

Children's Coordination of Linguistic and Numeric Units in Mathematical Argumentative Writing

Karl Wesley Kosko ^{1*}, Rashmi Singh ¹

¹ Kent State University, USA

* CORRESPONDENCE: ✉ kkosko1@kent.edu

ABSTRACT

Mathematical argumentation and proof has long been identified with algebraization. Much literature discusses the relationship between the two, but with little specificity on how particular semiotic features in argumentation relate to coordination in early algebra. Further, there is a particular lack of research on this topic in the elementary/primary years of schooling. The present study examines how children's unit coordination in early algebra (particularly the concepts of equivalence and multiplicative reasoning) co-occurs with their coordination of grammatical information units. Coordination of information units was examined through reference use via the semiotic tool of detailing. Results suggest that second and third grade students who coordinate reference chains to support a mathematical claim in their argumentative writing tend to have higher multiplicative reasoning and conception of equivalence scores on several tasks. However, features of certain tasks may influence whether and how such unit coordination interacts with reference use.

Keywords: mathematical argumentation, early algebra, unit coordination

INTRODUCTION

Research examining the development of argumentation and proof in schooling has predominately focused on the relationship between generalization in argument and algebraic contexts for problem solving (e.g., Fosnot & Jacob, 2009; Morris, 2009; Tall et al., 2012). Findings from such research has generally found that when students are able to engage in more sophisticated forms of algebraization in elementary grades they are more likely to engage in clarifying claims and specifying justifications, and vice versa (Blanton & Kaput, 2011; Russell, Schifter, & Bastable, 2011). Despite such findings, the development of mathematical argument in the elementary grades is relatively undertheorized. Although various descriptions and models of how children engage in mathematical argumentation and/or proof have been proposed (Blanton & Kaput, 2011; Krummheuer, 2007; Morris, 2009; Tall et al., 2011), the vast majority of such descriptions focus exclusively on generalized arguments, particularly in the upper elementary and middle grades. This bias towards generalization is understandable, given the necessity of generalization in the desired development of proof processes in schooling. However, such focus has neglected aspects of children's mathematical argumentation that may precede inference and generalization, particularly in early elementary grades. Furthermore, though much of the literature base provides evidence for a connection between early algebra tasks and mathematical argument in the elementary grades, there has been little to no focus on the specific semiotic features (such as language, signs, or diagrams) that interrelate with algebraization.

The present study examines the relationship between researchers' cognitive models for children's algebraic reasoning and children's enactments of mathematical argument. Specifically, we examined children's

Article History: Received 29 November 2018 ♦ Revised 31 December 2018 ♦ Accepted 1 January 2019

© 2019 by the authors; licensee Modestum Ltd., UK. Open Access terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) apply. The license permits unrestricted use, distribution, and reproduction in any medium, on the condition that users give exact credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if they made any changes.

coordination of mathematical information through their use of reference, an action Kosko (2016) defined as *detailing*, and investigated how such actions related to their demonstrated multiplicative reasoning and understanding of equivalence. Thus, the purpose of the present study is to investigate whether children demonstrating certain early algebra actions are more likely to engage in detailing in their mathematical argumentative writing across multiple tasks.

MATHEMATICAL & ARGUMENTATIVE WRITING

Writing in mathematics has been advocated in elementary grades as early as the 1980s (Cockroft, 1982) and various studies it has a positive effect on students' mathematical thinking and learning (Bicer, Perihan, & Lee, 2018). Research on more general argumentative writing suggests that students as young as five can make simple arguments (Macedo, 2011). Studying writing of elementary students, O'Halloran (2014) noted that "counter to claims that second graders are not capable of formulating Reasons, nearly half of the second graders in this study did so successfully" (pp. 225-226). Focusing on mathematical argumentation, Nordin and Boistrup (2018) observed that many grade 3-5 Swedish children were also able to incorporate reasons and rationales to support claims in whole class arguments. However, Knudson (1992) found that many students do not incorporate backings in general arguments. This is supported by evidence from Kosko & Zimmerman's (in press) analysis of K-3 students in the United States. Specifically, Kosko & Zimmerman (in press) found that while some students did provide rationales as early as Kindergarten and first grade, this did not begin to show prominence until second and third grade. Knudson's (1992) work suggests that it is still relatively uncommon in upper elementary grades. A central question, particularly in the realm of mathematical argument, is why such differences in argumentative writing emerges in elementary students, and why might it persist? The present study focuses on the first part of this question.

THEORETICAL PERSPECTIVE

There is a useful distinction between a mathematical argument as a product, mathematical argumentation as a process, and discursive objects that are mathematically argumentative in nature (Cirillo et al., 2015). The present study focuses specifically on discursive objects (in the form of children's writing) that are mathematically argumentative, but which may not have all features of what some would refer to as a mathematical argument. Thus, mathematical argumentative writing (MAW), as described in the present study, includes students' written language, mathematical symbols, diagrams, etc. that convey a mathematical claim. Children can use a number of semiotic resources to convey such a claim, and to warrant its validity. In prior work, we have discovered that variations in how mathematical claims are warranted in children's MAW illustrate different levels of sophistication in how mathematical meaning is conveyed (Kosko, 2016; Kosko & Zimmerman, in press).

Described in detail in the following sections, the observed variation in sophistication appears to center on how reference and operate with mathematical objects in a semiotic manner. At a surface level, this distinction can be easily dismissed as whether a child references certain numbers or not. However, there is a great deal of nuance in the coordination of mathematical reference akin to Steffe and colleagues' (Hackenberg, 2010; Steffe, 1994; Steffe & Cobb, 1986) descriptions of unit coordination. Various theoretical perspectives and analytical tools have been introduced to study mathematical argumentation and proof including, but not limited to, Toulmin's (1958/2003) scheme (Krummheuer, 2007), Halliday and Matthiessen's (2004) Systemic Functional Linguistics (Morgan, 2006), and the Writing Across the Curriculum movement (Bossé & Faulconer, 2008). Although research stemming from such perspectives has provided useful findings, these perspectives are limited in how they have thus far drawn connections between representations associated with linguistic signs and cognitive-situated actions. In the pages that follow, we articulate our theoretical perspective in detail. In doing so, we attempt to limit the theoretical jargon introduced to only those terms necessary. In various places, we refer to more specific terms for readers interested in drawing further connections, but do so without necessarily utilizing those more specified terms otherwise.

Semiotics of Detailing References

Peirce (1903/1998) defined *argument* as a sign that synthesizes various premises and includes inference towards some general claim. Although Peirce's description of argument is a discursive object we would not expect to commonly see in children's MAW, his conception of argument as a sign that abstracts various other signs is particularly useful. Specifically, Peirce (1908/1998) described a hierarchy of signs that can stand on

their own as a set of propositions or can be situated within other signs (such as an argument). This consideration of various propositions as conveying a singular sign is particularly useful, and one that is fundamental in mathematics. For example, $a + b + c$ conveys three distinct elements in relation to another, but also conveys one expression. The abstraction of subsets of propositions metaphorically as another proposition is facilitated by various semiotic resources. In the aforementioned example, the addition signs are semiotic resources that convey a particular relationship between two different variables. However, the presence of multiple addition signs points to other conveyed meanings (commutative property, associative property, etc.) that signal that not only are there three variables, but a singular expression.

Though various semiotic resources are operationalized by students in their mathematics, the present study focuses on the particular tool referred to as *detailing*. Kosko (2016) describes detailing as the operationalization of given information through the construction of a reference chain through the warranted propositions supporting a claim. According to Halliday and Matthiessen's (2004) approach to SFL, reference serves the role of establishing cohesion for a text (or discourse). References may be *exophoric*, in which a nominal element (colloquially serving the role of noun) is introduced as new information. Once introduced, references are *endophoric*, since they are recoverable from within the text, and may be referenced *anaphorically* (i.e., using it to point to a more specified referent). In this manner, reference chains can be used in argumentative and factual texts to establish *information units*. Information units are semiotic units that connect given (endophoric) and new (exophoric) referents to convey information¹. Thus, information units can be considered as a type of proposition². We build off of Halliday and Matthiessen's (2004) descriptions of information units and reference to define detailing. To summarize, detailing establishes reference chains between endophoric (*given information* within a text) and exophoric (*new information* to the text) references in two or more information units (i.e., a form of proposition) that allow for them to be conveyed holistically as a copulative proposition (Liszka, 1996) – a set of propositions metaphorically serving the role of one proposition.

