Check for
updates

Assessing the Influence of Syntax, Semantics, and Pragmatics in Student Interpretation of Multiply Quantified Statements in Mathematics

Paul Christian Dawkins[1] ⬤ · Kyeong Hah Roh[2]

Abstract

This study compares the relative influence of syntax, semantics, and pragmatics in university students' interpretation of multiply quantified statements in mathematics, both before and after instruction. Like previous studies, results show that semantics plays a heavy role in student interpretation, especially before instruction. Unlike previous studies, our data suggests that the patterns of student interpretation rely more upon the mathematical context than upon the order of the quantifiers. We operationalize two of Grice's (1975) pragmatic Maxims to evaluate whether they help explain which interpretations are harder for students to adopt for various statements. Our data support the claim that students find it easier to construct relevant interpretations, but do not support the claim that students find it easier to construct interpretations that render the statement true. Finally, based on our sample from six Transition to Proof classes across the US, we observe that after their experiences in such courses students became more sensitive to syntax in their interpretation of the statements.

Keywords Logic · Quantifiers · Syntax · Semantics · Pragmatics · Transition to proof

Advanced mathematical language involves a number of very particular conventions of syntax and interpretation because mathematicians strive to communicate precise meanings with fidelity. Previous studies have particularly investigated how students make sense of statements that combine universal ($\forall$) and existential ($\exists$) quantifiers, which we

✉ Paul Christian Dawkins
   Pcd27@txstate.edu

   Kyeong Hah Roh
   khroh@asu.edu

[1] Department of Mathematics, Texas State University, 337 Mary Max Cir, San Marcos, TX 78666, USA

[2] School of Mathematical and Statistical Sciences, Arizona State University, Tempe, AZ, USA

shall call multiply quantified (MQ) statements. Such statements appear quite frequently in advanced mathematics and experts almost always use them in a consistent manner, though the precise nature of the relationships conveyed varies in important ways (c.f. Durand-Guerrier and Arsac 2005). Prior studies have assessed students' naïve readings of such statements (Dubinsky and Yiparaki 2000) and have proposed and evaluated certain methods of teaching students to interpret MQ statements as mathematicians do (Dubinsky et al. 1988; Dubinsky and Yiparaki 2000; Durand-Guerrier and Arsac 2005; Roh and Lee 2011). This study seeks to extend our insights into student interpretation of MQ statements by contributing a theoretical framework of the interpretation process. We evaluate this framework through carefully designed survey instruments administered to Transition to Proof students both before and after relevant instruction. We investigate the constituent influences syntax, semantics, and pragmatics each may play in the ways students construct meaning for MQ mathematical statements. In particular, we address the following research questions:

- To what extent do quantifier order, mathematical context, truth of the normative construal,[1] and relevance of the normative construal help explain variations in student interpretations of MQ statements in mathematics?
- How does student interpretation of MQ statements in mathematics change after experiencing Transition to Proof instruction?

Insights from Prior Literature

Interpreting MQ statements in mathematics resides at the interface between mathematical logic and mathematical language. We concur with previous researchers that while there exist formal rules for trying to render mathematical language purely syntactic (able to operate by precise rules ignorant of subject matter), mathematicians rarely operate in such a manner (Durand-Guerrier and Arsac 2005; Weber and Alcock 2005) and teaching novices will almost certainly require some balance between syntactic rules and semantic sense-making (Durand-Guerrier 2003; Durand-Guerrier et al. 2012).

   To portray this duality between syntax and semantics, consider the following paradigm examples of the kinds of "pairwise" relationships conveyed by MQ statements: identity and inverses. The identity relationship is a relationship between one object $e \in S$ and all others in a set $x \in S$. To express the idea that this one object interacts with all others in a particular way, mathematicians write "$\exists e \in S$ such that $\forall x \in S, \ e * x = x * e = x$." Placing the existential quantifier before the universal quantifier is understood to convey such a "one to every" relationship, as portrayed in Fig. 1. The inverse relationship is between pairs of objects such that each member of a set (x) has a corresponding object with which it is paired (x^{-1}). To express this relationship, mathematicians write "$\forall x \in S, \exists x^{-1} \in S$ such that $x * x^{-1} = x^{-1} * x = e$."

---

[1] We use the term "normative construal" to describe the way mathematicians interpret statements of a given form. We use the term "normative" to convey that the understandings we intend for students to develop are those generally agreed upon within the mathematical community.

Placing the universal quantifier before the existential is understood to convey such an "each to some" relationship as portrayed in Fig. 1.[2] Following previous studies we shall henceforth refer to the former type of statement as "EA" and the latter as "AE."

The Role of Syntax

Syntax describes the way that the grammar and structure of the statement itself influences the meanings it conveys. Students learning advanced mathematics are often taught particular rules for dealing with MQ statements in general. Durand-Guerrier and Arsac (2005) present some rules of "natural deduction" from Copi (1954) that could be used to work with these statements in precise, rule-based ways. More simply, those authors explain that the "dependence rule" states that quantities that appear later in a statement may depend upon those that appear earlier (as the choice of inverse depends on the choice of x). Roh and Lee (2011) also point out the "independence rule" that states that quantities appearing with earlier quantifiers should not depend upon those later in the statement (as the identity e does not vary with choice of x). These are ways syntax may dictate interpretation of the statement, and many previous studies recommend teaching such rules (e.g. Dubinsky and Yiparaki 2000; Epp 2003). The importance of sentential order in determining the meaning of MQ statements has been taught rather successfully using game theoretic ideas (Dubinsky and Yiparaki 2000; Glivická 2018) and using analogies that relate sentential order to temporal order (Dawkins and Roh 2016; Roh and Lee 2011).

The Role of Semantics

Semantics refers to the ways that a reader's understanding of the ideas referenced in a statement give meaning to that statement. Imagine a student read the definitions of identity and inverse for the first time without prior instruction on mathematical quantifiers and the assumed meaning of quantifier order. In such cases, students may make sense of the definitions by thinking about familiar instances of identities (0 or 1) and inverses ($-x$ and $1/x$). Considering these examples, students may infer that the former are "one to every" relationships and the latter are "each to some" relationships. In this way, students may use their knowledge of the relevant mathematical relationships to give meaning to the formal statements (c.f. Pinto and Tall 2002), as opposed to drawing this information from the syntactic form of the statement itself. Another example of this is the way that students have been observed to think epsilon depends upon delta in limit definitions in the same way that the function value depends upon the input value (Swinyard 2011). David, Roh, and Sellers (2019) found a similar effect in the ways students made meaning for the Intermediate Value Theorem. In these two cases, the semantic understanding leads students to interpret the definition in a non-

---

[2] We acknowledge that the example of inverses (and our diagram thereof) could be overgeneralized in the sense that "each to some" relationships need not include every element in the second in a pair and multiple elements of the first set may be paired to the same element. Also, the pairing may be fixed for each member of the first set (functional) or more flexible (Durand-Guerrier and Arsac 2005).
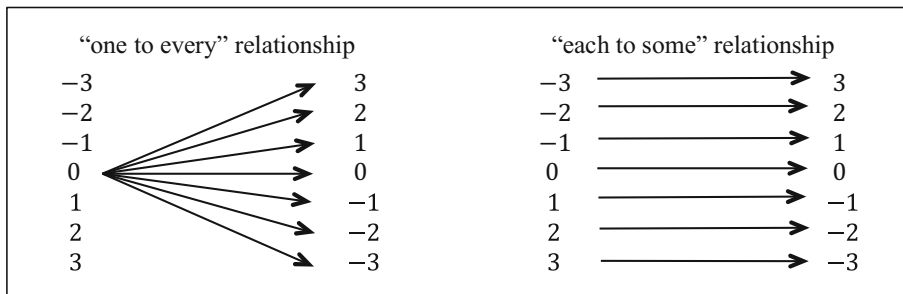
Fig. 1 Two types of pairwise relationships conveyed through MQ statements

normative way, but it demonstrates how quantification structure can be induced from semantic features. Durand-Guerrier and Arsac (2005) listed some particular qualities of a semantic context that they anticipate would exacerbate or alleviate problems with using the dependence rule, but they did not make predictions or test them regarding specific statements as is done in this study.

