Journal of Mathematical Behavior 80 (2025) 101268

Contents lists available at ScienceDirect

Journal of Mathematical Behavior

journal homepage: www.elsevier.com/locate/jmathb

Using test cases to refute incorrect existentially quantified propositions: An exploratory study

John Griffith Tupouniua[a,*], John Smith[b]

[a] Institute of Education, Massey University, Auckland, New Zealand
[b] Russell Street School, Palmerston North, New Zealand

ARTICLE INFO

Keywords:
Existentially quantified propositions
Test cases
Counterexamples
Domain of validity

ABSTRACT

Towards the goal of extending the applicability of test cases to the context of existentially quantified propositions, the present study explores how test cases might support learners with refuting their incorrect existentially quantified propositions. We present and analyze data from two separate instances in which two in-service primary school teachers initially made incorrect existentially quantified propositions and then were asked to find a valid example of their respective propositions (i.e., an element of the subject that satisfies the predicate). The participants were given, and sometimes generated their own, test cases which led to an iterative process of ruling out potential examples and classes of potential examples. Our analysis of this iterative process as it emerged within our specific research setting, comprising among aspects, particular researcher-participant interactions, sheds light on how these test cases afford and support the development and refinement of the learners' respective existentially quantified propositions.

1. Research background and rationale

Counterexamples are particularly useful for helping learners refute incorrect propositions. In theory, counterexamples prompt learners to confront their current incorrect propositions, reject them, and construct universally correct propositions (Balacheff, 1991; Klymchuk, 2010; Peled and Zaslavsky, 1997). For instance, the proposition 'multiplying any two real numbers a and b always yields a larger number c', can be countered with $a = 1/2$ and $b = 1/3$. In presenting the learner with this counterexample, one hopes that the learner rejects their proposition and realizes that multiplying any two real numbers does not always yield a larger number.

The counterexamples approach to supporting students refute and revise their incorrect mathematical propositions can be thought of as an approach whereby particular cases (examples) are used to test the validity of the proposition. These test cases may or may not lead to the refutation and revision of the proposition. The present study is concerned with the use of these test cases in supporting students to refute incorrect existentially quantified propositions. To clarify and distinguish the present study's context of interest from those that have been explored in past research, consider the following propositions:

(1) $a^2 + 2ab + b^2 < 0$ *for all real numbers a, b*
(2) $a^2 + 2ab + b^2 < 0$ *for any real numbers a, b such that a and b are greater than zero*
(3) $a^2 + 2ab + b^2 < 0$ *for some real numbers a, b*

* Corresponding author.
E-mail addresses: j.g.tupouniua@massey.ac.nz (J.G. Tupouniua), john@russellst.school.nz (J. Smith).

https://doi.org/10.1016/j.jmathb.2025.101268
Received 17 July 2024; Received in revised form 29 May 2025; Accepted 30 May 2025
Available online 12 June 2025

The truth value of all three propositions is false. However, in the first two propositions, the domains within which the propositions are said to be true have been specified. In this study, we use the term *domain of validity (DoV)* to refer to this domain. In other words, the first two propositions are *universally quantified* (i.e., the statement is said to be true for all elements within a certain set). As such, in either of the first two propositions, one can ascertain a concrete counterexample—an element of the DoV for which the proposition is false (e.g., $a = 2$ and $b = 3$). While the respective DoVs are clear for the first two propositions, this is not so for the third proposition (i. e., “some real numbers”), which is *existentially quantified*. As such, for the third proposition, one cannot ascertain a single concrete counterexample because the DoV has not been specified. So, the third proposition represents a scenario that complicates the use of test cases, because a teacher for example cannot ascertain any counterexamples for the learner’s proposition.

So how might a teacher use test cases to support learners in refuting their incorrect propositions in situations where the DoV is unspecified? As we will summarize in the Literature Review, the situation represented by the first two propositions above has been explored throughout the literature, that is, in terms of using specific test cases (counterexamples) to support learners in refuting incorrect propositions. In contrast, much less is known about the applicability of test cases in situations represented by the third proposition. This gap in the literature was noted by Buchbinder and Zaslavsky (2019) who argued that the majority of past research that has explored how students use examples “have restricted their scope to universal (“for all”) statements, without attending to students’ understanding of the roles of examples in existential (“there exist”) statements” (p. 130). Buchbinder and Zaslavsky also referenced other studies (e.g., Barkai et al., 2008; Tabach et al., 2012) that found that both primary and secondary mathematics teachers were often not aware of existential quantifiers, and were unaware of how an example could prove or disprove an existentially quantified proposition. This aligns with our observation that secondary school students in our context of New Zealand have very little exposure to existential statements in the mathematics curriculum, and thus enter first-year university mathematics courses with little to no understanding of the fundamental position that existential quantifiers occupy in mathematical reasoning, and how examples could be used to reason about existentially quantified statements. Thus, we argue that there is a need for research to extend the applicability of the test cases (examples) approach to the context of existentially quantified propositions. More specifically, how do test cases (examples) impact students’ reasoning with respect to existentially quantified propositions?

