, however, future recruitment efforts should ensure that all participants have support from peer teachers. Mark and Francine were not partnered with another teacher, and were the only participants who could not describe the ways in which they interacted with the group. Georgina and Ashley were both evaluated as peripheral members, while Monica and Rachel were both identified as active. These relationships suggest that a strong mentor-mentee relationship can help facilitate growth and development of a new teacher and lead to active participation.

This study offers additional insights for designing hybrid vCoPs, particularly in the post-pandemic era, building on the work of Ghamrawi (2022). Variability in knowledge gains, engagement levels, and perceived CoP participation suggests that virtual and hybrid PDs must be designed to actively bridge participation gaps and tailor support for different participants.

Conclusion

This study explores the complex dynamics of climate literacy, and climate participation within a CoP framework. As with previous research, the differences in both actual and perceived CoP participation levels among the teacher participants, had implications for their climate content knowledge, self-efficacy, and engagement. Core CoP membership, characterized by planning, coordination, and leadership within the community, was facilitated primarily through the program coordinator, thus, crucial to fostering sustained engagement among future PD (Baker & Beames, 2016; Borzillo et al., 2011; Wenger et al., 2002). However, for long-term impact of climate change PD programs, it is clear that for teachers to transition from peripheral to core participation roles as they gain experience with the curriculum, ongoing support and mentorship is necessary, and PD programs should consider interventions that can bridge this gap, including pairing teachers from the same school to build in-person peer support, incorporating structured mentorship to guide early-career teachers, and leveraging hybrid models that blend virtual flexibility with school-based collaboration.

While most of the participants within the study reported feeling connected to the CoP, lack of perceived connectedness may be influenced by isolation from peer teachers at her school. This further supports the need for peer collaboration in fostering a sense of belonging and active engagement within CoPs, aligning with previous research and the importance of school-based teacher teams in sustaining participation.

The study revealed variations in climate content knowledge development among participants, with active members demonstrating increases in knowledge while some peripheral members showed no change or even decreases. These findings suggest differences between actual and perceived participation in shaping learning outcomes within CoPs. Additionally, mentor-mentee relationships highlight the potential for peer support to influence participants' climate perceptions and engagement levels, suggesting the need for structured support mechanisms within CoPs. While the development of mentor/mentee relationships was not intentional in the recruitment process, they were impactful on engagement. Future design should consider a nested CoP structure, in which each school has a predetermined core or active teacher to draw other peripheral teachers in and provide on-site support to positively influence teacher participation. As the program norms and practices evolve, program leaders should emphasize in-person collaboration within school environments and support expert teachers within a group to share expertise.

The role of virtual tools in facilitating CoP with virtual check-ins served as the primary mechanism for communication and collaboration among participants, the COVID-19 pandemic bolstered the case for virtual CoPs. Jocius et al. () uncovered unique impacts of online CoPs when the pandemic forced them to shift their face-to-face teacher PD to a virtual platform. Previous in-person sessions had high levels of engagement, so they worried "switching to a virtual experience might limit opportunities for community building" (p. 11). We had similar concerns as Sustainable WATERS was forced to shift modes. We found that interactions were initiated primarily by the university team, indicating a need for interventions to foster deeper engagement and collaboration beyond this framework. Leveraging virtual tools that mimic in-person interactions and provide immediate feedback may enhance CoP development and support participants' sense of belonging and self-efficacy.

Our findings suggest the connection between hybrid CoP participation, climate literacy, and teacher self-efficacy are useful tools. However, more research is needed to define how and with what specific tools this is fostered. That said, what happens at schools as they interact in-person may be more impactful than the CoP, as teachers' perceived CoP participation was aligned with school team participation. CoP impact hinges on teachers' abilities to contribute to and receive feedback in real time, therefore, hybrid and virtual programs should be designed with tools that enable that type of interaction.

Future Research

Overall, the study outlines the complex interactions between participation and climate literacy within CoPs. Future research should examine strategies for promoting active participation and mentorship within CoPs, as well as the effectiveness of virtual tools in fostering collaboration and knowledge sharing among participants. Additionally, efforts should be made to incorporate peer support structures and on-site collaboration within school environments to enhance teacher engagement and learning outcomes within CoPs.

Declaration of Interest Statement

No potential conflict of interest was reported by the authors.

Data availability statement

This study was approved by the University's Human Subjects' IRB. The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to restrictions as information could compromise the privacy of research participants.

Supplemental online materials

https://sites.google.com/view/sustainablewaters

Appendix A:

Pre-Assessment: https://forms.gle/QEzaacGDmdwoTmZZ8

Appendix B:

Post-Assessment: https://forms.gle/8hJzMPg43nGTsiwD7

Funding

This work was supported by the NOAA's Gulf of Mexico B-WET Program under Grant number AWD-000271 -GR00348

Molly Nation (mnation@fgcu.edu) is a professional at the intersection of science, policy, and education. Molly served as an AAAS Science and Technology Policy Fellow in Washington, DC from 2022-2024, contributing her expertise to the Millenium Challenge Corporation as a Climate Change and Nature-based Solutions Advisor and as a Translational Climate Science Fellow at the EPA Office of Research and Development. Molly is an Associate Professor at the Florida Gulf Coast University, specializing in Environmental Education within the Department of Ecology and Environmental Studies.

Heather Skaza Acosta (hskaza-acosta@fgcu.edu) is the Director of the Whitaker Institute for STEM Education at Florida Gulf Coast University and an Associate Professor of Environmental Education. She has extensive experience teaching and developing curriculum to support the next generation of environmental educators. Her teaching and research focuses on the impact of STEM field experiences on the teaching practices of formal and non-formal educators.

References

Althauser, K. (2015). Job-embedded professional development: Its impact on teacher self-efficacy and student performance. *Teacher Development*, *19*(2), 210-225. https://doi.org/10.1080/13664530.2015.1011346

Bhattacharya, D., Steward, K. C., Chandler, M., & Forbes, C. (2020). Using climate models to learn about global climate change. *The Science Teacher*, *88*(1), 58-66. https://doi.org/10.1080/00368555.2020.12293558

Baker, A. & Beames, S. (2016). Good CoP: What Makes a Community of Practice Successful? *Journal of Learning Design*, *9*(1), 72-79. https://doi.org/10.5204/jld.v9i1.234

Bandura, A. (1997). Self-efficacy the exercise of control. New York: *H. Freeman & Co. Student Success*, *333*, 48461. https://doi.org/10.1891/0889-8391.13.2.158

Banilower, E. R., Smith, P. S., Weiss, I. R., Malzahn, K. A., Campbell, K. M., & Weis, A. M. (2013). *Report of the 2012 national survey of science and mathematics education*. Horizon Research, Inc.

Borko, H. (2004). Professional development and teacher training: Mapping the terrain. *Educational Researcher, 33(*8), 3–15. https://doi.org/10.3102/0013189x033008003

Boling, C. J. & Martin, S. H. (2005). Supporting teacher change through online professional development. *The Journal of Educators Online*, *2*(1), 1–15. https://doi.org/10.9743/jeo.2005.1.1

Borzillo, S., Aznar, S., & Schmitt, A. (2011). A journey through communities of practice: How and why members move from the periphery to the core. *European Management Journal*, *29*(1), 25-42. https://doi.org/10.1016/j.emj.2010.08.004

Cambridge, D., Kaplan, S., & Suter, V. (2005). Community of practice design guide: A step-by-step guide for designing & cultivating communities of practice in higher education. *EDUCAUSE Learning Initiative (ELI)*, *2*.

Cartier, J., Rudolph, J., & Stewart, J. (2001). The nature and structure of scientific models.

Chen, X. & Weko, T. (2009). Students who study science. *Technology and Engineering*.

Creswell, J. W. & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3rd ed). SAGE.

Darling-Hammond, L. (2006). *Powerful teacher education: Lessons from exemplary programs* (1st ed.). San Francisco, CA: Jossey-Bass. https://doi.org/10.15581/004.12.25329

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational researcher*, *38*(3), 181-199. https://doi.org/10.3102/0013189x08331140

Duschl, R. A., Schweingruber, H. A., Shouse, A. W., & National Research Council. (2007). Committee on Science Learning Kindergarten Through Eighth Grade. *National Research Council (US)*.

Feldman, A., Nation, M., & Laux, K. (2021). The effects of extended action research-based professional development on the teaching of climate science. *Educational Action Research*, 1-17. https://doi.org/10.1080/09650792.2021.1981417

Fischer, C., Fishman, B., Dede, C., Eisenkraft, A., Frumin, K., Foster, B., Lawrenz, F., Levy, A. B., & McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education*, *72*, 107-121. https://doi.org/10.1016/j.tate.2018.02.011

Gavora, P. (2010). Slovak pre-service teacher self-efficacy: Theoretical and research considerations. *The New Educational Review, 21*, 17–30. https://doi.org/10.15804/tner.10.21.2.01

Ghamrawi, N. (2022). Teachers' virtual communities of practice: A strong response in times of crisis or just another fad?. *Education and Information Technologies, 27*(5), 5889-5915.