How then do students make meaning of unfamiliar, multiply quantified statements? There is not a single answer to this question and we do not expect the answer is the same in every semantic context. Nevertheless, prior studies provide some insight. Dubinsky and Yiparaki (2000) provide the most extensive study of students' untrained interpretations of a range of multiply quantified statements in various contexts. Some of their primary findings were that:

- students were unaware of their interpretation process,
- students interpreted everyday statements using their view of the world (semantics) often without attending to the syntax of the statement (the quantifiers at times seemed to be ignored),
- students discussed statements largely in terms of their understanding of the context/ situation that it referred to, rather than in terms of the statement itself, as a result they had trouble thinking of an alternative situation in which the truth-value of the statement would change,
- students found it easier to interpret AE statements and tended to interpret vague everyday statements as conveying "each to some" relationships, and
- students had more difficulty interpreting multiply quantified statements in mathematics.

These findings hold some important implications for how researchers analyze novice students' interpretations. If students read pre-consciously (i.e., are not aware of inter-pretation and do not analyze their own interpretive process) and do not attend directly to the role of the quantification phrases (or their order), then some of the "structure" that dictates the intended meaning of a mathematical statement for mathematicians remains inert in their reading (Dawkins and Cook 2017).

It is thus important to maintain that the conventions that mathematicians use to interpret MQ statements are neither necessarily correct (though they are useful) nor embedded directly in language itself. To see that these conventions are not necessarily correct, Epp (2003) provides examples of everyday statements that do not abide by mathematical conventions (matching EA wordings to "each to some" relationships). To

see that these conventions are not in language, one may consider a follow-up that Dubinsky and Yiparaki (2000) introduced to their initial interview study. The authors tried to teach students to see the influence of quantifiers order (EA vs. AE) by engaging them in a game. Each player selected a value of a variable in the role of a quantifier, and the order of quantifiers in the statement dictated the order of play (see also Glivická 2018). While some students recognized how changing the order of play changed whether a statement was necessarily true or false (i.e. which player had a winning strategy), at least one protested that these rules of the game were not "in the statement" (Dubinsky and Yiparaki 2000, p. 44). While some may object that the order of quantifiers is "in the statement," the convention that this order carries particular meaning is not "in the statement." Rather, any reader goes through an (unconscious or conscious) interpretation process to make meaning for a statement. Mathematicians have simply refined their interpretations so there is a tight match between their use of syntax and semantics.

The Role of Pragmatics

Finally, Dawkins and Cook (2017) found that, in addition to syntax and semantics, pragmatic issues sometimes influenced how undergraduate students interpret mathematical statements. In particular, some students decided that disjunctions such as "16 is even or 15 is odd" are false because the statement should be "16 is even and 15 is odd." In this case, Dawkins and Cook argue that students are using "false" to mark that the statement is an inappropriate speech act rather than to say that it is untrue. Those authors used Grice's (1975) Maxim of Quantity that states "Make your contribution as informative as is required" (p. 45) to explain these students' reasoning. Since a speaker ostensibly knows that both "16 is even" and "15 is odd" are true, then it is reasonable to expect them to use the connective *and* between them. The truth-functional use of *or* conflicts with the normative use of *or*, which is understood to convey alternatives or uncertainty (Dawkins 2019). The implication that someone making an *or* assertion is conveying some level of uncertainty or alternative is conventionally understood as part of the pragmatics of conversation, rather than syntax or semantics.

Theoretical Framework for Interpretation

In this section we shall outline ways in which we conceptualize the various elements of the interpretation process to facilitate our investigation into the ways syntax, semantics, and pragmatics contribute to students' meaning making for MQ statements. This involves presenting a theoretical framework for the interpretation process, building on the work of Stenning and van Lambalgen (2004, 2008).

Figure 1 above presents our basic analysis of the two types of relationships that MQ statements generally convey. Previous studies have used the language AE ("for every-there exists") and EA ("there exists-for every") to alternatively refer to 1) the structure of a mathematical statement, 2) the normative interpretation shared among mathematicians, and 3) a student's interpretation of those statements. While we continue to use those two-letter codes for convenience, we adopt a different terminology to distinguish

these constituents of the analytical process. Consider the definition of identity stated above: "$\exists e \in S$ such that $\forall x \in S, e * x = x * e = x$." The *wording of the statement* clearly exhibits EA structure. We refer to the meaning an individual makes for any such wording as their *construal* of the statement. The construal shared among mathematicians –EA means "one to every" – we call the *normative construal*. Each student will construe a statement in ways tantamount to "one to every," "each to some," or something else.

We relate the statement and student construal using Stenning and van Lambalgen's (2004, 2008) notions of *reasoning toward an interpretation* and *reasoning from an interpretation*, as portrayed in Fig. 2. For adults, reading is primarily a pre-conscious process by which the reader transforms the sequence of words into some mental representation (*reasoning toward an interpretation*), which we shall call a student construal. Obviously, students can engage in this process consciously, inasmuch as reading mathematical statements can be highly effortful and require rereading (e.g. Fletcher et al. 1999). However, we understand Dubinsky and Yiparaki's (2000) claim that many students in their study were unaware of their own interpretation process as saying that many novice mathematics students reason toward an interpretation pre-consciously. In our example of identity, most students very likely do not consciously consider "one to every" versus "each to some" pairings, but rather generalize from their knowledge of paradigm identities such as 0 or 1. It is in this process of reasoning toward an interpretation that we hypothesize that syntax, semantics, and pragmatics each have constituent influences upon the ways students construe a statement. Because this process is often pre-conscious, we must infer the role of these various aspects by comparing students' reading of closely related statements. Stated another way, since reading a statement often takes less than one second, much of the mental processing that goes into forming a mental representation is "under the hood" and is not consciously available to the reader. Once the reader has interpreted the statement, they then use their construal to draw inferences – *reasoning from an interpretation* – such as the whether the statement is true/false and why.
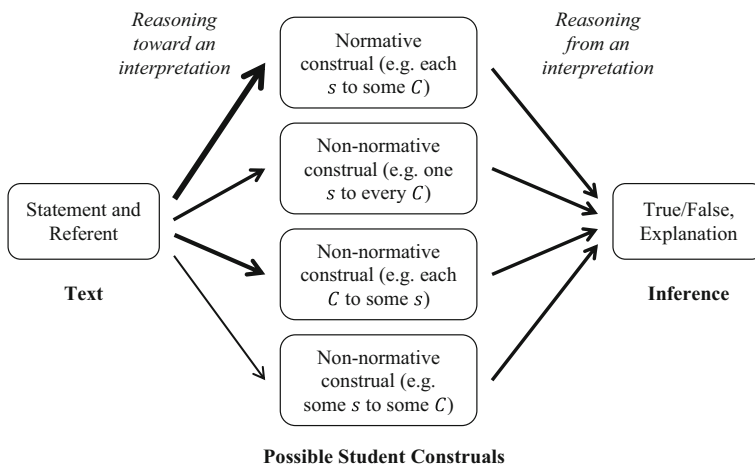