Consider the example proof for the sum of two even integers is even. A completed proof may state that: Let n and m be even integers, such that $n = 2x$ and $m = 2y$. So, $2x + 2y = 2(x + y)$, which is divisible by 2. Thus, the sum of two consecutive even integers is even.

The proof includes several information units. For example, n and m are provided initially as exophoric references and identified as even integers. According to Halliday and Matthiessen (2004), the inclusion of information units without endophoric references is common since “discourse has to start somewhere, so there can be discourse-initiating units consisting of a New element only” (p. 89). In the second information unit, expressions $2x$ and $2y$ also serve as exophoric references, and are conveyed in reference to the, now, endophoric references of n and m . Specifically, these exophoric references are transformations of n and m , where n and m were initially exophoric reference but, following its establishment within the proof are considered as endophoric for the remainder of the proof. Each subsequent information unit further operationalizes these referents in such a manner that allows for each information unit to connect with the other. For example, various propositions in the proof are connected with the referent n . Yet, n is used as a singular referent only in the initial information unit. Every other incorporation of n is done through an operationalization of n such that the new reference is, technically, a different nominal object.

The prior example may appear over-analyzed to some, as the transformation of n and m across the proof appears obvious to those with mathematical orientations. However, by distinguishing between references that are endophoric and exophoric, we are able to describe when and how mathematical meaning is conveyed differently. Detailing occurs when references across different information units are connected to metaphorically convey a singular proposition. However, we have observed that children may do much of the work to convey such a copulative proposition, but their detailing is incomplete (Kosko, 2016; Kosko & Singh, 2016b). This is often due to the introduction of new information (an exophoric reference) that does not connect with any endophorically referenced information. Thus, by focusing not only on how students coordinate their mathematical references through the semiotic resource of detailing, we are able to see how certain propositions are synthesized, or why they are not.

¹ The information unit is parallel to the clause as a grammatical unit. However, Halliday and Matthiessen (2004) note that “a single clause may be mapped into two or more information units; or a single information unit into two or more clauses” (p. 89), and therefore caution against identifying clauses and information units as identical.

² Halliday and Matthiessen (2004) define propositions as when clauses involve the exchange of information. However, information units are not always identifiable as clauses, and therefore the distinctions between clause and information units in regards to propositions is less clear.

Detailing may be more easily observable in the example proof just provided than in other forms of mathematical communication, simply due to the abstractions incorporated. As will be demonstrated in this paper, however, the examination of detailing is easily applied to children's MAW. The examination of detailing originated with analysis of children's MAW (Kosko, 2016; Kosko & Zimmerman, in press). As defined, detailing facilitates conveyance of reference chains which link together endophoric and exophoric references in two or more information units. Thus, detailed reference chains serve the role of synthesizing information units into a copulative proposition.

Early Algebra and Colligation in Mathematical Argument

Tall et al. (2011) suggest that as students explore various operations on mathematical objects, they observe patterns which eventually are described as rules of arithmetic. Describing and examining these rules is said to lead to the development of deductive reasoning and mathematical argument. Various studies have observed connections between generalizing involved in early algebra and improved precision in claims and conjectures (e.g., Blanton & Kaput, 2011). Specific facets of early algebra lead towards this generalizing, such as looking for and describing patterns and examining the structure involved in arithmetic operations. The latter includes, among other concepts, understanding of equivalence and multiplicative reasoning, of which the present study focuses particularly. As with other areas of early algebra, equivalence (Fosnot & Jacob, 2009) and multiplicative reasoning (Morris, 2009) have been observed to co-occur with more sophisticated argumentation. Specifically, various authors (Fosnot & Jacob, 2009; Morris, 2009) describe the importance of students examining the relationships between mathematical objects, or between two or more collections of such objects. In this paper, we argue that such a focus can be extended to explain the actions of students' MAW in tandem with their demonstrating an ability to abstract other mathematical objects.

The present study posits a more specified view of the interplay between early algebraization and mathematical argumentative objects. Specifically, children's detailing in the formation of copulative propositions points to an ability to abstract multiple meanings (i.e., information units) as one synthesized information unit. In a very similar fashion, children's coordination of quantitative units in ways mathematics education researchers would consider as algebraic also points to an ability to abstract multiple meanings as one. Although these two types of coordination point to the same ability to abstract meaning via information units, they are not necessarily identical, but we conjecture that they are similar enough to co-occur. To help articulate this co-occurrence, we briefly describe such interplay for the context of multiplicative reasoning and conception of equivalence.

Although various models of multiplicative reasoning populate the literature, we consider it from the perspective of scheme theory, with particular attention to unit coordination (Hackenberg, 2010; Steffe, 1994). Initial multiplicative unit coordination involves the coordination between two units, such as through skip counting. Thus, solving the task 19×3 , a student could skip count 19 three times (or 3 nineteen times). As students begin to coordinate three levels of units, they begin disembedding numbers. Such a student solving the task 19×3 might recognize that 10×3 is 30, and then count-down-from 30 to 27 to add 27 and 30 together. Yet, a student with a more sophisticated coordination of three levels of units may solve 19×3 by recognizing two different math facts (10×3 and 9×3) and adding the products together.

Like multiplicative reasoning, our view of equivalence also involves unit coordination. Specifically, students' conception of equivalence, or their meaning of the equals sign, is typically described as either operational or relational (Byrd, McNeil, Chesney, & Matthews, 2015; Falkner, Levi, & Carpenter, 1999). An *operational* view of equivalence interprets the equals sign as a sign to add/subtraction or to produce the answer. A *relational* view of equivalence interprets the equals sign as identifying the "sameness" of two expressions on each side of the sign. An operational view of equivalence relies on additive reasoning, in which students consider one unit at a time. Yet, a relational view requires students to consider an expression, which can include multiple units as a unit itself. For example, the expression $a + b$ includes the elements a and b as units in and of themselves, but the expression itself is also a unit that happens to contain other units. From an SFL perspective, this interpretation is described as nominalization, but from a neo-Piagetian perspective, it is described as the coordination of two levels of units.

As previously described, detailing involves the coordination of multiple information units into what Pierce (1903/1998) refers to as a copulative proposition. Both multiplicative reasoning and relational conceptions of equivalence, at their initial levels, involve the coordination of two different units as if they are one. In the case of detailing, exophoric and endophoric references are linked together across information units in a detailed reference chain. Likewise, a child who has coordinated six 3s into a composite of 18 has referenced, through

their actions, the unit 3 endophorically upon each iteration to find the resultant 18 (3, 6 as two 3s, 9 as three 3s, etc.). Building from this perspective, we argue further that such activity serves the role of unit coordination with quantity that is similar to detailing in a predominately linguistic context (i.e., writing). Further, we suggest that the ability to abstract in such ways in the quantitative and linguistic contexts may point to a more general ability to abstract meaning.

RESEARCH QUESTIONS

We hypothesize that a more sophisticated ability to coordinate units, often demonstrated via different aspects of early algebra, associate with an increased likelihood to engage in detailing, and vice versa. To test this hypothesis, we compared second and third grade students' scores on two early algebra assessments (multiplicative reasoning and equivalence) with their demonstrated detailing on six argumentative tasks embedded in early algebra. Specifically, we asked the following research question:

Do children who detail in mathematical argumentative writing have higher scores on early algebra assessments, particularly in regards to their ability to coordinate units multiplicatively and in regards to equivalence?