Green, C. A., Tindall-Ford, S. K., & Eady, M. J. (2020). School-university partnerships in Australia: A systematic literature review. *Asia-Pacific Journal of Teacher Education*, *48*(4), 403-435. https://doi.org/10.1080/1359866x.2019.1651822

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching, 8*(3), 381–391. https://doi.org/10.1080/135406002100000512

Heritage, M. 2008. Learning Progressions: Supporting Instruction and Formative Assessment. Washington, DC: National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Hestness, E., McGinnis, J. R., Breslyn, W., McDonald, R. C., & Mouza, C. (2017). Examining science educators' perspectives on learning progressions in a climate change education professional development program. *Journal of Science Teacher Education*, *28*(3), 250-274. https://doi.org/10.1080/1046560x.2017.1302728

Holden, M. E., Groulx, J., Bloom, M. A., & Weinburgh, M. H. (2011). Assessing teacher self-efficacy through an outdoor professional development experience. *The Electronic Journal for Research in Science & Mathematics Education*, *15*(2).

Holthuis, N., Lotan, R., Saltzman, J., Mastrandrea, M., & Wild, A. (2014). Supporting and understanding students' epistemological discourse about climate change. Journal of Geoscience Education, 62, 374–387. https://doi.org/10.5408/13-036.1

Jocius, R., O'Byrne, W. I., Albert, J., Joshi, D., Blanton, M., Robinson, R., ... & Catete, V. (2022). Building a virtual community of practice: Teacher learning for computational thinking infusion. *TechTrends*, *66*(3), 547-559. https://doi.org/10.1007/s11528-022-00729-6

Jones, A. L., & Stapleton, M. K. (2017). 1.2 million kids and counting—Mobile science laboratories drive student interest in STEM. *PLoS biology*, *15*(5), e2001692. https://doi.org/10.1371/journal.pbio.2001692

Kirschner, P. A., & Lai, K. W. (2007). Online communities of practice in education. *Technology, Pedagogy and Education*, 16(2), 127–131. https://doi.org/10.1080/14759390701406737

Kunkle, K. A., & Monroe, M. C. (2018). Cultural cognition and climate change education in the U.S.: Why consensus is not enough. Environmental Education Research. https://doi.org/10.1080/13504622.2018.1465893

Leal Filho, W., & Hemstock, S. L. (2019). Climate change education: An overview of international trends and the need for action. Climate change and the role of education, 1-17. https://doi.org/10.1007/978-3-030-32898-6_1

Lee, J. F. (2008). A Hong Kong case of lesson study—Benefits and concerns. *Teaching and teacher education*, *24*(5), 1115-1124. https://doi.org/10.1016/j.tate.2007.10.007

Li, C. J., Monroe, M. C., Oxarart, A., & Ritchie, T. (2021). Building teachers' self-efficacy in teaching about climate change through educative curriculum and professional development. Applied Environmental Education & Communication, 20(1), 34-48. https://doi.org/10.1080/1533015x.2019.1617806

Lieberman, A., Miller, L., Wiedrick, J., & von Frank, V. (2011). Learning communities: The starting point for professional learning is in schools and classrooms. *The Learning Professional*, 32(4), 16.

Lynch, D., & Smith, R. (2012). Teacher education partnerships: An Australian research-based perspective. *Australian Journal of Teacher Education, 37*(11), 132-146. https://doi.org/10.14221/ajte.2012v37n11.7

Marton, F. (1986). Phenomenography—a research approach to investigating different understandings of reality. *Journal of thought*, 28-49.

McCauley, C. D., & Guthrie, V. A. (2007). Designing relationships for learning into leader development programs. In B. R. Ragins & K. E. Kram (Eds.), The handbook of mentoring at work: Theory, research, and practice (pp. 573-592). Thousand Oaks, CA: SAGE. https://doi.org/10.4135/9781412976619.n23

McNeal, K. S., K. St. John, & S. B. Sullivan. 2014. "Introduction to the Theme: Outcomes of Climate Literacy Efforts (Part 1)." Journal of Geoscience Education 62 (3): 291–295. doi:10.5408/1089-9995-62.3.291. https://doi.org/10.5408/1089-9995-62.3.291

Moore, J., & Barab, S. (2002). The inquiry learning forum: A community of practice approach to online professional development. *TechTrends*, 46(3), 44–49. https://doi.org/10.1007/bf02784841

Morris, D. B., Usher, E. L., & Chen, J. A. (2017). Reconceptualizing the sources of teaching self-efficacy: A critical review of emerging literature. *Educational psychology review*, *29*, 795-833. https://doi.org/10.1007/s10648-016-9378-y

Munn, M., Griswold, J., Starks, H., Fullerton, S., Viernes, C., Sipe, T., Brown, M., Dwight, C., Knuth, R., & Levias, S. (2018). Celebrating STEM in Rural Communities: A Model for an Inclusive Science and Engineering Fesitval. *Journal of STEM Outreach*, *1*(1), 1-11.

Nation, M. T., & Feldman, A. (2022). Climate change and political controversy in the science classroom: How teachers' beliefs influence instruction. *Science & education*, *31*(6), 1567-1583. https://doi.org/10.1007/s11191-022-00330-6

Nation, M. T., & Feldman, A. (2021). Environmental education in the secondary science classroom: How teachers' beliefs influence their instruction of climate change. *Journal of Science Teacher Education*, *32*(5), 481-499. https://doi.org/10.1080/1046560x.2020.1854968

National Research Council. (2013). *Monitoring progress toward successful K-12 STEM education: A nation advancing?*. National Academies Press. https://doi.org/10.17226/13509

NGSS Lead States. (2013). Next generation science standards: For states, by states. Washington, DC: National Academies Press.

Nilsson, P. (2014). When teaching makes a difference: Developing science teachers' pedagogical content knowledge through learning study. International Journal of Science Education, 36(11), 1794–1814. https://doi.org/10.1080/09500693.2013.879621

NOAA (2021). NOAA Meaningful Watershed Educational Experience https://www.noaa.gov/education/explainers/noaa-meaningful-watershed-educational-experience

Oversby, J. (2015). Teachers' learning about climate change education. *Procedia-Social and Behavioral Sciences*, *167*, 23-27. https://doi.org/10.1016/j.sbspro.2014.12.637

Plutzer, E., McCaffrey, M., Hannah, A. L., Rosenau, J., Berbeco, M., & Reid, A. H. (2016). Climate confusion among US teachers. Science, 351(6274), 664–665. https://doi.org/10.1126/science.aab3907

Puchner, L. D., & Taylor, A. R. (2006). Lesson study, collaboration and teacher efficacy: Stories from two school-based math lesson study groups. *Teaching and teacher education*, *22*(7), 922-934. https://doi.org/10.1016/j.tate.2006.04.011

Rutherford, T., Long, J. J., & Farkas, G. (2017). Teacher value for professional development, self-efficacy, and student outcomes within a digital mathematics intervention. *Contemporary Educational Psychology*, *51*, 22-36. https://doi.org/10.1016/j.cedpsych.2017.05.005

Schneider, R. M., and K. Plasman. 2011. "Science Teacher Learning Progressions: A Review of Science Teachers' Pedagogical Content Knowledge Development." Review of Educational Research 81 (4): 530–565. doi:10.3102/0034654311423382.

Schwarzhaupt, R., Liu, F., Wilson, J., Lee, F., & Rasberry, M. (2021). Teachers' Engagement and Self-Efficacy in a PK–12 Computer Science Teacher Virtual Community of Practice. *Journal of Computer Science Integration*, 4(1): 1, pp. 1–14. DOI: https://doi.org/10.26716/

Smith, C. L., and M. Wiser. 2015. "On the Importance of Epistemology–disciplinary Core Concept Interactions in LPs." Science Education 99 (3): 417–423. doi:10.1002/sce.21166.

Van Driel, J. H., & Berry, A. (2012). Teacher professional development focusing on pedagogical content knowledge. Educational Researcher, 41(1), 26–28. https://doi.org/10.3102/0013189x11431010

Vescio, V., Ross, D., & Adams, A. (2008). A review of research on the impact of professional learning communities on teaching practice and student learning. *Teaching and teacher education*, *24*(1), 80-91. https://doi.org/10.1016/j.tate.2007.01.004

Wenger, E. (2010). Communities of practice and social learning systems: the career of a concept. In *Social learning systems and communities of practice* (pp. 179-198). Springer, London. https://doi.org/10.1007/978-1-84996-133-2_11

Wenger, E., McDermott, R. A., & Snyder, W. (2002). *Cultivating communities of practice: A guide to managing knowledge*. Harvard business press. https://doi.org/10.1108/bl.2002.17015bae.001

White, S., Bloomfield, D., & Le Cornu, R. (2010). Professional experience in new times: Issues and responses to a changing education landscape. *Asia-Pacific Journal of Teacher Education, 38*(3), 181-193. https://doi.org/10.1080/1359866x.2010.493297

Wilson, M. (2009). Measuring progressions: Assessment structures underlying a learning progression. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching, 46*(6), 716-730. https://doi.org/10.1002/tea.20318

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 29, NO. 4, 81-101

Incorporating Participatory Science in Elementary Schools: Teacher and Student Experiences with Outdoor Learning

Sarah J Carrier
North Carolina State University

Danielle R. Scharen
Horizon Research, Inc.

Meredith L. Hayes
Horizon Research, Inc.