Possible Student Construals

Fig. 2 Our model of the interpretation process with example construals of S3

The model of interpretation in Fig. 2 is a researcher model in that it does not attempt to reflect what is consciously available to the reader. Students often do not distinguish the set of words in the statement from their construal of the statement's meaning, as Dubinsky and Yiparaki (2000) observed. The multiple construals in the diagram represent elements of the interpretation process that are available to us as experts and not necessarily to participants in the study. We can anticipate the most common student construals because they will approximate the normative construal of some permutation of the given statement (switching quantifier order and/or variable order). However, the final example of non-normative construal is not of this type, as some student construals will not be. We assume that no construal is inaccessible to adult students in advanced mathematics courses, but syntactic, semantic, and pragmatic factors make certain construals more easily accessible than others. We portray this in the diagram through the thickness of the arrows to various construals. We further anticipate that shifting construals once one is adopted can be rather difficult and effortful (e.g. Dawkins and Cook 2017), which is not portrayed in the diagram.

We parse the task-based elements that students may use to reason toward an interpretation in the following way: *quantifiers*, *predicate*, and *referent*. In the definition of identity, the quantifiers are "$\exists e \in S$ such that $\forall x \in S$," the predicate is "$e * x = x * e = x$," and the referent is a particular set S, operation $*$, and choice of e. We consider the influence of quantifiers in student construal as reflecting the role of syntax in interpretation. If students construct meanings based primarily in their understanding of the predicate and referent, we considered this most directly an influence of semantics. In both cases, we can evaluate the role of each element by systematically changing the quantifier order, the predicate, and the referent to observe shifts in student interpretation. Also, if students explain the definition of inverse with reference to paradigm examples, this constitutes evidence of the role of semantics in interpretation.

To assess the role of pragmatics, we operationalize two of Grice's (1975) pragmatic maxims. Grice's maxims express rules by which interlocutors in discourse may draw reasonable implications from another's statements that may reach beyond the strict meaning conveyed in the statement. The two we consider are a Maxim of Quality "Try to make your contribution one that is true" (p. 46) and a Maxim of Relation "Be relevant" (p. 46). If the Maxim of Quality influences student reasoning toward an interpretation, then we expect students would be more likely to construe statements so as to make them true. For normatively true statements, this would have no effect or aide in normative construal. For normatively false statements, we expect this would nudge students toward a non-normative construal under which the statement is true. From our standpoint as researchers, this means the maxims may constitute criterion by which a certain construal is more or less accessible to students as they read (easier or harder to construct).

Regarding the Maxim of Relation, we observe that certain construals convey meanings that are more semantically reasonable or interesting. If the Maxim of Relation influences student reasoning toward an interpretation, then we expect students will find it easier to construe statements in ways that make them convey semantically interesting information (i.e., statements that are worth saying). For statements that are normatively relevant, this would have no effect or aide in normative construal. For statements whose normative construals are uninteresting or absurd, we expect this would increase the frequency of non-normative construal. We assume that these maxims are operative in

the process of *reasoning toward an interpretation*, which we understand as primarily pre-conscious. Thus we are not invoking the Maxim of Relation with respect to students' subjective sense of interestingness, but rather we hypothesize that certain relations are more easily conceived of in the first place and alternatives will simply not come to mind without some conscious effort. Assessing the normative construals by this criterion involves a certain amount of expert judgment, which we shall address when we present our research tasks in the methods section.

Methodology

In this section we summarize the design of our survey instrument, data gathering methods, coding process, and analysis process.

Design of the Survey Instrument

In line with our framework for interpretation, we designed tasks that would vary the order of quantifiers, the context and predicate, and the referents. We desired to keep the set of tasks relatively short to maximize students' voluntary participation when they completed the assessment online outside of class time. Figure 3 presents the set of four statements, each with two referents, that comprise the eight research tasks. For each task, students were given the following prompt: "Determine whether the following statements are true or false and explain why. Try to explain in simpler terms what each statement says about the [given function/segment/ray]." For the geometry tasks, we added "We use the notation $d(A, C)$ to mean the distance between points A and C" and a diagram of a ray or segment.

These tasks varied the order of quantifiers in each statement to assess the role of syntax in interpretation. They varied the mathematical context and the referents to assess the role of semantics. We attempted to vary the relevance of the normative construal to assess the role of the Maxim of Relation. Statement S1's normative construal is mathematically interesting as it defines a function being bounded above. Statement S2's normative construal merely depends upon the fact that the set of real numbers is unbounded above, and is true of every real-valued function. We judge this at least a minor breach of the Maxim of Relation, since the choice of referent is irrelevant. This implies that S2 should be harder to construe normatively than S1. Statement S3's normative construal is mathematically interesting as it conveys the fact that rays extend

S1. "There exists a real number M such that for all real numbers x, $f(x) < M$." S2. "For all real numbers x, there exists a real number M such that $f(x) < M$."	referents	$f(x) = 3x + 2$ $f(x) = \sin(x)$
S3. "For every positive real number s, there exists a point C on the segment [ray] such that $d(A, C) = s$." S4. "There exists a point C on the segment [ray] such that for every positive real numbers s, $d(A, C) = s$."	referents	segment $\overline{AB}$ ray $\overrightarrow{AB}$

Fig. 3 The four statements and four referents comprising the study tasks

to all positive distances away from the endpoint (unlike segments). This is a form of the Real Ray Axioms in Blau (2008). The normative construal of S4 is in contrast rather absurd because it states that two particular points have infinitely many different distances between them. S4 violates the pattern that many relationships in geometry are functional (distance is a function of the two points compared), as was pointed out by Durand-Guerrier and Arsac (2005). We judge this a clearer violation of the Maxim of Relation and thus expect that S4 is harder to construe normatively than S3. In selecting contexts, we also made sure that in one context the relevant statement was of EA form (S1) and in the other it was of AE form (S3).

Finally, to assess the influence of the Maxim of Quality in student interpretation, we varied whether the first task in each context was normatively true or normatively false. By switching the order of presentation of the two referents in each context, we created two versions of the survey instrument: T-First version and F-First version (see Table 1). These two versions of the survey instrument contain the same tasks: four function tasks followed by four geometry tasks. In both contexts, the T-First version initially presents the referent that makes the first statement true and the F-First version initially presents the referent that makes the first statement false. Table 1 also presents the normative truth-value and normative construal for each statement/referent pair.