METHODS

Sample and Measures

Data were collected from 168 second and third grade students in two suburban schools in a Midwestern U.S. state in May 2015. Second grade students were enrolled in four different teachers' classrooms ($n = 76$) and third grade students were enrolled in three different teachers' classrooms ($n = 92$). Further, 48.2% of students identified as female and 51.8% identified as male. No descriptive data regarding race, ethnicity or socio-economic status were collected. Participants completed two packets: one packet included two assessments for early algebra (multiplicative reasoning and conception of equivalence) and the other packet included six mathematical argumentative writing tasks³.

Packet 1: Early Algebra Assessments

The first packet was administered one week prior to the second packet and included items assessing equivalence and multiplicative reasoning. The equivalence portion included 15 items adapted from Fyfe, DeCaro, and Rittle-Johnson (2014) with eight missing value tasks (e.g., $3+4=\square+5$) and seven true/false items. Nine of the 15 items were designated as requiring a relational interpretation of the equals sign ($5+\square=6+2$: True or False $67+86=68+85$), with other items requiring a more operational interpretation ($7=\square+3$; True or False $7+6=0$). Thus, higher scores on the assessment are interpreted as having a more relational interpretation of equivalence while lower scores are interpreted as having a more operational interpretation of equivalence. As with prior researchers' application of the assessment, we found it to have sufficient internal reliability ($\alpha = .89$). Descriptive statistics ($M = 10.70$, $SD = 4.00$) indicate that approximately half of students demonstrated a more relational interpretation of equivalence. However, half of students scored 12 or less, suggesting their interpretation of equivalence was not consistently relational.

The assessment on multiplicative reasoning was developed by Kosko and Singh (2018), and is based on the notion that more abstracted unit coordination is representative of more sophisticated multiplicative reasoning (Hackenberg, 2010; Norton, Boyce, Ulrich, & Phillips, 2015). The 12-item assessment used length models to assess different aspects of unit coordination (see **Figure 1** for examples with included student work). For example, two items on the assessment examined whether students could successfully coordinate two-levels of units by skip-counting non-1 units (i.e., given an iterable unit of 3, being able to determine a length of 18). Other items assessed iterating, partitioning, and disembedding of various units. Higher scores represent more sophisticated unit coordination (i.e., more sophisticated multiplicative reasoning). Our use of the assessment showed sufficient internal reliability ($\alpha=.79$). Descriptive statistics of the assessment ($M = 4.88$, $SD = 2.88$) indicate that approximately half of participants could coordinate units to at least skip count, with third grade participants demonstrating this more frequently.

³ Preliminary analyses of the same data are reported in published conference proceedings (Kosko & Singh, 2016a, 2016b).

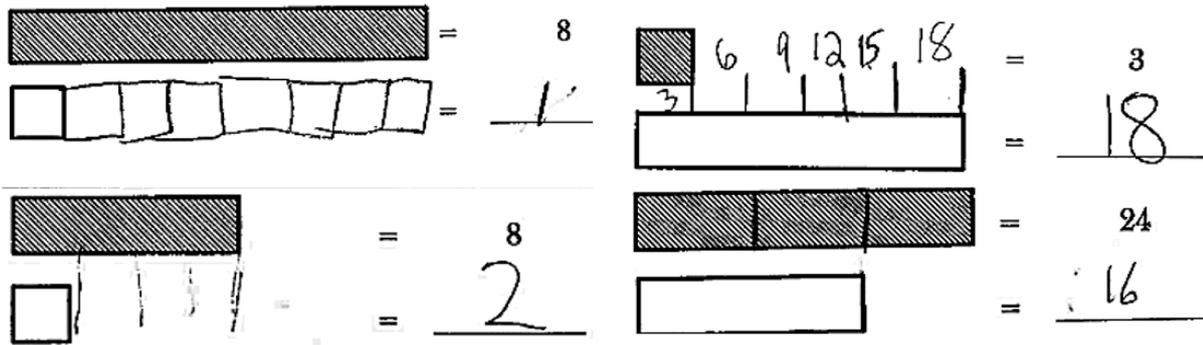


Figure 1. Example items from Kosko & Singh's (2018) multiplicative reasoning assessment

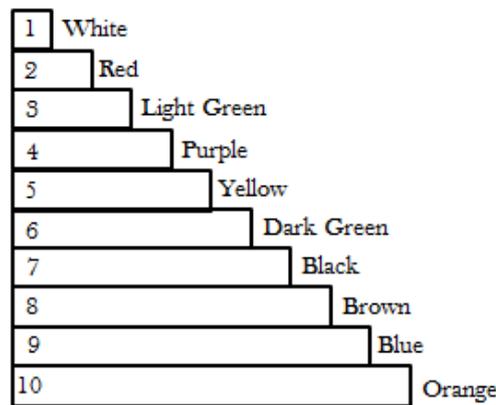


Figure 2. Cuisenaire rods are rods of color-coded lengths

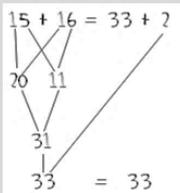
Table 1. Mathematical argumentative writing tasks focusing on numeric relationships

Description	Canonical Mathematical Equations(s)
Task 1 <i>If a red rod is 2 long, how long is a brown rod?</i>	$2+2+2+2=8$ $4 \times 2=8$
Task 2 <i>Paul says 2 red rods make a yellow rod. This is impossible because...</i>	$2+2 \neq 5$ $2 \times 2 \neq 5$
Task 3 <i>If a red rod is 5 long, a yellow rod can't be 9 because...</i>	$5+5 \neq 10$ $2 \times 5 \neq 10$

Packet 2: Mathematical Argumentative Writing Tasks

MAW tasks included six tasks with three incorporating length model representations involving quantitative comparisons (see Table 1) and three focusing on arithmetic involving relationships between equations and expressions (see Table 2). The order of tasks was varied so that half of participants completed tasks 1 through 6 in sequence and the other half completed them in reverse order. Participants were provided with Cuisenaire rods as a resource for completing the tasks. Cuisenaire rods are color coordinated length-based models used to represent relationships between different numbers. Shown in Figure 2, each color represents a specific length in centimeters, but can be referenced mathematically as any number (e.g., referring to a yellow rod as $\frac{1}{2}$).

Table 2. Mathematical argumentative writing tasks focusing on arithmetic relationships

Tasks	
Task 4	<p>Omar was asked to find the answer for $19 + 19$. This is what he wrote on his paper:</p> $19 + 19$ $20 + 20 = 40$ $40 - 2 = \textcircled{38}$ <p>Explain how Omar solved the problem and why it does (or doesn't) work.</p>
Task 5	<p>Tom told a friend that "$15 + 16 = 33 + 2$." His friend didn't believe him, so Tom wrote the following:</p>  <p>Do you agree with Tom? Write to Tom and explain why or why not:</p>
Task 6	<p>Vanessa was asked to find the answer for $38 + 22$. This is what she wrote on her paper:</p> $38 + 22$ $30 + 20 + 10$ 60 <p>Vanessa's answer is correct. Look at how Vanessa solved the problem and decide whether this will work for <u>any addition problem</u>. Explain to Vanessa why her strategy will (or won't) always work.</p>

The length model tasks focused on quantitative comparisons between different numbers. For example, students completing Task 1 are provided with the given information that a red rod is 2 long. This implicitly tasks the child with iterating four 2s (four red rods) to find the total length of a brown rod (8 long). The child could do this by adding all four 2s or multiplying 4×2 . However, a child may also choose to use white rods (length 1) or some other combination to represent alternate relationships. Task 2 provides a similar context, but notes that the comparison is not possible and tasks the child to explain. Task 3 is nearly identical to Task 2 in how the physical manipulatives could be manipulated, but designates that a red rod is now 5 long. Each of these tasks have been used in prior work with children's MAW (Kosko, 2016; Kosko & Zimmerman, in press). Of particular note is analysis conducted by Kosko (2016) observing Task 3 solicited more detailing from students than Task 2. In examining clinical interview data with two students completing these tasks, Kosko (2016) observed that across tasks, how students considered the given information appeared to interact with their multiplicative reasoning on the tasks.