P. Sean Smith
Horizon Research, Inc.

Christine Goforth
North Carolina Museum of Natural Sciences

Laura Craven
Horizon Research, Inc.

Lindsey Sachs
Horizon Research, Inc.

ABSTRACT

Science instruction in elementary school provides a base for student understanding of the natural world, yet policies prioritizing mathematics and reading have marginalized science. In response, some teachers have enhanced their science instruction by introducing students to participatory science (PS) projects. Using data from a larger study that examines the development of educative support materials for two existing PS projects, this embedded mixed methods study focuses on teachers' and students' experiences learning outdoors. We compare teachers' weekly log data, surveys, interviews, observations, and student focus groups to document teachers' applications of PS in their science classrooms and outdoors. Teachers report benefits (e.g., purposeful science learning) and challenges (e.g., time constraints, testing pressure) of implementing outdoor PS projects. Teacher and student data document cognitive and affective benefits of students' participation. Implications support the potential for PS projects that include schoolyard activities to supplement elementary science teaching and learning.

Keywords: participatory science, citizen science, elementary school, outdoor education, schoolyard

Introduction

Science instruction in elementary school provides a base for students understanding of the natural world and prepares students for future learning (Appleton, 2013; Curran & Kellogg, 2016). Despite the benefits of early science learning, accountability policies emphasizing mathematics and reading in elementary classrooms have marginalized science instruction (Banilower et al., 2018; Plumley, 2019). In response, some teachers have chosen to enhance their science instruction by introducing students to participatory science (PS) projects, where students have an opportunity to engage in real-world projects as they collect and make sense of the data (Jones et al., 2012; O'Donnell, 2023). The term *citizen science* has been widely debated in the United States, where the study took place; however, we have chosen to use the term *participatory science* for this manuscript to identify the practice of non-professional scientists collecting and contributing scientific data. Our use of participatory science aligns with the primary organization in the United States that recently changed its name from *Citizen Science Association* to *Association for the Advancement of Participatory Sciences*. *Community science* is another term that is sometimes used. In school-based participatory science, students can learn science content, and another important benefit of student engagement in PS is the potential to engage in learning outdoors (Carrier et al., 2013; Szczytko et al., 2018; Feille, 2021; Shume & Blatt, 2019), connecting students with the natural world outside their classroom doors.

In this study, we present data from a larger research project that examined teachers' and students' experiences with PS projects that incorporate outdoor learning experiences. Our research team prepared educative curriculum support materials that are designed to support teacher and student learning (Arias et al., 2016; Davis et al., 2017) for two existing PS projects: Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS) and Lost Ladybug Project (LLP).

We collected data from two groups of teachers who were asked to teach both participatory science projects. Teachers who had access to support materials for the CoCoRaHS project were identified as the CoCoRaHS treatment group, and teachers who had access to LLP support materials were the LLP treatment group. Both groups served as control groups for the project for which they did not have support materials. The key goals of this larger research study were to document whether and how the support materials contribute to teachers' and students' engagement in the projects. In addition to promoting student data collection and sense-making activities, our support materials included opportunities for teachers to expand teaching and learning beyond the four walls of the classroom to the schoolyard. In the present study, we focus on the outdoor learning experiences of teachers and students in our study, and our research questions ask:

1. How do teachers describe their implementation of PS project outdoor instruction?
2. What are teachers' views of their students' experiences with PS in the outdoors?
3. How do students describe their outdoor experiences?

Literature Review

Science Education in Elementary Schools

The elementary school years are a critical time in children's development that build the foundation for students' future learning, and science instruction offers multiple opportunities to connect students' school experiences and their lives outside of school (Irish & Kang, 2018). As teachers seek to provide their students with authentic science experiences, PS projects can offer students opportunities for engaging with their science instruction (Jones et al., 2012).

Despite the importance of early science instruction, policies in recent decades that align school accountability with reading and mathematics testing have resulted in a marginalization of science instruction time and resources in many elementary schools (Banilower et al., 2018; Plumley, 2019).

When the Next Generation Science Standards (NGSS Lead States, 2013) were released in 2013, the primary goals were to engage students in learning about science content and practices by doing science. While 20 states have adopted the standards, others have attempted to adapt their state's standards to align with the *Framework for K-12 Science Education* (National Research Council [NRC], 2012), and at this writing, one state has done neither (National Science Teaching Association [NSTA], 2023). Despite these efforts, few elementary science curriculum materials are aligned to the NGSS (Explore Reports, n.d.; Lowell et al., 2021), often leaving school districts and teachers scrambling to find instructional resources for science. In fact, elementary teachers most frequently create their own instructional materials (Doan et al., 2022). In the context of teachers seeking support for their science, we explore the potential for teacher and student engagement in science and the outdoors through PS projects.

Participatory Science

Participatory science (PS) has been described as the public's participation in science by contributing to the research of professional scientists through data collection and sharing (Bonney et al., 2016). In addition to engaging the public in science research, another goal of PS has been to increase the quantities of data far beyond what can be collected by professional scientists.

Science education reform efforts encourage a shift from teacher-centered to student-centered classrooms, and PS has the potential to help teachers learn to design instruction that centers on students and includes asking questions, data collection, and making sense of data (Shah & Martinez, 2016). While there is limited research on supporting teachers to include PS data collection in schools, PS offers dual potential for increasing the quantity of PS data while enhancing the science education of young learners.

When school-based PS projects connect with existing education standards (Lucky et al., 2014), they can be woven into school activities rather than added on as separate instruction. Although few PS projects include specific supports for teachers, the blending of PS in formal education can help students to envision themselves as contributors to, and part of, the larger science community (Esch et al., 2020). PS projects have also been found to motivate students (Dunn et al., 2016), and nature-based PS can foster connections of students' lives with the natural world and, importantly, in their own outdoor spaces at school (Schuttler et al., 2019). In this study, we present one effort to support elementary teachers' science instruction by introducing them to PS projects that include ongoing data collection across a school year and offer opportunities for connecting students with authentic data collection and sense-making aligned with their academic standards.

Outdoor Education

Learning in the outdoors has a long history. In the early 1900s, open-air schools and outdoor education were promoted for health and hygiene (Quay & Seaman, 2013). Interestingly, prior to World War II, science education primarily referred to nature studies (Appleton, 2013), and while most instruction today occurs indoors, learning about the natural world while in the outdoors has been found to contribute to students' cognitive and affective development (Carrier et al., 2014; Rios & Brewer, 2013; Szczytko et al., 2018).

Outdoor instruction is not limited to science, and importantly, it connects learning across discipline areas (Tan & So, 2019). Outdoor learning is often experiential, connecting with both the body and the mind, and such full body connections have been found to positively influence learners' cognition and emotions (Thorburn & Marshall, 2014). When active outdoor learning experiences are connected to indoor lessons, learning is strengthened. Such experiential learning is beneficial for all students and has been found to be especially beneficial for students who struggle with learning or behavior in traditional classroom settings (James & Williams, 2017; Szczytko et al., 2018). As teachers plan to move instruction outdoors, situating outdoor instruction in the familiar schoolyard can decrease the novelties of field trips (Ayotte-Beaudet et al., 2019; Feille, 2021; Martin, 2003).

The schoolyard expands learning beyond the classroom to a setting where students can engage in the practices of science (NRC, 2012; NGSS Lead States, 2013) that include ongoing observations and investigations in the natural world (Ayotte-Baudet, 2017; Carrier et al., 2014; Rios & Brewer, 2014). The schoolyard is readily accessible and avoids field trip costs and logistical challenges of traveling to outdoor parks or nature centers. The accessibility of the schoolyard facilitates opportunities for student engagement with science practices such as observations, data collection, and sense making aligned with science content studies (e.g., life cycles, seasonal changes, weather) in a setting familiar to students across the entire school year.

Theoretical Framework

This study of students' and teachers' PS experiences in their schoolyard is framed in situated learning theory that acknowledges learning is positioned in context, and the context influences learning (Giamellaro, 2017; Sadler, 2009). When students' experiences take place in outdoor environments, their science learning is situated in the context of study (Giamellaro, 2017) and can include examinations of life, earth, and physical science content. Importantly, moving science instruction from the classroom to the schoolyard offers opportunities for students to connect learning with local phenomena (Lloyd et al., 2018).

In addition, sociocultural learning (Vygotsky, 1978) frames teacher and student learning together in the outdoors, which informed Rogoff's (1990) notion of cognitive development with reciprocal contributions of teachers and students when sharing the dual familiarity of the schoolyard. As suggested by Vygotsky's (1978) sociocultural learning theory, knowledge is constructed in one's community, which includes the local context and social interactions to learn science in their schoolyard with their classmates. Importantly, findings from these data and data from the larger study have informed an emerging theory of school-based participatory science (Smith et al., 2025).