Administration of the Survey Instrument

Six instructors of Transition to Proof courses from five different universities in the United States allowed their students to participate to our research study in Spring 2018. Such Transition to Proof courses have become increasingly common in mathematics programs in the United States. David and Zazkis (2019) identified such courses in the mathematics programs of 179 out of the 215 US colleges and universities classified as having high or very high research output. Those authors found that 81% of such courses covered a standard set of topics: "symbolic/formal logic, truth tables, propositions, quantifiers, methods of proof (including contradiction and induction), number systems, set relations and functions, infinite sets, and cardinality" (p. 6). Cook et al. (2019) provide comparative analysis of the most commonly used texts for such courses. We did not gather more specific data about the content or instruction in the six courses we studied. Students in these courses are usually in their second or third year of university.

In order to gather the student responses from multiple universities, we created an online platform that all participants could access to complete the version of the survey assigned to them. We administered the survey instruments twice for the same participants, before and after their classes covered topics related to MQ statements. We refer the former to pretest and the latter to posttest, respectively. We only invited students to complete the posttest at least one week after they had completed instruction and summative assessments on MQ statements. Some professors agreed to offer a minor incentive to students who completed both surveys. Students could earn this incentive while opting out of the study to make sure study participation was voluntary. In total, 119 students completed the pretest and among them, 77 students completed the posttest and agreed to take part in the study. In this paper, we report our results from the 77 students who completed both pretest and posttest. We randomly assigned the student participants into the two groups and students remained in the same group for pretest and posttest. The T-First group had 43 respondents while the F-First had 34.

Table 1 Normative truth-values, construals, and task order for each task

Context - Predicate	Function - $f(x) < M$				Geometry - $d(A, C) = s$			
Referent	Sine function		Linear function		Ray		Segment	
Task name	EA Sine	AE Sine	EA Line	AE Line	EA Ray	AE Ray	AE Seg	EA Seg
Normative truth-value	T	T	F	T	T	F	F	F
Normative construal	One M to every x	Each x to some M	One M to every x	Each x to some M	Each s to some C	One C to every s	Each s to some C	One C to every s
Task order t-first group	1	2	3	4	5	6	7	8
Task order f-first group	3	4	1	2	7	8	5	6

Data Coding

Our codes for student responses describe the nature of the construal we infer that the student constructed for each statement. According to our model of interpretation (Fig. 2), this involves mapping back from the inferences students drew about each statement to the most likely construal that would lead to those inferences. The first source of evidence for coding was the explanation combined with the chosen truth-value. If there was vagueness or uncertainty about our initial code, we also compared to the students' responses to sister tasks (e.g. compare EA sine to EA line and AE sine) to gain more insight into their likely construal. We first randomly selected a small portion (20%) of the data that the two authors coded independently. Through follow-up discussion, we agreed to adopt three basic codes (EA, AE, Other) for our models of student construal. These corresponded respectively to whether students expressed a "one to every" construal, an "each to some" construal, or anything else. Responses fell into the third category if they entailed some other quantification structure, only quantified one variable, or if they construed the predicate in a manner incompatible with the normative construal (see Table 2). We use the phrase *rate of normative construal* to refer to the percentage of students who construed a statement/referent pair in a manner compatible with the way mathematicians do.

Table 2 Examples of student responses and codes assigned (one example of each code in each context)

Researcher-assigned code	Task	Student-assigned Truth-value	Student explanation
Function context			
AE-X	AE line	True	For any given x, I can provide an M that is larger than $f(x)$ by, for example, simply defining M as $f(x) + 1$.
EA-X	EA line	False	The function $f(x)$ increases infinitely. I, therefore, cannot provide a number M such that all outputs $f(x)$ fall below that M.
AE-M	AE line	True	There is always a larger real number
EA-M	EA sin	True	This is true for all numbers $M > 1$.
Other	EA line	False	Let M be 0, so if x is 1, then $M < f(x)$
Geometry context			
AE-X	AE ray	True	For every number, there is a point on the line where the distance from A to that point will equal the number.
EA-X	EA ray	False	distance of A and C can only be one number not every positive real number
AE-M	AE ray	True	There's a point C that would create such distance called S
EA-M	EA ray	False	If C lies 5 away from A, then d(A,C) =5; if we let s = 6, then clearly d(A,C) does not equal s.
Other	AE ray	True	when C is equal to s

As stated above, we adopted a fully online data gathering approach in part so we could gather more data from various sites across the United States. Naturally, conducting interviews would have improved our ability to infer the nature of student construals of the statements. However, according to our model of the interpretation process, this would not have improved our ability to assess how students reasoned toward an interpretation since we do not assume students are consciously aware of or have conscious control over that process (at least initially). Thus, interviews would not have improved our analysis of the rates of normative construals, they would only have improved our confidence in our inferences about their construals.

We did not code the dependence that students conveyed between the two variables because we could not reliably code all of the data in this manner. As was predicted in our framework for interpretation (specifically that syntax does not solely dominate interpretation), we observed that some students relied heavily on a semantic meaning (e.g. boundedness of the sine function) such that their explanation left some quantification implicit. For example, many students responded to the EA sine task with a value or range for M without explicitly noting that the relation held for all x. To capture this distinction among the responses coded either AE or EA, we further coded explanations as conveying explicit (X) quantification if they attended to the quantification of both variables or implicit (M) quantification otherwise. Once we agreed upon this two-part coding, each author coded another 40% of the data independently. Throughout the rest of the coding process, we selected complex responses to discuss as a research team, which allowed us to continually negotiate and refine the codes to maintain the reliability of the coding process.

Data Analysis

Our first research question considers the influence of "quantifier order, mathematical context, truth of the normative construal, and relevance of the normative construal" in student interpretation. To address these, we compared student construals and the rates of normative construal across tasks that varied only by quantifier order, context, and referent. Further we compared student responses between the T-First and F-First groups. Regarding our second research question about changes in student interpretation over the course of instruction, we analyzed the patterns of interpretation from pretest to posttest. We present each type of analysis we conducted in detail in the results section alongside the findings for each.

We anticipated that our data would demonstrate that, at least prior to instruction, semantics and pragmatics would strongly influence student interpretation. Dubinsky and Yiparaki (2000) claimed that students found AE statements easier to interpret, and were more likely to construe EA statements with a "each to some" construal. This suggests that the syntax was dominant in student interpretation. We rather hypothesized that semantic relevance would be more dominant in interpretation, so students would find it easier to construct "one to every" construals on the function tasks and "each to some" construals on the geometry tasks. Accordingly, we hypothesize that the rate of normative construal would be higher on the EA function tasks and the AE geometry tasks, according to the Maxim of Relation. We hypothesize that the T-First group would have a higher rate of normative construal on the first task in each context than would the F-First group, according to the Maxim of Quality. We did not register any of these hypotheses, so we consider our findings exploratory.