As such, in this paper, we present and analyze data from two separate instances in which two learners (in-service primary school teachers) initially made incorrect existentially quantified propositions and then were asked to find a valid example of their respective propositions. The participants were given, and sometimes generated their own, test cases which led to an iterative process of ruling out potential examples and classes of potential examples. Our analysis of this iterative process as it emerged within our specific research setting, comprising among other aspects, particular researcher-participant interactions, sheds some light on how the learners respond to the test cases and how these test cases contribute to the development and refinement of the learners’ respective existentially quantified propositions.

2. Literature review

This section is in two parts. We begin by clarifying the key concepts used in the present study. Then, we summarize key findings from literature within which we situate our study, in addition to clarifying and distinguishing the focus of this study from those of previous related studies.

2.1. Key concepts

A *proposition* is a statement consisting of two related terms—a *subject* and a *predicate*—and is either true or false. For example, “even numbers are greater than or equal to three” is a proposition that connects the subject “even numbers” to the predicate “greater than or equal to three”. A proposition can include *quantifiers* (e.g., “one”, “all”, “some”) which indicate for how many elements of the subject domain, the predicate holds. There are two types of quantifiers. A *universal quantifier* indicates that the predicate holds for every element of the subject. An *existential quantifier* indicates that the predicate holds for at least one element of the subject.

We use the term *domain of validity* (of a proposition) to refer to the set of all elements of the subject (i.e., a subset of the subject) for which the predicate is proposed to be true. The DoV of a proposition is sometimes specified (i.e., the subject is universally quantified), that is, we can construct a concrete element of the DoV. Here, the DoV is equivalent to the subject. At other times, the DoV is unspecified (i.e., the subject is existentially quantified), that is, we cannot construct a concrete element of the DoV. In Table 1, the first two propositions have domains of validity which have been specified (i.e., they are universally quantified). However, for the last two examples, the domains of validity are unspecified (they are existentially quantified). Regardless of whether the DoV is specified or unspecified, the predicate holds for all elements of the DoV by definition. If a proposition is universally quantified, then the DoV is equivalent to the subject. If a statement is existentially quantified, it does not rule out the possibility that the predicate holds for all

Table 1
Examples of propositions.

Proposition	Domain of validity	Predicate
1. All cars have four wheels	All cars	Four wheels
2. All cars in New Zealand have four wheels	All cars in New Zealand	Four wheels
3. $f(a, b) = a^2 + b^2 + 2ab < 0$ for some real values of a and b	Some real values of a and b	$f(a, b) < 0$
4. The product of at least one pair of prime numbers is also a prime number	At least one pair (a, b) of prime numbers	a x b is a prime number

elements of the subject.

Some propositions can include multiple quantifiers (e.g., "for every prime number x, there exists a real number y such that $x = 2y$") but this is beyond the scope of the present study. As a small exploratory study, we focus this study on propositions that have a single existential quantifier. In light of the definitions above, we can say that a *counterexample to a proposition* is an element of the proposition's domain of validity for which the predicate does not hold. For instance, consider the following proposition: "Multiplying two numbers a and b always yields a number that is at least as large as both a and b, where a and b are both greater than 1." The pair of numbers $a = 1/2$ and $b = 5/6$ is not a counterexample to the above proposition since it lies outside the DoV (i.e., a and b are both greater than 1). However, $a = 1/2$ and $b = 5/6$ would be a counterexample if the proposition is: "Multiplying two numbers a and b always yields a number that is at least as large as both a and b, where a and b are real numbers." A counterexample can be ascertained only when the subject is universally quantified (i.e., the DoV is specified).