Methods

Context

We begin by presenting the context of the larger study for which we selected two PS projects because of their content area alignment with the state's science standards. One PS project, the Community Collaborative Rain, Hail, and Snow Network (CoCoRaHS) aligned with the state's fifth-grade weather standards. The second project, Lost Ladybug Project (LLP) aligned with the state's fifth-grade ecosystems standards. CoCoRaHS was launched in 1998 and currently has over 26,000 active observers in the United States and beyond. Observers collect precipitation data to share with the CoCoRaHS community and help scientists learn more about precipitation patterns (CoCoRaHS, n.d). Examinations of the CoCoRaHS project describe its potential for engaging participants in science (e.g., Jones, 2022; Lackstrom et al., 2022; Mahmoudi et al., 2022), but few promote connections with schools (Sheppard et al., 2022). LLP began in 2000 and was designed to help entomologists document the types and numbers of ladybug species in the US with a goal to conserve declining ladybug populations (The Lost Ladybug Project, n.d.). As with CoCoRaHS, research on LLP primarily identifies participants in the public rather than specific connections with schools (Gardiner et al., 2012; Losey et al., 2012; Lynch et al., 2018; Marchante et al., 2024). The context of the present embedded mixed methods study focuses on teacher and student experiences with the PS projects with a special focus on outdoor teaching and learning.

Participants

Twenty-three fifth-grade teachers and approximately 450 students across one state in the southeastern United States agreed to participate in the first year of this study. Each teacher was asked to incorporate both PS projects - CoCoRaHS and LLP - with their fifth-grade classrooms, but they received support materials for only one of the projects. By providing support materials for only one of the two PS projects to treatment group teachers, our research team sought to learn if and how having support materials contributed to teachers' incorporation of PS with their students. The educative support materials (Arias et al., 2016; Davis et al., 2017) for both projects were designed for this study, created by the research team, and located on a website. Both projects' materials include monthly lessons (August–June), content resources for science, mathematics, and literacy for teachers, and a media guide that includes content-specific readings, videos, books, and interactive guides that related to the project. Both PS projects' support materials include the same teacher recommendations for outdoor instruction such as establishing expectations prior to taking the class outdoors for learning, using multiple senses to make observations outside, collecting and recording data, and exploring students' wonder and curiosity.

The teacher participants were introduced to the two PS projects by attending an in-person professional development for one day over the summer. The professional development included opportunities to begin to consider the inclusion of PS in their classrooms in the upcoming school year. We asked the teachers to implement both projects, though they had support materials for only one project, their treatment group, and teachers served as control group teachers for the PS project for which they did not have educative support materials.

Teachers also attended virtual sessions once a month across the school year for one hour with the research team and other teachers in their treatment group to preview the upcoming month's support materials and discuss their projects. These meetings were held on the same days for each treatment group on the first Wednesday or Thursday of the month after school hours (4:30pm). Meetings for both projects followed the same structure, and while they only discussed the instructional materials of projects for which teachers had support, the meeting schedule included time for teachers to talk about both projects.

Data Collection

This mixed methods study includes both qualitative and quantitative data from teachers and their students (See Table 1). Quantitative data were collected from 23 teacher participants who completed baseline and end-of-year surveys that included questions about how conducive their schoolyard was for outdoor learning related to the PS projects and the amount of time they spent teaching science outdoors in previous years. We also asked teachers about their prior experiences or understanding of PS and their reasons for joining the project. Once the school year began, we asked the 23 teachers to submit weekly instructional logs that asked them to document their activities and classroom decisions related to the PS project, including their estimates of the frequency and approximate percentage of students who spent time outdoors.

Table 1

Data Sources

	Data Sources	Participants: All Teacher Participants	Participants: Only Case Study Teachers	Participants: All Students	Participants: Only Students in Case Study Classrooms
Quantitative	Baseline Survey	X		X	
	Instructional Logs	X		X	
	End-of-year Survey	X		X	
	Content Area Pre-test Ecosystems/ Weather	X		X	
	Content Area Post-test Ecosystems/ Weather	X		X	
Qualitative	Beginning and End-of-Year Interviews		X		
	Six Classroom Observations		X		X
	Post-observation interviews		X		
	Three Focus Groups (Beginning, Middle, and End of Year)				X

To provide a closer look at project implementation, we collected qualitative data with 11 of the 23 teachers as case study participants. From the teachers who expressed interest in participating as case study teachers, we purposefully selected teachers to represent the range of the state's geographic regions, school characteristics (e.g., rural, urban, size), and student populations. While the demographics of teachers in this study were typical of elementary school teachers (i.e., white, female) (Plumley, 2019), we use gender neutral pseudonyms for case study teachers in this paper (Table 2).

Table 2

Participant Design

Group	Treatment Project (Received Support Materials)	Control Project (Did Not Receive Support Materials)	Case Study Teachers
Group A Teachers	LLP	CoCoRaHS	Astor, Jordan, Lian, Taylor
Group B Teachers	CoCoRaHS	LLP	Dana, Morgan, Kody, Perry, Kai, Asa

Note. We use gender neutral pseudonyms for case study teachers in this paper

Project researchers *observed* each case study teacher's "target" science class (for instances when a teacher taught science to more than one class) at least six times over the school year and documented the observations using an observation guide that was developed by the research team. The observation guide helped researchers record the type of activity (e.g., data collection, class discussion, group work), setting (indoor or outdoor), grouping structure (e.g., whole class, small group), proportion of students engaged in the activity (e.g., 50-75%, 75-100%), and interdisciplinary connections (e.g., mathematics, language arts). Researchers recorded field notes on the observation guide using time stamps to document the length of each of the activities observed. Researchers *interviewed* the case study teachers seven times starting with a baseline interview prior to the first observation, then conducted post-observation interviews following each of the six observations. The interviews asked teachers about their PS lesson planning, implementation, reflections on their students' experiences with PS in the classroom and schoolyard, and their interactions with the support materials. In each case study teacher's class, the project researchers also conducted three *focus group* interviews with students near the beginning, middle, and end of the school year. The teachers selected four to six students with parent consent and student assent, and each of the three focus groups consisted of different student participants.

Data Analysis

Quantitative data included teacher baseline and end-of-year *surveys* that examined teachers' prior experiences with PS and outdoor instruction. Weekly *instructional log data* were analyzed for patterns using the PS projects both in the classroom and the schoolyard. Log data documented teachers' accounts of the frequency and time of their engagement with the PS projects, use of support materials, and teachers' considerations that informed their planning and instruction.

Qualitative *classroom observation* field notes combined with *teacher interviews* were recorded and transcribed. We used inductive coding to identify initial codes and organized them by themes (Riger & Sigurvinsdottir, 2016). Following discussions on initial themes and code meanings (Chandra et al., 2019), two researchers independently coded the same interviews, discussed differences, and through negotiated agreement (Belotto, 2018) reconciled the remaining differences. The researchers identified common interpretations of themes that captured teachers' and students' experiences with outdoor learning. Themes include teacher views of outdoor instruction, their views of students' outdoor experiences, and student impressions of outdoor instruction. *Student focus group* data document students' cognitive and/or affective reactions to the outdoors, classroom connections, connections to their lives, and students' feelings about learning outdoors (Table 3).

Table 3

Themes, Descriptions, and Sample Quotes

Code	Description	Sample Quotes
	Theme: Teacher Views about Outdoor Instruction	
Authentic learning experiences in the outdoors	Teachers' descriptions of the authenticity of teaching outdoors	*Now I am taking them out for more of a purpose. (Astor)*
Schools' focus on test preparation	Teachers' views on standardized testing pressures	*It gave me permission, quite frankly because I was part of this study…I got permission from up high to walk away from teaching to the test. So that's a really good thing because it felt more authentic. (Morgan)*
Situating learning outdoors	Teachers' views on moving instruction outdoors	*Outdoor learning reinforces that learning can be anywhere. (Kody)*
Benefits of outdoor instruction	Teachers' views on positive aspects of outdoor instruction	*I think they always love the outdoors. I got outdoors way more this year than I did before, and I really think that that was a good idea and I will be doing that again next year. (Kai)*
Challenges of outdoor instruction	Teachers' views on obstacles to outdoor instruction	*I had great intentions but one of my barriers was I don't really have a safe way for the students to go out and check the rain gauge without me. (Lian)*
	Theme: Teacher Views about Student Experiences in the Outdoors	
Student enthusiasm for the outdoors.	Teachers' views on students' excitement about outdoor lessons	*They had the freedom to move around in a space that they don't normally get to run around and move in with the purpose. And I think anytime you can get outside you feel better. (Astor)*
Students' limited outdoor experiences	Teachers' views on their students spending little time outdoors	*The most rewarding aspect of Lost Ladybug was) getting them outside because they need to learn that you don't have to be behind a computer to learn. (Kody)*
Students' engagement in learning outdoors	Students' engagement in learning when outdoors	*Going outside is always very engaging, so they really enjoyed that. That was fun. And I haven't taken the class outside in a while, so it was helping me remember, oh, kids really, really do enjoy going outside.' As long as we have a specific purpose for why we're outside. (Perry)*
	Theme: Student Views about the Outdoors	

Cognitive reactions to the outdoors	Students sharing cognitive connections outdoors.	*I feel more thoughtful about things in the outdoors*
Affective reactions to the outdoors	Students' descriptions of feelings when outdoors.	*Outside is more calm than in the classroom. Classrooms are wild*
Classroom connections	Students connecting indoor lessons with outdoor learning.	*[PS Connects to] what we do in social studies, we do latitude and longitude, prime meridian, equator.*
Connections to students' lives	Students' reflections on outdoor learning with their own lives.	*It's changed me a lot, because now every time I go outside, I always look up on the clouds and now it's a habit.*
Learning in the outdoors	Students' comments about learning when outdoors.	*I've never seen clouds that look like that before, and before this I never even noticed anything. So, this has helped me learn a lot.*

Limitations

We acknowledge that the teachers chose to participate in this study, introducing self-selection bias. We further recognize that elementary school outdoor areas vary widely and can limit student access for data collection with some PS projects, thus limiting the generalizability of these findings. While there were urban schools in this study, we recognize that one factor that influenced teachers' decisions to participate in this study was the potential of their schoolyard for outdoor data collection. This includes presence of vegetation and attention to safety concerns.