Results

We organize our results by the various types of analyses we conducted. First, we present findings regarding the rates of normative construal by task, time, and group. Second, we describe analyses comparing individual student construal across task, specifically whether quantifier order and referents influenced student construal. Third, we consider how often the students' explanations gave explicit attention to the quantification structure of the task.

Rates of Normative Construal

Figure 4 presents the rates of normative construal by time and task, juxtaposing the performance of the two groups on each chart. We remind the reader that in order to compare the two groups' performance by task, the order of the tasks on the charts does not match the order in which the F-first group responded to the tasks.

The first trend these data show is that students more frequently construed the first statement in each pair normatively. This appears visually as the jagged appearance of each graph. There are two possible explanations for this phenomenon. First, this pattern may confirm our hypothesis about the Maxim of Relation, namely that students were less likely to construct the normative construal when its meaning was either uninteresting (the EA function statement) or patently false (the AE geometry statement). We claim it is uninteresting to note that each function output is exceeded by some real number, since it makes the choice of function irrelevant (thus specifying a function might be viewed as violating a maxim). Further, we view it as absurd to claim that the distance between two particular points is every positive real number, inasmuch as points and distances on rays naturally invoke an "each to some" relation (Durand-Guerrier and Arsac 2005). Under this explanation of the data, the pattern demonstrates the role of both semantics and pragmatics in student interpretation. The alternative
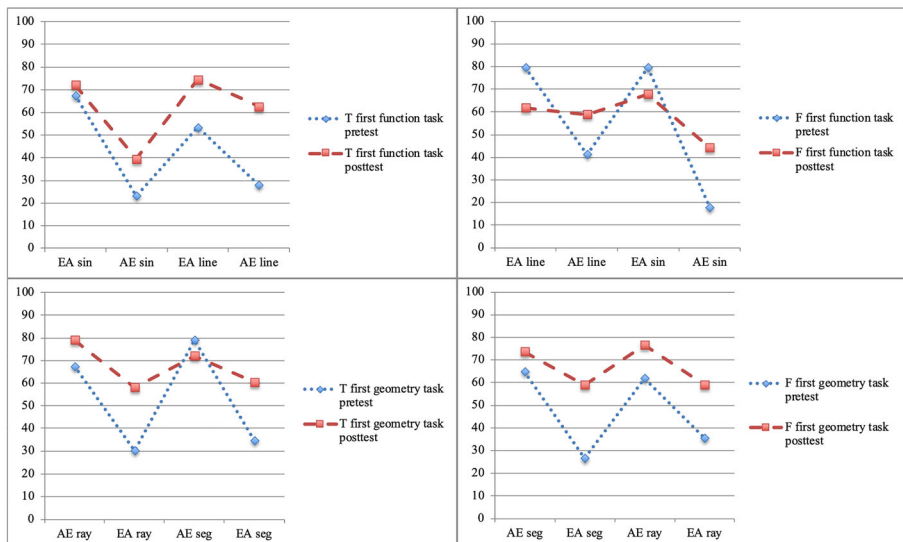
Fig. 4 Percentages of normative construal, organized by mathematical context and group

explanation for this data pattern relates to task order, because students always saw the more "natural" (according to normative construal) statement first. It may be that students construed the second statement less normatively because they had to develop a new construal for a very closely related statement, which was more challenging. The first explanation is content specific, while the second is content general. Our task design does not directly provide a way to distinguish between these alternative explanations, but we shall later provide evidence that the second explanation alone cannot account for the data.

The second pattern we notice in these data is that the AE sine task resulted in some of the lowest percentages of normative construal overall. The AE sine task's low normative construal rate should be viewed in part as a product of the analytical method. Since the sine function is bounded (the EA statement is true), then a single value of M satisfies the predicate for all real numbers x. Thus while the AE statement entails a slightly different construal (e.g. M could be .5 when $f(x) = 0$), the statement can be verified by selecting $M = 2$ for all x. Under either construal the statement is true, and students declared it so 88% of the time overall. When a student explains their interpretation of the AE sine task by noting that $M = 2$, this is insufficient evidence to indicate whether the student held a "one to every" or "each to some" construal. Without clear evidence that students understood how M could depend upon x, we did not code their responses as a normative "each to some" construal. Thus, it is likely that more students responded to the AE sine task according to a normative construal, but their explanation did not provide enough evidence for us to discern it. Many other explanations provided clearer evidence of either a "one to each" or an "every to one" construal.

A third pattern we observe in Fig. 3 is that on the posttest the rate of normative construal greatly increased for the more difficult statements (function EA and geometry AE), resulting in a more consistent rate of normative construal across group and context. Indeed, the rate of normative construal was between 58% and 80% on all of the posttest tasks except the AE sine task. This can be explained in two ways as above. Either students' improved their ability to construct less relevant construals, which was the greater challenge on the pretest, or they became more able to shift construals for similar statements.

To further illustrate the first and third data trends, Table 3 presents the difference between normative construal rates on each AE/EA task pair. On the pretest, these differences ranged from 25.6% to 61.8% with an average difference of 39.5%. On the posttest, these differences ranged from 2.9% to 32.6% with an average difference of 16.9%.

Table 3 Difference in normative construal between more and less relevant statements

Group	EA-AE sine pre	EA-AE line pre	EA-AE sine post	EA-AE line post
T-First ($N = 43$)	44.2% (19)	25.6% (11)	32.6% (14)	11.6% (5)
F-First ($N = 34$)	61.8% (21)	38.2% (13)	23.5% (8)	2.9% (1)
	AE-EA ray pre	AE-EA seg pre	AE-EA ray post	AE-EA seg post
T-First ($N = 43$)	37.2% (16)	44.2% (19)	20.9% (9)	11.6% (5)
F-First ($N = 34$)	26.5% (9)	38.2% (13)	17.6% (6)	14.7% (5)

Influence of Task Order

One of our primary hypotheses about student interpretation considered the influence of the Maxim of Quality, which assumes that speakers will say something true. If this maxim influenced students' interpretations of the given statements, then reading (normatively) false statements first would make students more likely to search for a construal that rendered the claim true. We would thus expect that the F-First group would have a lower rate of normative construal on such tasks. We only consider the first task in each context to account for the ways their initial reading influenced how students read subsequent, closely related statements. The rate of normative construal differed between groups by 10% or more on the following tasks: EA sine pre, EA line pre, AE line pre, AE segment pre, and EA sine post. Thus, the strongest evidence that the order of presentation affected student construal appeared on the function tasks prior to instruction. In this case, we see that the F-first group was actually more successful in finding a normative construal for both EA function tasks. This suggests that, in this context, reading the statement with reference to the linear function first aided students in construing the definition of bounded above with reference to both functions. This disconfirms our hypothesis regarding the Maxim of Quality.
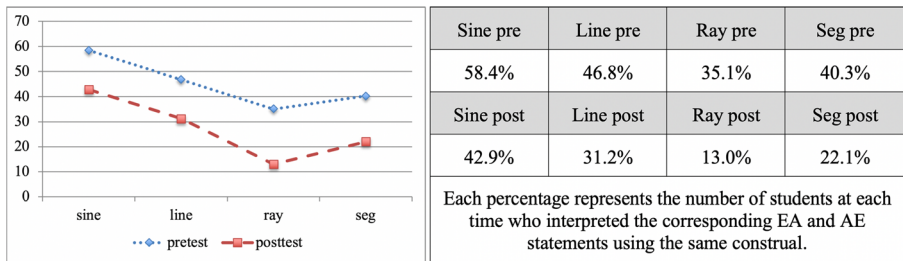
However, the geometry tasks caution against a general explanation that seeing a false statement first helps interpretation. The direction of the differences between the groups was exactly the opposite on the four geometry tasks than on the four function tasks. The group who saw the ray first (of which S3 is true) more frequently construed the geometry tasks normatively than did the group who saw the segment first (of which S3 is false). So, while there seemed to be some effect due to order of presentation, it varied with semantic content and not merely with the truth-value of the statement. This provides evidence for the role of semantics in interpretation, though in a manner inconsistent with our operationalization of the Maxim of Quality.