Tasks 4 through 6 included more canonical representations of mathematics in the form of expressions and equations (see **Table 2**). The tasks focused on varying arithmetical relationships. For example, in Task 4, the fictional student Omar added 1 to each 19 and then subtracted them both at the end of their arithmetic. Similarly, Task 6 involves use of the decomposing 38 and 22 into $30 + 20 + 8 + 2$ (a step not explicitly shown). Although these two tasks involved arithmetic manipulation through composing and decomposing numbers, prior work with a small group of second grade students on these tasks revealed a reliance on an understanding of equivalence. Specifically, those students who demonstrated an emerging relational interpretation could make sense and work with these tasks while students demonstrating an operational interpretation did not relate one expression with another. Thus, while Tasks 1 through 3 appeared to tacitly relate to multiplicative reasoning, Tasks 4 through 6 appeared to tacitly relate to conception of equivalence. However, we do not view these previously observed relationships as excluding other potential interactions with various mathematical concepts.

Table 3. Descriptive Statistics for Coding Reference Use in Tasks

	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6
Kappa	.81	.43	.67	.90	.64	.76
No Reference	.28	.06	.31	.16	.16	.07
Simplistic Reference	.01	.01	.01	.19	.03	.38
Incomplete Detailing	.28	.41	.46	.27	.43	.34
Detailing	.43	.52	.22	.38	.39	.21

Analysis

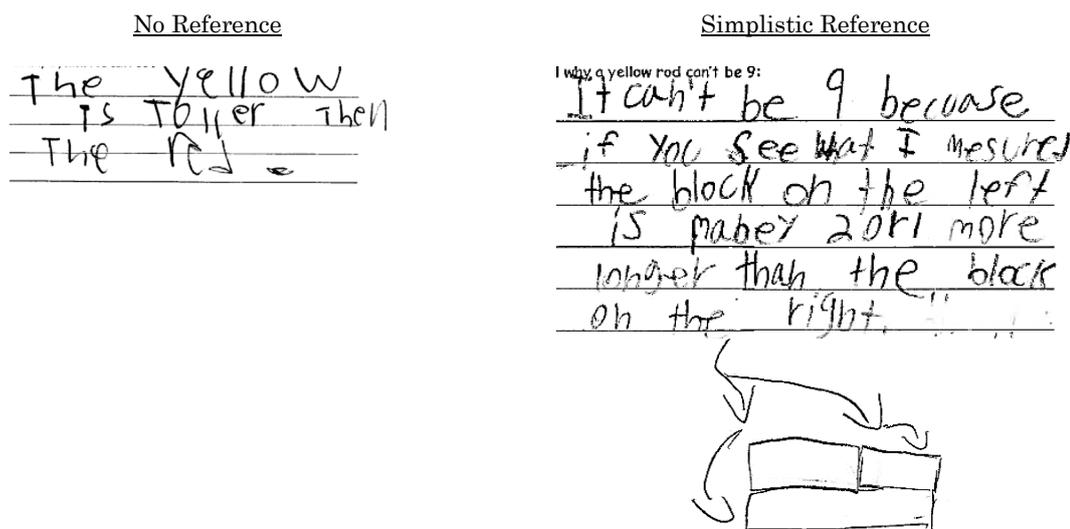
The present study incorporated a data-transformation variant of convergent mixed methods design (Creswell & Plano Clark, 2011). A convergent mixed methods design uses both quantitative and qualitative methods during the same phase of the research process in order to triangulate the results of a single topic. The data-transformation variant emphasizes the quantitative strand of research by merging data through a transformation process. Specifically, Creswell and Plano Clark (2011) state that “after the initial analysis of the two data sets, the researcher uses procedures to quantify the qualitative findings” (p. 81). This allows for mixing of the design through quantification of qualitative findings.

The qualitative strand in the current design included the use of SFL to examine students’ use of reference within their MAW for each task. This involved two facets of examination. First, both authors examined the tasks for mathematical referents conveyed and identified themes for how reference use was elicited from such tasks. We distinguished information units from each other to allow for a finer grain analysis before examining for the use of reference and information units in students’ MAW. We transformed these themes into codes, with example reference chains that appeared most common in independent analysis. Next, approximately 10% of the 168 participants’ writing samples, across tasks, were examined by both authors in an iterative process. Specifically, we co-examined a portion of writing samples in concert with our detailed reference chains coding scheme to look for trends in what information units emerged and in what ways. The four themes that emerged in regards to reference use are identified as categorical codes below:

0. No reference to given(s) provided, and no parts of a reference chain conveyed.
1. Simplistic reference: Given(s) are endophorically referenced, but no connections between references across information units is observed.
2. Incomplete detailing: both endophoric and exophoric references are conveyed to establish a reference chain between information units. However, the reference chain either has missing connections between references or introduces extraneous exophoric references.
3. Detailing conveyed via a complete detailed reference chain.

Specific examples of each theme are provided in the description of results and findings. Themes are presented here to report on the inter-rater reliability following analysis of the remaining 90% of writing samples. Following co-analysis of an initial 10% of MAW artifacts, both authors independently analyzed the remaining 90% of writing samples, by task, for presence of different themes. Interrater reliability was assessed using the Kappa statistic (Landis & Koch, 1977), before both authors reconciled coding (see **Table 3**). Kappa statistics calculated for each task range from moderate to near perfect in size (Landis & Koch, 1977), suggesting that the analysis of themes has sufficient reliability. Additionally, the frequency of each theme observed by tasks is illustrated in **Table 3**. Descriptive statistics suggest that the themes emerged differently dependent by task, which is supportive of prior findings (Kosko, 2016; Kosko & Zimmerman, in press).

The quantitative strand of the mixed methods analysis involved comparison of scores on early algebra assessments (i.e., conception of equivalence and multiplicative reasoning) in regards to children’s use of reference within their MAW. In order to examine how students’ use of reference was related to their early algebraic reasoning, we used Analysis of Variance (ANOVA) to examine the differences, by task, between students’ reference use and their scores on the equivalence and multiplicative reasoning assessments. This required a quantification of the themes (see **Table 3**) into nominal variables. However, given that the observed frequencies for simplistic reference were extremely small in certain tasks, no reference and simplistic reference were combined into a singular nominal variable (no or simplistic reference). Thus, three nominal variables were used and treated as groups for an ANOVA of each task: no or simple reference; incomplete detailing; detailing.



“The yellow is taller than the red” (#212).

“It can't be 9 because if you see what I measured the block on the left is maybe **2 or 1 more longer** than the block on the right.” (#344)

Figure 3. Examples of No Reference and Simplistic Reference use on Task 3

FINDINGS AND RESULTS

Qualitative Findings

Qualitative analysis revealed four distinct patterns in how children primarily operationalized mathematical references within their MAW: no reference used; simplistic reference used; incomplete detailing; detailing. For simplicity, we describe these themes in the context of Task 3, which prompted “if a red rod is 5, a yellow rod can't be 9 because...”. *No reference* signified that a student did not refer to mathematically quantitative information provided within the text of the task. For example, student #212 wrote the claim that “the yellow is taller than the red” which does reference the rods provided in the task, but does not provide a reference to quantity (see **Figure 3**). To be clear, the child makes what Kosko & Zimmerman (in press) would refer to as a mathematical claim. However, the mathematics referenced is qualitative in nature. Students' writing within the no reference classification typically either provided only a mathematical claim or a qualitative recounting of what they did (i.e., stating they lined up the rods, but not conveying a manipulation of number).