Findings

Quantitative – Surveys and Instructional Logs

Survey data

At the start of the study, teachers completed a baseline survey that included questions about their feelings of preparation to use their school grounds to teach science and how frequently they took their students outside for science instruction in the previous year. Teachers answered the same questions on the end-of-year survey. On the baseline survey, about 70 percent of the teachers said they did not have science instructional materials designated by their school/district. Sample data in Table 4 reveal slight variations in survey data with no significant changes in teachers' feelings of preparedness for outdoor instruction after one academic year.

Table 4

Sample Survey Data

	How well prepared do you feel to use your school grounds to teach science?			
	Not adequately prepared	Somewhat prepared	Fairly well prepared	Very well prepared
Baseline Survey (N=23)	3	4	10	6
End-of-year Survey (N=22)	1	5	11	5

	How often do you take your students outside for science instruction?				
	Never	Rarely	Sometimes (once a month)	Often	All or almost all
Baseline Survey (N=23)	1	6	13	3	0
End-of-year Survey (N=22)	0	7	8	6	1

Instructional Log Data

Weekly instructional log data revealed variations in time spent on each PS project and time outdoors. Instructional log data for *all* participants documented teachers spending little time on projects for which they did not have support materials. For CoCoRaHS treatment teachers who had CoCoRaHS support materials, weekly log data show that their students spent an average of about 26 minutes each week on activities related to CoCoRaHS (e.g., reading the rain gauge; submitting data), and 18 minutes each week for activities related to LLP (e.g., searching for ladybugs; submitting data). For teachers who had LLP support materials, weekly log data show that their students spent an average of about 38 minutes each week on activities related to LLP and 13 minutes each week on activities related to CoCoRaHS.

Teachers were asked, "How much time did a typical student in this class spend outdoors as a part of [LLP or CoCoRaHS] this week?" Log data indicated that students went outdoors more frequently for CoCoRaHS (5,565 total minutes) than LLP (3,486 total minutes), collecting precipitation data most days with one to four students reading the rain gauge. Interestingly, while the teachers' log data indicated that students went outdoors less frequently for LLP compared to CoCoRaHS, when they went outside, more students (the entire class of approximately 20 students) went outdoors to search for ladybugs, and they also spent more time outdoors when compared to CoCoRaHS. These log data were reinforced by the observation and interview data collected with the case study teachers, helping us learn more about teachers' views of their own and their students' experiences learning outdoors.

Qualitative - Interviews, Observations, and Student Focus Groups

In the baseline interview, we asked about their school's science curriculum, their familiarity about PS, and enactment of the PS projects. Teachers reported that, while their state posted elementary science standards, most districts did not provide curriculum or instructional materials for science. When asked about curriculum and instructional resources, Morgan reported using a "Scholastic magazine, and various online resources, but it's up to me how to design it and use those resources." A baseline interview question asked case study teachers about their familiarity with PS, and Perry said "I was familiar with the idea of it, but not fully firm on what it meant. I just thought it was doing kinds of outdoorsy stuff with your kids that benefited the outdoors, like ecosystems." As with teachers' log data, field notes from researchers' observations of the case study teachers' lessons and teacher interviews found teachers spent little time on projects for which they did not have support materials. In an end-of-year interview, when asked about LLP, Perry, who had support materials for CoCoRaHS said, "I haven't worked with that (LLP) as much or hardly at all," and explained that without support materials, they were unsure what to do.

Observations

Here we describe sample observation data from two case study teachers' enactment of their PS project's introductory activity. CoCoRaHS: Morgan, who had materials for CoCoRaHS, used an introductory activity from the materials and took the entire class outside on a "weather walk" to orient students to using their senses and rain gauge measurements to collect weather data and introduced students to the location of the rain gauge. Students' sensory observations included looking at clouds and identifying cloud types, discussing how the air felt on their skin, touching dew on the grass and speculating how the dew got there. After the whole-class introduction of the CoCoRaHS project in the outdoors, observation data from the remainder of the school year documented that Morgan sent only one student or a small group (2-4 students) outdoors to collect rain gauge data at the start of the school day. Many of the monthly CoCoRaHS activities were designed to give students experiences organizing and making sense of precipitation data; the observation data documented that, in Morgan's class, these monthly frequently occurred indoors.

An example of Taylor's enactment of LLP documented their adaptation of an introductory activity from the support materials designed to orient students to the schoolyard as an outdoor classroom and connect with state science standards on ecosystems. In the activity, students were asked to map their schoolyard and identify its features and types of vegetation. The goal of this initial activity was for students to the use their maps throughout the school year to document the locations where they found ladybugs. Rather than asking students to draw maps, Taylor chose to prepare a map template of the schoolyard in advance prior to taking the students outside and asked students to label features on their map and discussed using map keys. Although the purpose of this initial activity was designed to provide students with a map to plot the location of the ladybug sightings across the school year, its use was not documented in observations of Taylor's subsequent monthly activities. Next, we present the case study teachers' reflections on the benefits and challenges of situating instruction in the outdoors, as well as the impact outdoor learning had on their students.

Interviews

Benefits. In interviews, seven of the 11 case study teachers described how the PS projects created more purposeful opportunities for science and outdoor learning throughout the school year. Astor said, "I [used to] take them outside just to do work…but now I'm taking them out for more of a purpose…there's actually a reason." Perry's reflection on PS and purpose was, "Kids really, really do enjoy going outside, as long as we have a specific purpose for why we're outside."

Jordan contrasted the activities from the project's support materials with standards-focused school norms and described its impact on their teaching. They explained:

> This was so much more than just a science lesson or a math lesson. It was so much a part of our class. It was activities that we did together that we enjoyed…It has definitely changed my teaching in that [with PS activity] I wasn't standards-focused, I was purpose-focused.

Other teachers contrasted outdoor instruction with their school's culture focused on accountability and testing. Kyle explained, "I like to be outdoors…I think the kids learned a ton of things that you can't get from a test. I think that we're so tied to doing a test…It's just unfortunate." One case study teacher shared their school's repetitive focus on testing and reviewing, noting that outdoor learning broke up this pattern:

> I think it's something that I get more excited about. At the end of the day, this test is what drives, but it is nice to have that little break once a month to do something different that's not geared toward the test but still on topic. It's hard to justify that sometimes to a principal about doing something that's not on the test. They want us to just teach, teach, teach, review, review, review the whole time.

For some, viewing the outdoor areas of the schoolyard as a setting for learning prompted a shift in both teachers' and students' thinking. Kody reflected that "[PS] reinforces that learning can be anywhere. Outside doesn't have to be recess," and Morgan explained their students' learning "that outdoors can be a learning experience, not just for playtime." Taylor expressed their own intentions to grow:

> I definitely want to try to get them outdoors more…I'm one of those [teachers] - the desks are lined up…it was kind of neat to see how I could teach them outdoors. You know, they don't just have to be sitting behind a desk and listening to me, watching me on the Smart Board, so for sure getting them outside more.

Challenges. During post-observation interviews, case study teachers were asked about challenges they faced with outdoor instruction. Lack of time emerged as a prominent theme; six of the 11 teachers cited the amount of time it took out of the instructional day to go outside. Referring to a limited block of time for science instruction, Dana said:

> We definitely did not go outside as much as I would've liked. So, when it comes to my challenges, time is gonna be a big part of that and just having that time…one of those challenges that I'm gonna try to figure out this summer.

Kody similarly expressed time as the greatest challenge to including outdoor PS saying, "The biggest thing is getting outside and checking that rain gauge and mostly it's because of time constraints."

Other teachers expressed aspirations for planning more outdoor learning opportunities in following years. Lian hoped that their move to a different school would support these goals, "I had great intentions, but one of my barriers was I don't really have a safe way for the students to go out and check the rain gauge without me. But next year they'll be able [to]." For many teachers, the learning curve for adapting new initiatives takes time (Pak et al., 2020), as Morgan described "I didn't do hardly anything with it till the very end because my ecosystem unit is at the end." Promisingly, even the teachers like Jordan, who did not fully implement the PS projects outside in the first year, expressed future intentions, "I'm really excited for next year and the fact that I'm going to start out the year ready to roll with it, more so than I did last year.

Students' Outdoor Experiences

In interviews, many case study teachers described how their students have limited opportunities to experience the outdoors outside of school. Perry shared:

> Some kids don't get outside at all… [One student] didn't really think about being outside and connecting with nature and prefers to stay inside and play a lot of video games. I must remember that not all kids go home and play outside, so it's an important experience for them.