Considering the effect of order of presentation after students' experiences in Transition to Proof courses, the picture grows more complex. The two groups' rate of normative construal on the geometry tasks was nearly identical after instruction. On the function tasks, the T-First group outperformed the F-first group on all three tasks they previously underperformed on, and vise versa. This effect contributed to the fact that while the F-first group's overall rate of normative construal (maximum of 8) increased from 4.06 to 5.00, the T-first group showed greater gains by increasing from 3.84 to 5.19. Either we should infer that one task order benefits novices while the other benefits students with more experience, or the T-first group simply improved more from pretest to posttest. We do not have clear evidence to distinguish these two explanations. The F-first group showed lower gains at least in part because their rate of normative construal on the EA line task decreased from pretest to posttest by more than 17%. Their performance on the EA sine task also decreased by more than 12%.

Individual Student Analyses: Influence of Quantifier Order and Referent

We assessed the influence of syntax, focused on the quantifier part of the statement, by comparing each student's construal of statements that varied only in the order of quantifiers (EA function vs. AE function; AE geometry vs. EA geometry). Figure 5 presents the percentage of students who construed such pairs of statements in the same

	Sine pre	Line pre	Ray pre	Seg pre
	58.4%	46.8%	35.1%	40.3%
	Sine post	Line post	Ray post	Seg post
	42.9%	31.2%	13.0%	22.1%

Each percentage represents the number of students at each time who interpreted the corresponding EA and AE statements using the same construal.

Fig. 5 Frequency of students construing different quantifier order statements the same way

way. In other words, this is the percentage of students who interpreted corresponding EA and AE statements with the same construal (we ignored construals coded "other" in this analysis). The sine task showed the greatest frequency of invariant construal at both times, as might be expected due to the difficulty in coding responses to the EA sine task described above. Due to the methodological challenge posed by the sine tasks, we shall focus on the other three. On the pretest students construed the other three pairs of statements the same way between one third and one half of the time. On the posttest this rate dropped from between one eighth to one third of the time. Thus, it was quite frequent before instruction that reversing quantifier order did not elicit a novel construal and students became more sensitive to quantifier order by the time of the posttest.

Another pattern we notice in Fig. 4 is that students were better able to switch construals when the order of quantifiers changed on the geometry tasks than on the function tasks. On the posttest, about one third of students interpreted the function-related EA and AE statements the same, while only about one sixth of students interpreted the geometry-related EA and AE statements the same. This pattern suggests that students' ability to shift construals was context sensitive, reflecting the persistent role of semantics over pure syntax in interpretation. This finding helps distinguish between the two explanations given earlier for the first pattern in the rates of normative construal. If the only reason the second and fourth tasks in each context were harder was because it is difficult to shift interpretations for very similar statements, this could not explain why students were consistently better at doing so in geometry contexts. That means that some content-specific features of these statements render certain construals less accessible than others. We anticipated that this would be the case based on the Maxim of Relevance, but we anticipated that the AE function construals would be more easily accessible than the EA geometry construals. Thus, our evidence supports the claim that there are semantic factors that make some construals more and less easy to construct, but we cannot conjecture from our limited set of tasks what they are.

We also compared students' construals of the same statement regarding each pair of referents. One would hope that students would not completely shift their interpretation of a statement based on the object to which it currently refers, but Dubinsky and Yiparaki's (2000) reported that students interpreted statements largely in terms of their understanding of the state of affairs it describes. Figure 6 displays the percentage of students who shifted their interpretation from EA to AE or from AE to EA when only the referent changed. Notice, we ignored cases in which one or both of the construals was coded "other." As in other cases, the AE function rates were affected by the fact
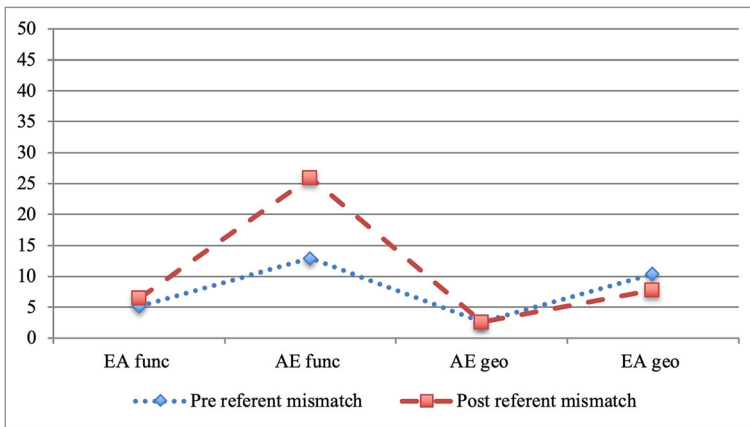
Fig. 6 Percentage of students who shifted interpretations when the referent changed

that many student responses to the AE sine task were too ambiguous to code reliably. Indeed, the AE sine and AE line tasks display a mathematical difference in the sense that a single value of M verifies the former and infinitely many values of M are necessary to verify the latter. There is thus a mathematical reason why student interpretations of those two tasks shifted based solely on the change of referent. On the other tasks, no more than 10.4% of the students shifted construals based on the referents (5.8% of students on average). This suggests that referents did not play a very prominent role in student interpretation of the statements, at least not to the point of completely changing the quantification structure.

Implicit and Explicit Quantification in Explanation

As noted in the methods section, some students' explanations did not explicitly reflect quantification of both variables. For instance, students sometimes responded to EA sine tasks by merely providing a value for M. Indeed, we conceptualized our study anticipating that some students would use a semantic conception such as the boundedness of sine to implicitly construct the quantification structure of the statement, rather than drawing on some content-general understanding of MQ statements. In such cases, we expected student responses to focus on the bound M without attending explicitly to the variation of x. To capture this aspect of student explanations, we coded each EA or AE construal as exhibiting implicit or explicit quantification.