Simplistic reference signified that a student references mathematically quantitative information, but does not create reference chains between information units. Student #344 in **Figure 3** provided a writing sample in this classification stating that “it can't be 9 because if you see what I measured, the block on the left is maybe 2 or 1 more longer than the block on the right.” The student then drew an arrow to a drawing of one longer block aligned with two stacked blocks. Although it may be apparent to a classroom teacher how the child was reasoning through the task, there are several gaps in the mathematical information conveyed in their writing. The child conveys the claim that the yellow can't be 9, but otherwise conveys a manipulation of the Cuisenaire rods, and not the numbers they signify. The information unit “the block on the left is maybe **2 or 1 more longer** than the block on the right” does convey a quantitative relationship between two rods aligned with the yellow. One can infer that the child is referencing a red rod (length 2) and a light green rod (length 3), but this is not explicitly conveyed. As such, the reader must make their own connections between the various information units provided.

The samples shown in **Figure 3** demonstrate the lowest entry points to what is referred to here as MAW. As such, they are products of the way a child communicates mathematically, but may not tell us everything about the mathematical processes the child engages. What can be assessed from these artifacts is how the child conveys mathematical information in this particular instance. By not creating a reference chain that connects quantity in one proposition with quantity in another, the MAW samples in this classification convey mathematical elements, but do not convey what is done mathematically with these elements.

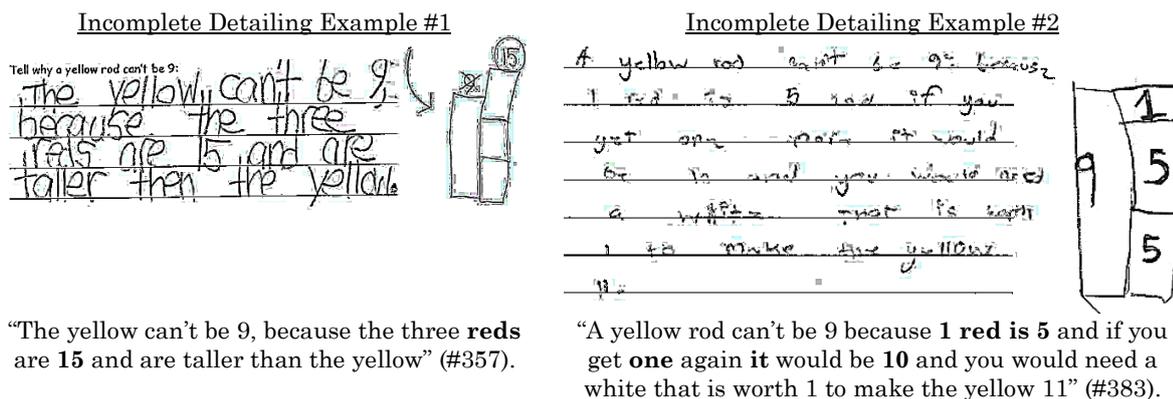


Figure 4. Examples of two students’ incomplete detailing on Task 3

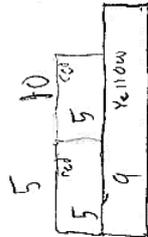
Incomplete detailing occurred when students provided some form of reference chain to warrant a provided claim, but the reference chain was incomplete. Specifically, a complete detailed reference chain conveyed how given mathematical information (endophoric reference) was transformed to the point that a clear and complete connection was made to warrant the provided claim. Incomplete detailing typically took the form of a procedural description of the mathematics, with some misalignment of how reference was utilized (for a more complete description of this difference, see Kosko & Zimmerman, in press).

Student #357 provides one useful example of incomplete detailing. The child makes the claim that a yellow cannot be 9 because “the three reds are 15” (see Example 1 in **Figure 4**). Although the child does not explicitly state that a red is 5, the given information is ellipsed. Specifically, the information unit “three reds are 15” endophorically references red as 5 by stating that three of them is 15. This information unit also conveys the transformation of 5 into 15. Although the child does an effective job conveying how they transformed 5 into 15, their final information unit that the 15 is “taller than the yellow” does not connect the prior information to support the claim that a yellow can’t be 9. Essentially, the child conveys that $15 > 9$. However, this information fails to establish that yellow cannot be 9 long. What is absent is an information unit conveying some sort of contradiction, which is a valid means of establishing why something cannot be. Thus, if the child had referenced 10, and not 15, for comparing to the yellow rod, they could have demonstrated the contradiction that $9 > 10$. By choosing 15 as a comparison point, student #357 does convey certain mathematical information and processes, but not in a manner that fully supports their claim.

Student #383 provides a useful example of incomplete detailing where exophoric references are introduced without appropriate connection to the references endophoric to the task (see Example 2 in **Figure 4**). The child makes the claim that a yellow can’t be 9 by stating the given endophoric reference that “1 red is 5.” The next information unit “if you get one again it would be 10” endophorically references a red as 5 with “one” and transforms it through an iteration of “again” to convey two reds as “10.” In such a manner, the student has provided evidence towards a reference chain that transforms given information towards warranting their claim. However, the child’s next information unit introduces an exophoric reference that is misaligned with the text, “a white that is worth 1” (i.e., for the current task, a white is half of a red 5, or 2.5 long). Since this particular exophoric reference does not connect with or represent a transformation of any endophoric reference provided in the given task, such reference of a white rod as 1 establishes an incomplete detailed reference chain. Phrased another way, by claiming that a yellow rod is not 9 because it actually is 11 long conveys a false claim established partially with an extraneous exophoric reference. Such instances of incomplete detailing can be referred to as provide examples of getting the right answer for the wrong reasons.

Evidence of detailing included complete detailed reference chains, but also included two distinct forms of reference chains. **Figure 5** provides two examples that reflect each form of detailing observed for Task 3. The first form of detailing, illustrated by example 1, represented the overwhelming majority of detailing (nearly all examples). Specifically, student 319 references the given information unit by endophorically referencing “1 red one is 5.” The next information unit includes multiple clauses to convey that by placing another 5 it produces a length of 10. These two information units were common across all detailing observed, and is even presented diagrammatically in example 2 in **Figure 5**. However, the final information unit provided by student 319 provides a necessary justification for why such information warrants the claim provided. The child states that “it’s smaller than the yellow” which conveys an observation that would be written as canonically

Detailing Example #1
 The yellow rod can't be 9 because 1 red one is 5 and that does not take it all up so you put another 5 and it's still not big enough but it's basically 10 and it's smaller than the yellow so the yellow can't be 9.



“The yellow rod can't be 9 because **1 red one is 5** and **that** does not take it all up so you put **another 5** and **it's** still not big enough but **it basically 10** and **it's** smaller than the yellow so the yellow can't be 9” (#319).

Detailing Example #2
 Tell why a yellow rod can't be 9:
 The yellow rod can't be 9 because $2\frac{1}{2}$ red blocks make $12\frac{1}{2}$.
 (red block) block block a red block

“The yellow block can't be nine because $2\frac{1}{2}$ red blocks make $12\frac{1}{2}$ ” (#343).

Figure 5. Examples of two students detailing in Task 3

as $9 > 10$. The information conveyed is a contradiction, which is a valid means of establishing why something cannot be (i.e., a yellow rod cannot be 9).

The second example in **Figure 5** illustrates a similar reference chain with a key distinction regarding one information unit. Student 343 includes a diagram that represents two reds and a half of a red. This diagram is referenced within the written description as $2\frac{1}{2}$ red blocks. The given information is endophorically referenced in the information unit that transforms $2\frac{1}{2}$ red blocks to “make $12\frac{1}{2}$.” Thus, the key information unit for student 343 is that $9 \neq 12\frac{1}{2}$. As with the key information unit conveyed in example 1 in **Figure 5**, this information unit also conveys a contradiction. Given that participants included second and third grade students, it is not surprising that this second form of detailing was less common for Task 3. However, what the two examples in **Figure 5** help illustrate is that, across tasks, detailing was evidenced with similar reference chains. Key differences in the manner in which detailing was observed tended to rely on a single information unit that justified prior warrants in particular ways to connect them with the claim.