Teachers shared how involvement in the projects provided students with new and meaningful ways to engage with the outdoors. Taylor explained, "[Students] are able to look up from the screens…they definitely have a better appreciation for the outdoors." At the end of the year, Astor shared the lasting impact that the LLP had on their students:

> What they learned is to be observers and aware of their natural surroundings and not just for ladybugs, for everything, because they had so many questions. They'd find another insect, or they'd find a certain flower and they were curious, and they were constantly asking questions. And I think this project has just made them more aware of the natural world around them and it's made them thoughtful. You don't just see a ladybug and or an ant and step on it. You know that everything has a purpose, and I hope that's what they'll take with them.

Other teachers described the ways students' outdoor learning experiences through these PS projects extended beyond searching for ladybugs or measuring precipitation. Lian shared how a student connected the project with her interests of both science and history:

> I had one young lady that brought in a field guide, and she's taking it outside and identifying all the plants and then telling me what the plants were used for in the Civil War. She said, "Did you know this one was used to dye cloth during the Civil War?" So, I saw them taking what we started with our ladybugs and then taking that into their own interests.

Teachers described students' connections of outdoor science with other subjects and to their lives were clear evidence of student enjoyment.

Student Engagement

In addition to sharing students' enjoyment, teachers also described students' enthusiasm for outdoor learning. In their interviews, many case study teachers shared that students were happy when they were outside, had a purpose outdoors beyond playtime, and felt better outdoors. Astor said, "They had the freedom to move around in a space that they don't normally get to run around and move with a purpose. And I think anytime you can get outside you feel better."

In end-of-year interviews, many case study teachers identified getting students outdoors to learn as the most rewarding part of the yearlong project. Jordan described high levels of student engagement in outdoor learning, "All of the students participated. I had 100% participation, 100% feedback. I don't know any other way in school that you get that return on the investment". In addition to their enthusiasm for outdoor instruction, students connected with science: "They like to be outdoors; they spent more time outside this year than ever before - science is their favorite." Astor talked about how impactful it was for students to be part of a community of outdoor scientists:

> I think [the projects] made [science] more real. I think they really believe, and they should believe, they were helping these scientists in the world because I kept telling them, "There's not enough scientists in the world to do all this research, and so we have to help them so they can figure out problems." And I think they felt very much a part of that community of scientists.

Jordan reflected, "It's like they had that purpose for being out there. And it wasn't intimidating to them to be in nature." Some students carried their excitement beyond the classroom and shared their joy for the PS projects at home. Astor described:

> [It was rewarding to] watch my children really grow differently this year than they had in the past. They always grow, but just to see the joy and the excitement when we found our first ladybug, it was just exciting to see them in it. And watching their families get into it and their little brothers and their sisters and all the pictures that were sent from all different places that they were thinking about this, beyond the five days that they go to school. That's impactful.

Student Focus Groups

To gather student perspectives on the projects and time outdoors, we conducted focus groups with different students from each class at the beginning, middle, and end of the school year. Focus groups explored students' impressions of learning in the outdoors, cognitive growth, and connections to their lives. Student quotes include their treatment group project in quotations, CoCoRaHS (CCR) or Lost Ladybug project (LLP).

Learning Outdoors

Many students expressed their appreciation for outdoor experiences. There were several student comments that compared learning outdoors to the classroom: "Being in nature...actually see what you're learning in the classroom (LLP)";" and "Outside is calmer than in the classroom. Classrooms are wild (CCR)." One student said, "It's great for kids to feel like they're more adult. We get to be outside and use our minds rather than sitting inside. Unlike what a normal teacher will do with three days of ladybugs on the screen (LLP)." Other descriptions contrasting outdoor learning experiences with classroom learning included, "I like the outside more. It just seems refreshing to be somewhere that's not class (LLP)"," and, powerfully, "I love going outside. Being free instead of being in jail (CCR)."

Cognitive Benefits

Some students described the cognitive benefits of learning outdoors such as "I feel more thoughtful about things in the outdoors" and that it's "easier to concentrate outside (CCR)." Another student explained, "I'm better at doing precipitation [now]. I just didn't understand that, so you feel like you're understanding the patterns, or even what it means (CCR)." Other students identified how being active contributed to their learning. One student said, "Learning doesn't have to be boring. We can learn more about nature, active learning. It is exquisite (LLP)." Another explained, 'I don't really like bugs a lot, but I like this project because I like doing, and everyone is involved, and I like to DO projects (CCR)." Student comments also included descriptions of student autonomy, "We're going and doing it on our own and we're able to look at these things and research these things on our own, like we're not sitting there watching what somebody else is doing (CCR)."

Students connected their activities to learning the project content and considered the larger impact of the PS projects. Students described how they compared their precipitation data to the data of CoCoRaHS stations across the state, noting, "We can pair the counties to see why they would have more or less rain than us (CCR)," and "I've learned that even though you're not far away from other counties, there is still a big difference in the precipitation they get (CCR)."

Students also considered the ways the PS projects connected to their learning in other content areas. One student referred to the opportunities their class had to graph the precipitation data they collected as part of the CoCoRaHS project, saying, "The graphing was my favorite, because it teaches us to do something. It's math, but like you're not learning it, you're doing something fun (CCR)." Others similarly shared, "I didn't know anything about decimals, and then I figured it out (CCR)" and "Graphing teaches math with fun (CCR)."

Connections to Students' Lives and Learning Outdoors

In each of the three focus groups, researchers asked students about their favorite and least favorite parts about being outdoors. Focus group conversations elicited clear examples of students connecting the projects to their lives and how they interact with the natural world. In their words: "I am more observant; I pay more attention [since participating in PS project] (LLP);" "It's changed me a lot, because now every time I go outside, I always look up at the clouds and now it's a habit (CCR);" and, "It was really cool to learn more about the rain because we don't normally think about it (CCR)." Other students talked about sharing outdoor experiences with their peers. One said "Discovering new insects is great. You get to work together to find out what it is (LLP)." Another student appreciated both the active and social aspects of learning outdoors, "We get to move and talk with each other when we're outside (CCR)."

The focus group data also documented students' least favorite parts of outdoor learning and these focused on inclement weather, bugs, or students' own fears. Sample student comments about their least favorite parts are, "I don't like the ladybug project. Ladybugs are scary (LLP)," and "I don't like my shoes getting dirty whenever we have to walk in the grass to the rain gauge (CCR)," and "sometimes way too hot or too cold (CCR)." As we continue to analyze teachers' interactions with the educative materials for PS, we consider additional ways to support teachers' efforts to reach all students.

Discussion

Extending science learning to the familiar and accessible schoolyard has been found to enhance both student learning and enthusiasm (Aflalo et al., 2020). Aligned with decades of research showing that outdoor learning is beneficial to students' cognitive achievement (e.g., Disinger, 1987), students in this study described both cognitive and affective benefits of situating instruction in the outdoors. Importantly, both teachers and students shared their appreciation for active learning experiences outdoors. Active learning (Mizokami, 2018; Vanhorn et al., 2019) has been defined with language such as "engagement" and "authentic" learning and contrasted with passive listening to a teacher's instruction. Here we argue that PS projects have the potential to support consistent and active learning outside the classroom.

There is a dearth of research on PS projects in elementary schools. One article written for elementary teachers provides a strong introduction of PS and LLP, including examples of PS projects and LLP activities (Harris & Ballard, 2018). Our study extends this research to suggest that developing educative curriculum materials for PS projects can support both teacher and student learning (Arias et al., 2016; Davis et al., 2017). Importantly, incorporating PS projects in formal education offers opportunities to extend learning beyond the classroom to the schoolyard.

In our study, teachers described how these two PS projects that included schoolyard experiences positively impacted students' enthusiasm for science and offered direct connections of science to students' lives (Ayotte-Beaudet et al., 2023; Berg et al., 2021). For example, students described comparing precipitation data from their own rain gauge with precipitation data collected in other geographic locations in their state, demonstrating their excitement for feeling connected to other communities through the CoCoRaHS PS project. These data also reveal examples of students connecting their PS experiences to other subject areas beyond science, including mathematics and social studies (Tan & So, 2019).

In focus group interviews, students shared their enjoyment of learning outdoors and their connections with nature (Ayotte-Baudet, 2017; Carrier et al., 2014; Rios & Brewer, 2014). Participatory science projects that include ongoing data collection in the outdoors can help students recognize their contributions to the work of scientists. "Students can then appreciate what their observations mean and how they might fit with those of others into the missions of broader science initiatives" (Esch et al., 2020, p. 5). In our study, data suggest that, as students learn that data collection in science is not limited to one teaching unit or time frame, they begin to learn more about the work of professional scientists. Such student participation and sharing data collection and sense-making opportunities in the classroom and outdoors with their classmates have been found to deepen collective learning (Krist & Shim, 2024).

Importantly, the PS projects and outdoor instruction in this study seemed to ignite both teacher and student interest (Dillon et al., 2016; Oberle et al., 2021; Rios et al., 2014).
Teachers in this study acknowledged the well-documented challenges of time, preparation for taking students outdoors, and the pressure to prepare their students for standardized testing which mirror challenges found in elementary classroom science education more broadly (Banilower et al., 2018; Plumley, 2019). However, the teachers' perceived benefits of engaging their students in PS projects appeared to motivate them to navigate past these obstacles, and many teachers in this study expressed intentions to engage more frequently or more deeply with the projects and outdoor instruction in the future. Findings from our study also suggest that providing teachers with support materials specific to a PS project can help teachers connect classroom and schoolyard instruction as they contribute data to the PS project.