Figure 7 presents the percentages of student responses that were coded with each type of quantification. We do not separate the groups for this analysis because we saw no clear reason why task order would influence the explicitness of quantification. The difference between each combined bar and 100% represents the number of construals coded "other." These data suggest that the percentage of codable responses (AE or EA) went up from pretest to posttest I every case and that the primary increase came in responses with explicit quantification structure (i.e. they attended to both variables in
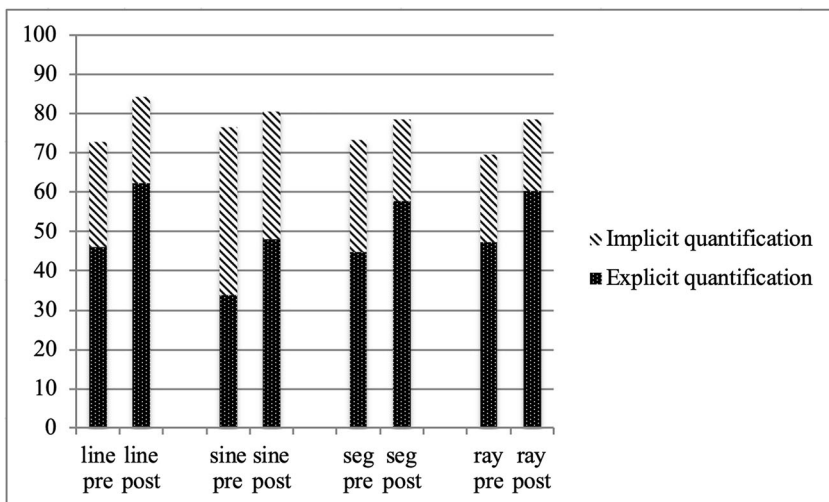
Fig. 7 Percentages of implicit or explicit quantification by time and task

some way). These results likely indicate that between the pretest and posttest students developed better tools for reasoning about the quantification structure of MQ statements and made them more articulate in explaining their reasoning. The sine tasks elicited the most implicit quantification, since so many students responded only in terms of M, which constituted a mathematically valid response.

Discussion

This study investigates the ways that students draw upon syntactic, semantic, and pragmatic features to give meaning to MQ statements in mathematics. Consistent with prior studies, we see that semantics played a prominent role in student interpretation, especially on the pretest. Our study adds to these prior studies in two particular ways. First, we constructed our instrument to help us quantitatively distinguish the influences of various aspects of interpretation, in addition to qualitative interpretation of student responses. Second, we sought to evaluate particular hypotheses about the role of pragmatics, specifically the influence Maxim of Quality and Maxim of Relevance. Our data did not support the role of the Maxim of Quality, since our study participants who saw a false statement first actually outperformed their peers in constructing the normative construal of the function items. Our data did support the role of the Maxim of Relevance to some extent. Students had less trouble constructing "one to every" construals for the function tasks and "each to some" construals for the geometry tasks. Because this matched our hypotheses around which the tasks were constructed, we infer that students had more trouble constructing the construals for statements that were less interesting (AE function) or relatively absurd (EA geometry). While students' rate of normative construal was higher on all such items after instruction, at both time periods students were better able to construct the (relatively absurd) normative con-strual for the EA geometry tasks than the (less interesting) normative construal for the AE function tasks. It is possible that the normative interpretations were less accessible

to students for reasons unrelated to our judgments of relevance, since this same pattern of normative construal could likely be explained using alternative criterion. Thus, the conservative interpretation is that some pragmatic criteria are at play to render certain semantic meanings more accessible than others, but we cannot from our data determine exactly what those criteria are.

Upon comparing their observations of student interpretations of MQ statements in everyday and mathematical contexts, Dubinsky and Yiparaki (2000) recommend that mathematics instruction "remain in the mathematical realm" (p. 55) to avoid the disanalogies between everyday contexts and mathematical ones. We accordingly conducted our study with exclusively mathematical statements. However, our findings emphasize the fact that some of the complexities of linguistic interpretation articulated by Grice's (1975) pragmatic maxims do not cease to influence students' reasoning toward an interpretation, even in mathematical contexts. We expect that further studies should explore more about the pragmatics of mathematical language, meaning the rules for what implications can reasonably be drawn beyond the explicit meaning of a given statement, both for experts and for novices. Further, we know little about the pragmatic aspects of interpretation that are specific to mathematics texts, which is a worthy arena for future investigation.

Our study also adds to prior research by documenting how experiences in Transition to Proof courses helped students improve their rates of normative construal, which we interpret as a shift toward syntax playing a larger role in student interpretation. On the pretest, students construed statements with reversed quantifier order in the same way between 35.1% and 58.4% of the time, consistent with Dubinsky and Yiparaki's (2000) finding that syntax remained relatively inert in many students' interpretive processes. However, experiences in their classes supported a 15.5–22.1% decrease in the rate at which students interpreted statements with reversed quantifiers in the same way. This suggests that our sample of six transition to proof courses from across the United States already exhibit some success in achieving Dubinsky and Yiparaki's (2000) recommendation to "help students learn to use the syntax of a statement as a tool for making sense of it" (p. 55).

More positively, relatively few students shifted their construal of a given statement based solely on a change in the referent. The exception to this appeared on the sine tasks in which the different referents make a meaningful difference in the standards for verification of the statement. This suggests that students' shifts in interpretation may reflect a meaningful difference in the relationship between statement and referent, though the referent by no means changes the normative construal.

It is worth comparing our findings regarding the role of semantic context in MQ statements to the conceptual analyses conducted by Durand-Guerrier and Arsac (2005). Those authors note that the dependence rule – objects quantified later in a statement may depend upon those quantified earlier in a statement – is so obvious in geometric contexts that quantification often remains implicit. For instance, the statement "all segments have a midpoint" (p. 157) hides the existential quantification on the midpoint itself, since the relationship of midpoint to segment is one of natural dependence (e.g. by construction). Those authors point out that this dependence is not so obvious in other settings, and thus the dependence rule is in more danger of being violated elsewhere. The geometric relationship between points and distance in our task is consistent with Durand-Guerrier and Arsac's observation, and we deemed the EA geometry statement relatively absurd because its normative construal violates that

dependence. Why then were students relatively successful in constructing that norma-tive construal and declaring it false? Durand-Guerrier and Arsac (2005) claim that the dependence rule is thus "nearly without interest in geometry" (p. 149) because it is semantically so obvious. This observation may shed some light on why this construal was more available to students than was the AE function normative construal, but we do not see a clear way to generalize the principle to anticipate on which other tasks less relevant construals will be more or less accessible to novice readers (Table 4).

Table 4 Summary of questions/ hypotheses, relevant comparisons, and primary findings

Question/ Hypothesis	Data compared	Findings
Influence of syntax	Rate at which students construed corresponding EA and AE pairs the same	- Students were less sensitive to syntax prior to instruction (35–47% of students construed reordered statements the same way, ignoring sine tasks). - Students were more sensitive to syntax after instruction (13–31% of students construed reordered statements the same way, ignoring sine).
Influence of semantics	Rates of normative construal across context and across referents	- The rates of normative construal were higher on EA function items than EA geometry items and on AE geometry items than on AE function items, especially on the pretest. - The rate at which students construed reorganized pairs the same way was lower on geometry items than on function items. - Students shifted interpretation due to reference relatively infrequently
Influence of pragmatics – Maxim of Quality	Rates of normative construal on the first task in each context by group	- This hypothesis was not supported. Rather, the primary difference between groups was that students in the F-first group performed better on the function items.
Influence of pragmatics – Maxim of Relation	Rates of normative construal on EA function and AE geometry tasks versus those on AE function and EA geometry tasks	- This hypothesis was supported, especially before instruction. Rates of normative construal were consistent with our predictions based on our judgment of which normative construal were more mathematically interesting or natural. - After instruction, the rates of normative construal were more consistent across the tasks.
Change in student interpretation over time	Pre and post normative construals	- The number of tasks construed normatively increased from pre to post by an average of 1.17 tasks out of 8 overall. - The primary increase in normative construal occurred on the "less natural" items AE function and EA geometry.
Attendance to quantification	Rates of explicit and implicit quantification	- The percentage of responses that could be coded and the percentage of explanations with explicit quantification were both higher on the posttest. Students became better at explaining their construals.