Quantitative Results

Following the qualitative analysis of MAW samples, demonstrated reference use was quantified into a nominal variable for each task ($0 = no\ or\ simple\ reference\ use; 1 = incomplete\ detailing; 2 = detailing$). Next, ANOVA was conducted for each task with both assessments as the dependent variable (i.e., conception of equivalence & multiplicative reasoning). This resulted in 12 different F-statistics (6 tasks \times 2 assessments), since the tasks were considered individually. Therefore, we adjusted the p-values in order to avoid a Type I error (false positive). Although the Bonferroni correction is commonly used, we used Simes' (1986) adjustment to avoid a Type II error (false negative) which the Bonferroni correction is prone. The Simes (1986) correction is robust, maintaining its power when the calculated statistics are not independent (Sarkar & Chang, 1997). Therefore, the Simes (1986) correction is reported for all F-statistics presented in **Table 4** and **Table 5**.

Table 4. Results for Independent t-test for Conception of Equivalence and Detailing

	No or Simplistic Reference	Incomplete Detailing	Detailing	F	p-value ^a	Levene Statistic
Task 1	M = 8.94 SD = .81	M = 10.24 SD = .66	M = 12.22 SD = .36	10.45 ^b	.000	13.88***
Task 2	M = 9.38 SD = 1.91	M = 9.46 SD = .55	M = 12.21 SD = .36	10.77 ^b	.000	11.76***
Task 3	M = 8.80 SD = .65	M = 11.14 SD = .45	M = 13.33 SD = .39	21.63 ^b	.000	15.58***
Task 4	M = 8.64 SD = .67	M = 11.30 SD = .54	M = 12.38 SD = .40	13.05 ^b	.000	10.76***
Task 5	M = 8.68 SD = .91	M = 9.88 SD = .53	M = 12.96 SD = .31	18.391 ^b	.000	23.41***
Task 6	M = 9.82 SD = .55	M = 11.52 SD = .51	M = 11.93 SD = .65	4.10 ^b	.027	3.31*

* $p < .05$, ** $p < .01$, *** $p < .001$, ^aSimes Adjusted p-value, ^bWelches F-Statistic.

Table 5. Results for ANOVA for Multiplicative Reasoning and Reference Use

	No or Simplistic Reference	Incomplete Detailing	Detailing	F	p-value ^a	Levene Statistic
Task 1	M = 4.16 SD = .47	M = 3.88 SD = .36	M = 4.45 SD = .30	.83	.438	1.20
Task 2	M = 4.38 SD = 1.07	M = 3.76 SD = .33	M = 4.56 SD = .28	1.24	.338	.40
Task 3	M = 3.02 SD = .29	M = 4.34 SD = .29	M = 5.67 SD = .48	11.93 ^b	.000	3.02*
Task 4	M = 3.04 SD = .27	M = 4.00 SD = .39	M = 5.29 SD = .34	15.08 ^b	.000	4.76**
Task 5	M = 3.72 SD = .52	M = 3.85 SD = .28	M = 4.83 SD = .36	2.41	.112	2.04
Task 6	M = 3.68 SD = .28	M = 4.23 SD = .36	M = 5.23 SD = .48	4.59	.018	2.15

* $p < .05$, ** $p < .01$, *** $p < .001$, ^aSimes Adjusted p-value, ^bWelches F-Statistic.

In addition to the Simes (1986) correction to p-values, we tested for homogeneity of variance for each ANOVA using the Levene statistic (reported in **Table 4** and **Table 5**). In several comparisons, the Levene statistic was statistically significant suggesting that initial ANOVAs violated the assumption of homogeneity of variance and normality. In these particular comparisons, we used Welch's ANOVA to account for this violation (Lix, Keselman, & Keselman, 1996; Welch, 1951). Specifically, Welch's ANOVA is less affected by skewness in data and does not assume homogeneity of variance for unequal group sizes with at least 10 observations (Lix et al., 1996). In the case of customary ANOVA, we used Tukey's post hoc to compare differences in groups. In the case of Welch's ANOVA, we used Games-Howell's post hoc as it assumes unequal variance between groups. Distinctions between which statistics are reported is provided in **Table 4** and **Table 5**. Otherwise, we report on the findings in text with an effort on providing parsimonious descriptions.

Comparisons for conception of equivalence

Results for the Welch's ANOVAs comparing conception of equivalence scores for students' reference use are reported in **Table 4**. We found a statistically significant difference in scores across all tasks. Findings from Games-Howell post hoc analysis suggests that the relationship between equivalence scores and differences in reference use varied by task. Notably, for each task, mean scores for the equivalence assessment illustrate a trend from lower to higher across the MAW classifications for that task, and the associated variance tended to decrease. One possible explanation for this is that students with more sophisticated reasoning may have written in a manner below their capabilities. Given that language is choice (Halliday & Matthiessen, 2004), this is not surprising. In examining findings from the post hoc analysis, it is useful to consider this general trend with attention to where significant differences indicate specific points of departure. These points of departure likely signify a level of reasoning where the nature of the task required a certain level of explanation from students (a phenomenon observed by Kosko, 2016). These points of departure are highlighted in the following paragraph.

For Tasks 1 and 5, students who engaged in detailing had statistically significant higher scores than students who engaged in incomplete detailing or used no or simplistic reference. However, the latter two groups were statistically similar in regards to score. This indicates that although a hierarchical trend is apparent in students' conceptions of equivalence, there is a clear distinction between students' scores when engaging in detailing. Task 5 presents Tom's false claim that two mathematical expressions are equivalent, when they are not. Thus, it seems apparent that a relational conception of equivalence may be necessary to provide a detailed description for why Tom is incorrect. Task 1 tasks students with describing how long a brown rod is, given that a red rod is 2 long. Students engaging in detailing typically referred to this context either through repeatedly adding 2 or multiplying 2 four times. In either case, these representations were typically conveyed as identical to the brown (length 8) either linguistically or diagrammatically. For Tasks 4 and 6, students who engaged in incomplete detailing and detailing had statistically similar scores, but both had scores that were statistically significantly higher than students who used no or simplistic reference. Both Tasks 4 and 6 include application of the associative property, which requires considering expressions as units in their own right. Therefore, these results indicate that a relational understanding may be more critical in engaging in attempting to describe these tasks procedurally (i.e., incomplete detailing). For Task 3, students who engaged in no or simplistic reference use had statistically significant lower scores than both incomplete and complete detailing groups, and students who engaged in detailing had statistically significant higher scores than both other groups (i.e., no/simplistic reference < incomplete detailing < detailing). Similar to interpretations for findings on Tasks 4 and 6, it appears that to engage in each form of MAW classification, a different level of understanding equivalence may be necessary.

For Task 2, students who engaged in detailing had statistically significant higher scores than students engaged in incomplete detailing, but were statistically similar to students who engaged in no or simplistic reference use. At face-value, these findings appears to contrast those relationships already identified and discussed. However, it when considering the variance associated with scores in the no or simple reference classification (see **Table 4**), a more likely interpretation is that several students with higher demonstrated conceptions of equivalence wrote in very simplistic ways. In other words, the task may not have pressed them enough to provide more information.

These findings provide support for the hypothesis in this paper that a more sophisticated ability to coordinate units associates with an increased likelihood to detail. Given that the descriptive statistics across tasks illustrate a hierarchical trend (see **Table 4**), the post hoc analysis suggests that there was higher variance in scores for certain tasks and not others. Specifically, students' reference use varied across tasks. Recall that language is choice (Halliday & Matthiessen, 2004) and students chose to convey meaning in certain ways, consciously or not, in some tasks and did so differently in other tasks. Findings presented in **Table 4** suggest that one influence in such reference use may stem from the nature of the task and another influence may stem from students' mathematical reasoning. This may not seem particularly surprising to some. However, the findings and results here provide evidence that such distinctions can be observed empirically and in a manner that can be replicated to apply across fairly different tasks.