Implications

Findings from this study suggest that including PS projects in elementary school classrooms can encourage regular outdoor science learning experiences that enhance elementary science instruction and increase students' enthusiasm for learning. Elementary school science programs can benefit from PS projects that include supports designed to meet teacher needs, creating a culture of learning outdoors frequently and with purpose (Barfod & Bentsen, 2018). We suggest that, in partnership with educators (Carrier et al., 2024), PS project leaders can design educative support materials (Arias et al., 2018; Davis et al., 2014) to enhance efforts to expand their PS projects to formal education settings. Such authentic data collection and sensemaking in the schoolyard can provide young learners with science experiences with their peers, as they collect and share data, participating in the enterprise of science. Data from the present study were collected in the first year of a larger study, and because many teachers in this first year described intentions to "do more" in the following school year, this project's continuing research, and future research on PS in formal education can extend our understanding of this study's data over time.

Acknowledgements

NSF Acknowledgement–This material is based upon work supported by the National Science Foundation under Grant No. 2009212. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

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

Sarah J. Carrier (sjcarrie@ncsu.edu) taught elementary school and is a Professor of Science Education at North Carolina State University. Her research has included studies of literacy and science, science teacher identity, awe in science, and has focused on issues of climate change and environmental education. Her true heart is supporting elementary teachers situating instruction in the outdoors.

Danielle R. Scharen (d.scharen@horizon-research.com) is a Lead Researcher at Horizon Research, Inc. where she focuses on K-12 science education research. Prior to joining HRI in 2023, Danielle taught undergraduate science education at North Carolina State University and was an elementary school teacher in Wake County Public Schools.

Meredith L. Hayes (Meredith.hayes15@gmail.com) is a research Associate at Horizon Research, Inc., has worked on several projects funded by the U.S. National Science Foundation and the NASA Science Mission Directorate, both as a researcher and external evaluator. In addition, she has taught upper elementary grades and served as the school mathematics/science specialist, working withK-5 students and teachers. Her current work focuses on investigating how teachers implement participatory science projects.

P. Sean Smith (ssmith62@horizon-research.com) is a Senior Executive at Horizon Research, Inc. (HRI), received a Bachelor's Degree in Chemistry, a Master's Degree in Science Teaching, and a Ph.D. in Curriculum and Instruction. He is currently Principal Investigator for the NSF-funded Supporting Elementary Teacher Learning for Effective School-Based Citizen Science (TL4CS). Dr. Smith also leads the evaluations of several projects funded by the U.S. National Science Foundation.

Christine Goforth is the head of citizen science at the North Carolina Museum of Natural Sciences, a role that engages members of the public with scientific research they can participate in as non-experts. Chris holds a BA in Biology from Colorado College and a MS in Entomology from the University of Arizona. Her research areas focus on aquatic insect ethology and physiology as well as using insects to determine water quality in a variety of settings.

Laura Craven (l.craven@horizon-research.com) is a Research Associate at Horizon Research, Inc. where she develops surveys and conducts data analyses for a variety of STEM education projects.

Lindsey Sachs (l.sachs@horizon-research.com) is a research Associate at Horizon Research, Inc. (HRI). She received her Bachelor of Arts degree in Elementary Education from the University of North Carolina at Chapel Hill and a Master of Education degree from North Carolina State University. Prior to joining HRI in 2020, she worked for the Wake County Public Schools.

References

Aflalo, E., Montin, R., & Raviv, A. (2020). Learning outdoors or with a computer: the contribution of the learning setting to learning and to environmental perceptions. *Research in Science & Technological Education*, *38*(2), 208-226. https://doi.org/10.1080/02635143.2019.1603141

Appleton, K. (2013). Science pedagogical content knowledge and elementary school teachers. In *Elementary science teacher education* (pp. 31-54). Routledge.

Arias, A. M., Bismack, A. S., Davis, E. A., & Palinscar, A. S. (2016). Interacting with a suite of educative features: Elementary science teachers' use of educative curriculum materials. *Journal of Research in Science Teaching*, *53*(3), 422–449. https://doi.org/10.1002/tea.21250

Ayotte-Beaudet, J. P., Chastenay, P., Beaudry, M. C., L'Heureux, K., Giamellaro, M., Smith, J., ... & Paquette, A. (2023). Exploring the impacts of contextualised outdoor science education on learning: the case of primary school students learning about ecosystem relationships. *Journal of Biological Education, 57*(2), 277-294. https://doi.org/10.1080/00219266.2021.1909634

Ayotte-Beaudet, J. P., Potvin, P., & Riopel, M. (2019). Factors related to middle-school students' situational interest in science in outdoor lessons in their schools' immediate surroundings. *International Journal of Environmental and Science Education*, *14*(1), 13-32. https://doi.org/10.29333/ijese/7815

Ayotte-Beaudet, J. P., Potvin, P., Lapierre, H. G., & Glackin, M. (2017). Teaching and learning science outdoors in schools' immediate surroundings at K-12 levels: A meta-synthesis. *EURASIA Journal of Mathematics, Science and Technology Education*, *13*(8), 5343-5363. https://doi.org/10.12973/eurasia.2017.00833a

Banilower, E. R., Smith, P. S., Malzahn, K. A., Plumley, C. L., Gordon, E. M., & Hayes, M. L. (2018). *Report of the 2018 NSSME+*. Horizon Research, Inc.

Barfod, K., & Bentsen, P. (2018). Don't ask how outdoor education can be integrated into the school curriculum; ask how the school curriculum can be taught outside the classroom. *Curriculum Perspectives*, *38*, 151-156. https://doi.org/10.1007/s41297-018-0055-9

Belotto, M. J. (2018). Data analysis methods for qualitative research: Managing the challenges of coding, interrater reliability, and thematic analysis. *The Qualitative Report*, *23*(11), 2622-2633. https://doi.org/10.46743/2160-3715/2018.3492

Berg, S., Bradford, B., Barrett, J., Robinson, D. B., Camara, F., & Perry, T. (2021). Meaning-making of student experiences during outdoor exploration time. *Journal of Adventure Education and Outdoor Learning, 21*(2), 172-183. https://doi.org/10.1080/14729679.2020.1769694

Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science?. *Public Understanding of Science*, *25*(1), 2-16. https://doi.org/10.1177/0963662515607406

Carrier, S.J., Scharen, D.R., Hayes, M., Sean, P.S., Bruce, A., Craven, L. (2024). Citizen science in elementary classrooms: A tale of two teachers. *Frontiers in Education, 9*. https://doi.org/10.3389/feduc.2024.1470070

Carrier, S.J., Thomson, M.M., & Tugurian, L.P., Stevenson, K.T. (2014). Elementary science education in classrooms and outdoors: Stakeholder views, gender, ethnicity, and testing. *International Journal of Science Education*, *36*(13), 2195-2220.

Carrier, S.J., Tugurian, L.P. & Thomson, M.M. (2013). Elementary science indoors and out: Teachers, time, and testing. *Research in Science Education*, *43*(5), 2059-2083.

Chandra, Y., Shang, L., Chandra, Y., & Shang, L. (2019). Inductive coding. *Qualitative research using R: A systematic approach*, 91-106. https://doi.org/10.1007/978-981-13-3170-1

CoCoRaHS. (n.d.). *About us. https://www.cocorahs.org/content.aspx?page=aboutus.*

Curran, F. C., & Kellogg, A. T. (2016). Understanding science achievement gaps by race/ethnicity and gender in kindergarten and first grade. *Educational Researcher*, *45*(5), 273-282. https://doi.org/10.3102/0013189x16656611

Davis, E., Palinscar, A. S., Arias, A. M., Bismack, A. S., Marulis, L., & Iwashyna, S. (2014). Designing educative curriculum materials: A theoretically and empirically driven process. *Harvard Educational Review*, *84*(1), 24-52. https://doi.org/10.17763/haer.84.1.g48488u230616264

Dillon, J., Rickinson, M., & Teamey, K. (2016). The value of outdoor learning: evidence from research in the UK and elsewhere. In T*owards a convergence between science and environmental education* (pp. 193-200). Routledge. https://doi.org/10.4324/9781315730486

Disinger, J. F. (1987). *Cognitive learning in the environment: Elementary students.* ERIC/SMEAC Environmental Education Digest No. 2, 1987.