Limitations and Future Directions

Our study design reflects our desire to study changes in students' interpretation using a sample drawn from a number of classrooms. To conduct the study at multiple sites without taking class time and with minimal incentives for student participations, we had to keep the number of items small and accept the relative brevity of many student explanations. Clearly, an interview setting would allow for more fidelity in coding student construals, but it would afford neither the number of participants we recruited nor the range of locations from which data was gathered. Nevertheless, the ambiguity of many student responses poses a clear limitation to our findings. Future work could use the responses we gathered in this study to validate a multiple choice form of the research tasks in which students select a provided response that most closely approximates their thinking.

One of the primary challenges posed by ambiguity of student response was distinguishing "one to every" and "each to some" construals on the AE sine task. However, this difficulty is rooted in the logic of the task since the EA statement's truth implies the corresponding AE statements' truth, which makes their normative construals rather difficult to distinguish. It seems that this same difficulty would arise with any context in which the EA form of a statement is true.

We chose only to vary our task order in very controlled ways in order to test our hypothesis about the Maxim of Quality (that speakers make their contributions true). We could have conducted other interesting analyses had we randomized our task order. This likely would have required larger sample sizes, since there would be more implicit groups for comparison. Though our choice was motivated, our explanations of the data are limited by the fact that the later statements in any context are likely harder to interpret than the first, simply due to the challenge of shifting interpretations. Our study did not provide direct ways to control for that effect. Especially as we did not find evidence supporting the Maxim of Quality's role in interpretation, future studies could adopt a more thorough randomization of task order to better distinguish the effects of relevance and shifting construals. Naturally, adding other mathematical statements and contexts could also extend our findings.

Our predictions based on our operationalization of the Maxim of Relation were supported by the data. This lends some support to the claim that this maxim influences students' interpretations of MQ statements in mathematics, but alternative explanations could be formulated. We are less committed to the application of this particular maxim in this particular way than we are to the general principle that researchers can use pragmatic maxims to interpret (and even predict) student interpretations. This study provides a "proof of concept" for this general approach. We hope that future work will attend more to the ways that this lens can be extended and improved to better understand student reasoning rather than to the precise accuracy of our operationalization thereof.

Compliance with Ethical Standards

References

Blau, H. (2008). *Foundations of plane geometry*. Island Park: Whittier Publications.

Cook, J. P., Dawkins, P. C., & Zazkis, D. (2019). How do transition to proof textbooks relate logic, proof techniques, and sets? In A. Weinberg, D. Moore-Russo, H. Soto, & M. Wawro (Eds.), *Proceedings of the 22nd annual conference on research in undergraduate mathematics education* (pp. 146–153). Oklahoma City, Oklahoma.

Copi, I. (1954). *Symbolic logic*. New York: The Macmillan Company.

David, E. J., & Zazkis, D. (2019). Characterizing introduction to proof courses: A survey of US R1 and R2 course syllabi. *International Journal of Mathematical Education in Science and Technology*, 1–17.

David, E., Roh, K.H., & Sellars, M. (2019). Teaching the representations of concepts in Calculus: The case of the intermediate value theorem. *PRIMUS*. https://doi.org/10.1080/10511970.2018.1540023.

Dawkins, P. C. (2019). Students' pronominal sense of reference in mathematics. *For the Learning of Mathematics, 39*(1), 18–23.

Dawkins, P. C., & Cook, J. P. (2017). Guiding reinvention of conventional tools of mathematical logic: Students' reasoning about mathematical disjunctions. *Educational Studies in Mathematics, 94*(3), 241–256.

Dawkins, P., & Roh, K. (2016). Promoting meta-linguistic and meta-mathematical reasoning in proof-oriented mathematics courses: A method and a framework. *International Journal of Research in Undergraduate Mathematics Education, 2*, 197–222.

Dubinsky, E., & Yiparaki, O. (2000). On student understanding of AE and EA quantification. *Research in Collegiate Mathematics Education, IV*, 239–289.

Dubinsky, E., Elterman, F., & Gong, C. (1988). The student's construction of quantification. *For the Learning of Mathematics, 8*(2), 44–51.

Durand-Guerrier, V. (2003). Which notion of implication is the right one? From logical considerations to a didactic perspective. *Educational Studies in Mathematics, 53*(1), 5–34.

Durand-Guerrier, V., & Arsac, G. (2005). An epistemological and didactic study of a specific calculus reasoning rule. *Educational Studies in Mathematics, 60*(2), 149–172.

Durand-Guerrier, V., Boero, P., Douek, N., Epp, S. S., & Tanguay, D. (2012). Examining the role of logic in teaching proof. In G. Hanna & M. De Villiers (Eds.), *Proof and proving in mathematics education* (pp. 369–389). Netherlands: Springer.

Epp, S. (2003). The role of logic in teaching proof. *The American Mathematical Monthly, 110*, 886–899.

Fletcher, C. R., Lucas, S., & Baron, C. M. (1999) Comprehension of mathematical proofs). In S. R. Goldman, A. C. Graesser, & P. van den Broek (Eds.), *Narrative comprehension, causality, and coherence: Essays in honor of tom Trobasso* (pp. 195–207). Mahwah: Lawrence Erlbaum Associates.

Glivická, J. (2018). Game semantics in teaching of classical first order predicate logic to students of computer science. In *MATEC web of conferences* (Vol. 210, p. 4007). EDP Sciences.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics: Speech acts* (Vol. 3, pp. 41–58). London: Academic.

Pinto, M., & Tall, D. (2002). Building formal mathematics on visual imagery: A case study and a theory. *For the Learning of Mathematics, 22*(1), 2–10.

Roh, K., & Lee, Y. (2011). The Mayan activity: a way of teaching multiple quantifications in logical contexts. *PRIMUS, 21*, 1–14.

Stenning, K., & van Lambalgen, M. (2004). A little logic goes a long way: Basing experiment on semantic theory in the cognitive science of conditional reasoning. *Cognitive Science, 28*, 481–529.

Stenning, K., & van Lambalgen, M. (2008). *Human reasoning and cognitive science*. Cambridge: The MIT Press.

Swinyard, C. (2011). Reinventing the formal definition of limit: The case of Amy and Mike. *The Journal of Mathematical Behavior, 30*(2), 93–114.

Weber, K., & Alcock, L. (2005). Using warranted implications to understand and validate proofs. *For the Learning of Mathematics, 25*(1), 34–38.