Comparisons for multiplicative reasoning

ANOVA and Welch's ANOVA statistics were calculated for comparisons between reference use and multiplicative reasoning scores (see **Table 5**). Findings from post hoc analysis indicate that scores were statistically significantly and different between reference use for three tasks (3, 4, 6), and statistically similar for the other three tasks (1, 2, 5). Phrased differently, it appears that Tasks 1, 2, and 5 may not have required as much multiplicative reasoning for students to engage in detailing, but Tasks 3, 4, and 6 did. Tasks 1, 2, and 5 essentially required that students demonstrate a capacity to perform basic arithmetic. While Task 5 does draw upon a relational conception of equivalence, there are varying degrees of relational understanding with some including multiplicative reasoning and others not (Singh & Kosko, 2016). For Task 3, the relationship with multiplicative reasoning seems apparent, given that students must consider a red rod as a composite unit that cannot be divided into 1s, and that they must operate within the task. For Tasks 4 and 6, findings from Singh & Kosko (2016) may be useful in explaining this observed relationship. Specifically, Singh & Kosko (2016) found that students capable of composing and decomposing expressions within expressions or equations also demonstrated higher multiplicative reasoning. In Tasks 4 and 6, students must engage in this process in order to relate the different expressions to one another. Thus, it may be that students need a certain level of multiplicative reasoning to engage in detailing.

DISCUSSION

The present study sought to examine whether second and third grade students' reference use to form copulative propositions, as demonstrated via their use of detailing, was related with their unit coordination involved in early algebra contexts in the context of equivalence and multiplicative reasoning. Findings support this hypothesis, but with specific caveats. First, statistically significant relationships were found across all tasks with conception of equivalence scores, but for only half the tasks for multiplicative reasoning scores. In addition to results of the post hoc analyses, these findings suggest that the manner in which students coordinate mathematical information in argumentative writing does interact with their early algebraic reasoning. However, the content knowledge solicited by the task is a key determinant of whether certain interactions may occur.

Although prior research on mathematical and argumentative writing suggests it is beneficial to students (Bicer et al., 2018), and that elementary students can and should engage in writing mathematically (Knudson, 1992; Kosko & Zimmerman, in press; Macedo, 2011), there has been little evidence in primary grades to explain the nuance in such relationships. Rather, while research has described how primary grades students engage in mathematical argumentation (Krummheuer, 2007; Nordin & Boistrup, 2018), alignment between their discursive and mathematical actions has not received much attention. The present study provides initial evidence towards establishing an empirical relationship between students' linguistic and mathematical actions. This is done by characterizing students' discourse moves through the context of linguistic unit coordination, and provides the field with a novel approach to examining argumentation.

Prior research on students' mathematics indicates that the nature of the task matters when soliciting mathematical descriptions (e.g., Kosko & Norton, 2012; Stein, Engle, Smith, & Hughes, 2008). Although confirmation of prior research is a necessary part of educational research, such a finding in and of itself is not particularly surprising. What is significant about the finding presented here is the manner in which it was examined. By considering mathematical argumentation as being composed of information units that can be synthesized or decomposed into other information units, the analyses presented here illustrate how these information units can be associated with unit coordination in particular mathematical contexts. This allows for a much finer grained analysis than what is generally used in studying the relationship between different kinds of students' mathematical actions.

Consider the first student's example of detailing in **Figure 5**. On the assessments, the child demonstrated a relational conception of equivalence (Score = 11 out of 15) and was capable of skip counting in multiplicative tasks (Score = 6 out of 12). Thus, she demonstrated an ability to coordinate and relate composites of units in different ways (within equations and multiplicatively). When examining her MAW in **Figure 5**, the same child demonstrates a similar ability with information units by conveying a detailed reference chain. Although there is no explicit application of equivalence in Task 3, a comparison of *two red rods* (length 10) as being less than a *yellow rod* (length 9) points towards the same sort of coordination that a relational understanding of equivalence does. The approach to analyzing students' MAW demonstrated in this paper allows for aligning information units in children's writing with their actions on other mathematical tasks not requiring argumentative writing. Such evidence supports our conjecture that children's detailing in MAW points to an ability to abstract multiple meanings also present in their children's algebraic coordination of quantitative units. However, the approach to analysis also allows for distinguishing how MAW tasks affected students' level of writing; a relationship also observed by Kosko (2016). The analytical approach of using information units to examine interactions and relationships with other mathematical actions is an important contribution of the present study, and one that can be applied to contexts in both written and spoken forms of mathematical argumentation.

Although the present study lends support to the claim that early algebraic reasoning is related to mathematical argumentation, and provides specific methodological means of further examining this connection, there are many areas in need of future study. Particularly, findings from the present study suggest that the role of the task children use to engage in MAW is critical, and certain features embedded in tasks may solicit certain aspects of children's content knowledge in ways that reflect in their MAW. Thus, future study is needed to study how mathematical meaning is conveyed from and solicited by such tasks. The present study included tasks with several differences, albeit with common content, and it may be worthwhile for future investigation to constrain the differences between tasks in order to better understand how particular differences affect students' argumentation.

The present study also extends the findings of others (Fosnot & Jacob, 2009; Morris, 2009) who suggest relationships between students' deductive reasoning and their application of equivalence and multiplicative relationships. Although prior research has suggested a connection between early algebra and mathematical argumentation, there has been little analysis regarding specific linguistic features evidence this connection. The results presented here provide much needed evidence regarding this connection.

The findings reported here are significant because for two primary reasons. First, it is still common for many educational researchers to consider the mathematical writing of young children as unsophisticated. However, the analysis presented here suggests there is a great deal of coordination and mathematics involved in second and third grade students' MAW. Second, while various researchers cite relationships between early algebra and argumentation, the connections cited are often at too large a grain size for certain types of analysis. Specifically, the presence of particular information units was consistently related to students' engagement with key early algebraic concepts. The present study presents both evidence and a theory-situated methodology for examining the fundamental relationship between early algebra and mathematical argument. However, this study represents only one particular approach. Although aspects of the methodology reported here may be useful to other researchers studying students' mathematical argument, we encourage researchers to integrate with novel techniques to further the field's understanding of how students engage in mathematical argumentation. The present study represents one such contribution.

Disclosure statement

No potential conflict of interest was reported by the authors.

Notes on contributors

Karl Wesley Kosko – Kent State University, USA.

Rashmi Singh – Kent State University, USA.

REFERENCES

- Bicer, A., Perihan, C., & Lee, Y. (2018). The impact of writing practices on students' mathematical attainment. *International Electronic Journal of Mathematics Education*, 13(3), 305-313. <https://doi.org/10.12973/iejme/3922>
- Blanton, M. L., & Kaput, J. J. (2011). Functional thinking as a route into algebra in the elementary grades. In J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 5-23). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-17735-4_2
- Bossé, M. J., & Faulconer, J. (2008). Learning and assessing mathematics through reading and writing. *School Science and Mathematics*, 108(1), 8-19. <https://doi.org/10.1111/j.1949-8594.2008.tb17935.x>
- Brousseau, G. (1997). Theory of didactical situations in mathematics: Didactique des mathématiques 1970-1990. (N. Balacheff, M. Cooper, R. Sutherland, & V. Warfield, Eds. and Trans.). Dordrecht, The Netherlands: Kluwer.
- Byrd, C. E., McNeil, N. M., Chesney, D. L., & Matthews, P. G. (2015). A specific misconception of the equal sign acts as a barrier to children's learning of early algebra. *Learning and Individual Differences*, 38, 61-67. <https://doi.org/10.1016/j.lindif.2015.01.001>
- Cirillo, M., Kosko, K. W., Newton, J., Staples, M., Weber, K., Bieda, K., ..., Strachota, R. (2016). *White paper: Conceptions and consequences of what we call argumentation, justification, and proof*. PME-NA Working Group.
- Cockcroft, W. H. (1982). *Mathematics counts*. London: Her Majesty's Stationary Office.
- Common Core State Standards Initiative [CCSSI]. (2010). *Common core state standards for mathematics*. Retrieved from Common Core Standards Initiative website http://corestandards.org/assets/CCSSI_Math%20Standards.pdf
- Creswell, J. W., & Plano Clark, V. L. (2011). *Designing and conducting mixed methods research* (2nd Ed.). Los Angeles: Sage Publications, Inc.
- Falkner, K. P., Levi, L., & Carpenter, T. P. (1999). Children's understanding of equality: A foundation for algebra. *Teaching Children Mathematics*, 6(4), 232.