Doan, S., Eagan, J., Grant, D., Kaufman, J. H., & Setodji, C. M. (2022). *American Instructional Resources Survey. 2022 Technical Documentation and Survey Results. Research Report.* RR-A134-14. In RAND Corporation. RAND Corporation. https://doi.org/10.7249/RRA134-14

Dunn, R. R., Urban, J., Cavelier, D., & Cooper, C. B. (2016). The tragedy of the unexamined cat: Why K–12 and university education are still in the dark ages and how participatory science allows for a Renaissance. *Journal of Microbiology & Biology Education*, *17*(1), 4-6. https://doi.org/10.1128/jmbe.v17i1.1049

Esch, R. K., Burbacher, E., Dodrill, E., Fussell, K. D., Magdich, M., Norris, H., & Midden, W. R. (2020). Citizen science in schools: Scientists' perspectives on promise and pitfalls. *Horizon Research, Inc.*

Explore Reports. (n.d.). EdReports. Retrieved March 2, 2024, from https://edreports.org/reports/science

Feille, K. (2021). A framework for the development of schoolyard pedagogy. *Research in Science Education*, *51*(6), 1687-1704. https://doi.org/10.1007/s11165-019-9860-x

Giamellaro, M. (2017). Dewey's yardstick: Contextualization as a crosscutting measure of experience in education and learning. *Sage Open, 7*(1). 1-11. https://doi.org/10.1177/2158244017700463

Harris, E., & Ballard, H. (2018). Real science in the palm of your hand. *Science and Children*, *55*(8), 31-37. https://doi.org/10.2505/4/sc18_055_08_31

Irish, T., & Kang, N. H. (2018). Connecting classroom science with everyday life: Teachers' attempts and students' insights. *International Journal of Science and Mathematics Education*, *16*, 1227-1245. https://doi.org/10.1007/s10763-017-9836-0

James, J. K., & Williams, T. (2017). School-based experiential outdoor education: A neglected necessity. *Journal of Experiential Education*, *40*(1), 58-71. https://doi.org/10.1177/1053825916676190

Jones, K. F. (2022). A citizen science freezing rain tree. In *Proceedings–Int. Workshop on Atmospheric Icing of Structures, IWAIS* (pp. 19-23).

Jones, G., Childers, G., Stevens, V., & Whitley, B. (2012). Citizen scientists: Investigating science in the community. *Science Teacher*, *79*(9), 36–39. https://doi.org/10.1080/21548455.2018.1475780

Krist, C., & Shim, S. Y. (2024). Which ideas, when, and why? An experienced teacher's in‐the‐moment pedagogical reasoning about facilitating student sense‐making discussions. *Journal of Research in Science Teaching*, *61*(2), 255-288. https://doi.org/10.1002/tea.21908

Lackstrom, K., Farris, A., & Ward, R. (2022). Backyard hydroclimatology: Climate scientists contribute to drought detection and monitoring. *Bulletin of the American Meteorological Society*, *103*(10), E2222-E2245. https://doi.org/10.1175/bams-d-21-0157.1

Losey, J., Allee, L., & Smyth, R. (2012). The Lost Ladybug Project: Citizen spotting surpasses scientist's surveys. *American Entomologist*, *58*(1), 22-24.

https://doi.org/10.1093/ae/58.1.0022
Lost Ladybug Project. (n.d.). *About us.* http://www.lostladybug.org/about.php
Lowell, B. R., Cherbow, K., & McNeill, K. L. (2021). Redesign or relabel? How a commercial curriculum and its implementation oversimplify key features of the NGSS. *Science Education*, *105*(1), 5-32. https://doi.org/10.1002/sce.21604
Lloyd, A., Truong, S., & Gray, T. (2018). Place-based outdoor learning: More than a drag and drop approach. *Journal of Outdoor and Environmental Education, 21*, 45-60. https://doi.org/10.1007/s42322-017-0002-5
Lucky, A., Savage, A. M., Nichols, L. M., Castracani, C., Shell, L., Grasso, D. A., ... & Dunn, R. R. (2014). Ecologists, educators, and writers collaborate with the public to assess backyard diversity in The School of Ants Project. *Ecosphere*, *5*(7), 1-23. https://doi.org/10.1890/ES13-00364.1
Lynch, L. I., Dauer, J. M., Babchuk, W. A., Heng-Moss, T., & Golick, D. (2018). In their own words: The significance of participant perceptions in assessing entomology citizen science learning outcomes using a mixed methods approach. *Insects*, *9*(1), 16. https://doi.org/10.3390/insects9010016
Mahmoudi, D., Hawn, C. L., Henry, E. H., Perkins, D. J., Cooper, C. B., & Wilson, S. M. (2022). Mapping for whom? Communities of color and the citizen science gap. *ACME: An International Journal for Critical Geographies*, *21*(4), 372-388. https://doi.org/10.14288/acme.v21i4.2178
Marchante, E., López-Núñez, F. A., Duarte, L. N., & Marchante, H. (2024). The role of citizen science in biodiversity monitoring: when invasive species and insects meet. In *Biological Invasions and Global Insect Decline* (pp. 291-314). Academic Press. https://doi.org/10.1016/b978-0-323-99918-2.00011-2
Martin, S.C. (2003). The influence of outdoor schoolyard experiences on students' environmental knowledge, attitudes, behaviors, and comfort levels. *Journal of Elementary Science Education*, *15*(2), p. 51-56.
Mizokami, S. (2018). Deep active learning from the perspective of active learning theory. Deep active learning: Toward greater depth in university education, 79-91. https://doi.org/10.1007/978-981-10-5660-4_5
National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* National Academies Press.
National Science Teaching Association (NSTA). (2023). *Science Standards.* https://www.nsta.org/science-standards.
NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. Washington, DC: The National Academies Press.
Oberle, E., Zeni, M., Munday, F., & Brussoni, M. (2021). Support factors and barriers for outdoor learning in elementary schools: A systemic perspective. *American Journal of Health Education, 52*(5), 251-265. https://doi.org/10.1080/19325037.2021.1955232
O'Donnell, C. (2023). *Exploring the Experiences of K-12 Science Teachers Using Citizen Science to Engage Student Learners: An Interpretive Phenomenological Analysis* (Doctoral dissertation, Northeastern University). https://doi.org/10.17760/d20483515
Pak, K., Polikoff, M. S., Desimone, L. M., & Saldívar García, E. (2020). The adaptive challenges of curriculum implementation: Insights for educational leaders driving standards-based reform. *Aera Open*, *6*(2), 2332858420932828. https://doi.org/10.1177/2332858420932828
Plumley, C. L. (2019). *2018 NSSME+: Status of elementary school science.* Horizon Research, Inc.
Quay, J., & Seaman, J. (2013). Outdoor education and indoor education. In *John Dewey and education outdoors* (pp. 13-44). Brill. https://doi.org/10.1007/978-94-6209-215-0_3

Riger, S., & Sigurvinsdottir, R. (2016). Thematic analysis. *Handbook of methodological approaches to community-based research: Qualitative, quantitative, and mixed methods*, 33-41. https://doi.org/10.1093/med:psych/9780190243654.003.0004

Rios, J. M., & Brewer, J. (2014). Outdoor education and science achievement. *Applied Environmental Education & Communication, 13*(4), 234-240. https://doi.org/10.1080/1533015X.2015.975084

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Sadler, T. D. (2009). Situated learning in science education: socio-scientific issues as contexts for practice. *Studies in Science Education*, *45*(1), 1-42. https://doi.org/10.1080/03057260802681839

Schuttler, S. G., Sorensen, A. E., Jordan, R. C., Cooper, C., & Shwartz, A. (2018). Bridging the nature gap: can citizen science reverse the extinction of experience?. *Frontiers in Ecology and the Environment*, *16*(7), 405-411. https://doi.org/10.1002/fee.1826

Shah, H. R., & Martinez, L. R. (2016). Current approaches in implementing citizen science in the classroom. *Journal of Microbiology & Biology Education*, *17*(1), 17-22. https://doi.org/10.1128/jmbe.v17i1.1032

Sheppard, S. A., Turner, J., Thebault-Spieker, J., Zhu, H., & Terveen, L. (2017). Never too old, cold or dry to watch the sky: A survival analysis of citizen science volunteerism. *Proceedings of the ACM on Human-Computer Interaction*, *1*(PSCW), 1-21. https://doi.org/10.1145/3134729

Shume, T. J., & Blatt, E. (2019). A sociocultural investigation of pre-service teachers' outdoor experiences and perceived obstacles to outdoor learning. *Environmental Education Research*, *25*(9), 1347-1367. https://doi.org/10.1080/13504622.2019.1610862

Smith, P. S., Goforth, C., Carrier, S., Hayes, M. (2025). An emerging theory of school-based citizen science. *Citizen Science: Theory and Practice, 10*(1), 1-10. https://doi.org/10.5334/cstp.755

Szczytko, R., Carrier, S., & Stevenson, K.T. (2018). Impacts of outdoor environmental education on attention, behavior, and learning outcomes for students with emotional, cognitive, and behavioral disabilities. *Frontiers,* 3, 46. DOI: 10.3389/feduc.2018.00046.

Tan, E., & So, H. J. (2019). Role of environmental interaction in interdisciplinary thinking: from knowledge resources perspectives. *The Journal of Environmental Education*, *50*(2), 113-130. https://doi.org/10.1080/00958964.2018.1531280

Thorburn, M., & Marshall, A. (2014). Cultivating lived-body consciousness: Enhancing cognition and emotion through outdoor learning. *Journal of Pedagogy*, *5*(1), 115-132. https://doi.org/10.2478/jped-2014-0006

Vanhorn, S., Ward, S. M., Weismann, K. M., Crandall, H., Reule, J., & Leonard, R. (2019). Exploring active learning theories, practices, and contexts. *Communication Research Trends*, 38(3), 5-25. https://scholarcommons.scu.edu/crt/vol38/iss3/1

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