- Fosnot, C. T., & Dolk, M. L. A. M. (2009). Young mathematicians at work: The role of contexts and models in the emergence of proof. In D. A. Stylianou, M. L. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 102-119). New York: Routledge.
- Fyfe, E. R., DeCaro, M. S., & Rittle-Johnson, B. (2014). An alternative time for telling: When conceptual instruction prior to problem solving improves mathematical knowledge. *British Journal of Educational Psychology, 84*(3), 502-519. <https://doi.org/10.1111/bjep.12035>
- Hackenberg, A. J. (2010). Students' reasoning with reversible multiplicative relationships. *Cognition and Instruction, 28*(4), 383-432. <https://doi.org/10.1080/07370008.2010.511565>
- Halliday, M. A. K., & Matthiessen, C. M. I. M. (2004). *An introduction to functional grammar* (3rd Ed.). London: Hodder Education.
- Knudson, R. E. (1992). The development of written argumentation: An analysis and comparison of argumentative writing at four grade levels. *Child Study Journal, 22*, 167-181.
- Kosko, K. W. (2016). Making use of what's given: Children's detailing in mathematical argumentative writing. *Journal of Mathematical Behavior, 40*, 68-86. <https://doi.org/10.1016/j.jmathb.2015.11.002>
- Kosko, K. W., & Norton, A. (2012). Relationships between the process standards: Process elicited through letter writing between preservice teachers and high school mathematics students. *School Science and Mathematics, 112*(6), 340-348. <https://doi.org/10.1111/j.1949-8594.2012.00151.x>
- Kosko, K. W., & Singh, R. (2016a). Children's linguistic and numeric unit coordination. In M. B. Wood, E. E. Turner, M. Civil, & J. A. Eli (Eds.), *Proceedings of the 38th annual meeting of the North American Chapter for the Psychology of Mathematics Education* (p. 605-612). Tucson, AZ: The University of Arizona.
- Kosko, K. W., & Singh, R. (2016b). Colligation and unit coordination in mathematical argumentative writing. In K. V. Adolphson, & T. M. Olson (Eds.), *Proceedings of the 43rd annual meeting of the Research Council of Mathematics Learning* (pp. 18-25), Orlando, FL: RCML.
- Kosko, K. W., & Singh, R. (2018). Elementary children's multiplicative reasoning: Initial validation of a written assessment. *The Mathematics Educator, 27*(1), 3-22.
- Kosko, K. W., & Zimmerman, B. (in press). Emergence of argument in children's mathematical writing. *Journal of Early Childhood Literacy*. <https://doi.org/10.1177/1468798417712065>
- Krummheuer, G. (2007). Argumentation and participation in the primary mathematics classroom: Two episodes and related theoretical abductions. *Journal of Mathematical Behavior, 26*, 60-82. <https://doi.org/10.1016/j.jmathb.2007.02.001>
- Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*, 159-174. <https://doi.org/10.2307/2529310>
- Liszka, J. J. (1996). A general introduction to the semeiotic of Charles Sanders Peirce. Indiana University Press: Bloomington, IN.
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of assumption violations revisited: A quantitative review of alternatives to the one-way analysis of variance F test. *Review of Educational Research, 66*(4), 579-619. <https://doi.org/10.3102/00346543066004579>
- Macedo, A. P. (2011). *The development of children's argument skills* (Doctoral Dissertation). London: Royal Holloway University of London.
- Matthews, P., Rittle-Johnson, B., McEldoon, K., & Taylor, R. (2012). Measure for measure: What combining diverse measures reveals about children's understanding of the equal sign as an indicator of mathematical equality. *Journal for Research in Mathematics Education, 43*(3), 316-350. <https://doi.org/10.5951/jresmetheduc.43.3.0316>
- Morgan, C. (2006). What does social semiotics have to offer mathematics education research? *Educational Studies in Mathematics, 61*(1-2), 219-245. <https://doi.org/10.1007/s10649-006-5477-x>
- Morris, A. K. (2009). Representations that enable children to engage in deductive argument. In D. A. Stylianou, M. L. Blanton, & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 87-101). New York: Routledge.
- Nordin, A-K., & Boistrup, L. B. (2018). A framework for identifying mathematical arguments as supported claims created in day-to-day classroom interactions. *Journal of Mathematical Behavior, 51*, 15-27. <https://doi.org/10.1016/j.jmathb.2018.06.005>

- Norton, A., Boyce, S., Ulrich, C., & Phillips, N. (2015). Students' units coordination activity: A cross-sectional analysis. *Journal of Mathematical Behavior*, 39, 51-66. <https://doi.org/10.1016/j.jmathb.2015.05.001>
- O'Halloran, C. L. (2014). *Supporting elementary English language learners' argumentative writing through a functional grammar approach* (doctoral dissertation). Ann Arbor, MI: University of Michigan.
- Peirce (1998). Nomenclature and divisions of triadic relations, as far as they are determined. In N. Houser, A. De Tienne, J. R. Eller, C. L. Clark, A. C. Lewis, & D. B. Davis (Eds.), *The essential Peirce: Selected philosophical writings* (Vol. 2; pp.289-299). Bloomington, IN: Indiana University Press (written originally in 1903).
- Russell, S. J., Schifter, D., & Bastable, V. (2011). Developing algebraic thinking in the context of arithmetic. In J. Cai & E. Knuth (Eds.), *Early algebraization: A global dialogue from multiple perspectives* (pp. 43-69). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-17735-4_4
- Sarkar, S. K., & Chang, C. K. (1997). The Simes method for multiple hypothesis testing with positively dependent test statistics. *Journal of the American Statistical Association*, 92(44), 1601-1608. <https://doi.org/10.1080/01621459.1997.10473682>
- Simes, R. J. (1986). An improved Bonferroni procedure for multiple tests of significance. *Biometrika*, 73(3), 751-754. <https://doi.org/10.1093/biomet/73.3.751>
- Singh, R., & Kosko, K. W. (2016). Elementary students' understanding of equivalence and multiplicative reasoning. In M. B. Wood, E. E. Turner, M. Civil, & J. A. Eli (Eds.), *Proceedings of the 38th annual meeting of the North American Chapter for the Psychology of Mathematics Education* (p. 212). Tucson, AZ: The University of Arizona.
- Steffe, L. P. (1994). Children's multiplying schemes. In G. Harel & J. Confrey (Eds.), *The development of multiplicative reasoning in the learning of mathematics* (pp. 3-39). Albany, NY: State University of New York.
- Steffe, L. P., & Cobb, P. (1986). *Construction of arithmetical meaning and strategies*. New York: Springer-Verlag.
- Stein, M. K., Engle, R. A., Smith, M. S., & Hughes, E. K. (2008). Orchestrating productive mathematical discussions: Five practices for helping teachers move beyond show and tell. *Mathematical Thinking and Learning*, 10, 313-340. <https://doi.org/10.1080/10986060802229675>
- Tall, D., Yevdokimov, O., Koichu, B., Whiteley, W., Kondratieva, M., & Cheng, Y-H. (2011). Cognitive development of proof. In G. Hanna & M. de Villiers (Eds.), *Proof and proving in mathematics education* (pp. 13-49). New York: Springer. https://doi.org/10.1007/978-94-007-2129-6_2
- Welch, B. L. (1951). On the comparison of several mean values: An alternative approach. *Biometrika*, 38(3/4), 330-336. <https://doi.org/10.2307/2332579>

<http://www.iejme.com>