As this study is focused on existentially quantified propositions, we prefer to use the term "test cases" rather than "counterexamples". Test cases comprise the four classes of *examples* introduced by Buchbinder and Zaslavsky (2019). They claimed that with respect to every proposition made up of a subject S and a predicate P, we can distinguish between four classes of examples according to whether an example (x) is an element of the subject or not ($x \in S$ or $x \notin S$) and whether it satisfies the predicate or not (P(x) or ¬P(x)). In the context of existentially quantified propositions, the four types of test cases are:

- test case is an element of the subject and satisfies the predicate
- test case is an element of the subject but does not satisfy the predicate
- test case is not an element of the subject and does not satisfy the predicate
- test case is not an element of the subject but it satisfies the predicate

Buchbinder and Zaslavsky (2019) refer to the first class above as *confirming examples,* because they confirm the validity of the existentially quantified proposition; the second class as *non-confirming examples* because they support but are insufficient for proving or refuting the validity of the proposition. They refer to the last two types of classes, as *irrelevant examples* because they are inconsequential to the validity of the proposition. We prefer the term *test case* instead of *example* because the latter is ambiguous in multiple ways. Saying that something is an *example*, carries an assumption that it is an element of some set (namely the domain or the subject of the proposition). However, we see that in Buchbinder and Zaslavsky's definitions, two classes of examples (irrelevant examples) include non-elements of the subject, in which case one is tempted to ask: what are they examples of? As such, the notion of a test case includes both elements and non-elements of the subject and applies to both existentially and universally quantified propositions (though the focus of the present study is the former). A test case is thus some sort of device that is used to test the validity of a proposition. It is not necessarily so that the test case would reveal a flaw in the proposition. In fact, the revelation of a flaw cannot to some extent be determined a priori; because even if in theory a test case should reveal a flaw, when learners conduct the test, it may not reveal a flaw to them. And even if the test case ought not to reveal a flaw, the test case might still reveal a flaw from the learner's perspective.

2.2. Summary of findings from relevant literature

Many past studies have found that when learners are confronted with examples that contradict their propositions, learners do not always conduct revisions that result in correct propositions (e.g., De Bock et al., 2002; Moala, Yoon & Kontorovich, 2019; Peled & Zaslavsky, 1997; Tupouniua, 2022b; Zazkis & Chernoff, 2008; Zazkis et al., 2008). One explanation for this offered by Zazkis and Chernoff (2008) is that while all counterexamples have the potential to reveal incorrect propositions, not all counterexamples invoke cognitive conflict; moreover, counterexamples that motivate cognitive conflict do not necessarily facilitate conflict resolution. In other words, while some counterexamples may lead learners to acknowledge their incorrect propositions, the ensuing actions of learners do not always lead to correct propositions. Zazkis and Chernoff (2008) go on to suggest that some counterexamples are more effective than others in facilitating correct propositions. Larsen and Zandieh (2008) suggested that different learners may respond to the counterexamples differently. Larsen and Zandieh's (2008) analysis focuses more on differentiating between how the students respond to the counterexamples, as opposed to distinguishing between the counterexamples themselves. They drew on Lakatos' (1976) three notions of *exception barring*, *monster barring*, and *proofs and refutations* to explain these differences in students' responses. Stylianides's (2008) framework explains how the creation of mathematical knowledge involves activities such as identifying patterns, making conjectures, and providing proofs. He uses the collective term *reasoning-and-proving* to describe these activities. Stylianides' framework distinguishes between two different kinds of proof: generic examples and demonstrations. A generic example is a proof in which a particular case represents a general case. A demonstration is a proof that does not rely on a particular representative case, but rather argumentation processes such as induction, contraposition, contradiction, and exhaustion. Dogan and Williams-Pierce (2021) further argue that using a specific case as a generic example requires that the learner focuses on characteristics that are applicable to a broader group (Mason & Pimm, 1984): "If a prover is using the number 4 as an example of the domain of positive integers, yet relies upon the divisibility by 2 as a component of their proof, they are in fact using 4 as a non-generic example. However, if the number 4 is an example of the domain of all even positive integers, divisibility by 2 indicates that the number 4 may, in fact, be a generic example" (p. 134).

As previously mentioned, in our survey of relevant literature pertaining to the use of test cases to support the development of learners' mathematical thinking, an overwhelming majority studies were concerned with universally quantified propositions. We present a few examples from the literature below:

- In Larsen and Zandieh (2008), the participants conjectured that any subset of a group that is closed under the group operation is a subgroup. Here the domain of validity is specified as *any subset of the group that is closed under the group operation.* The counterexample offered was: *real numbers under addition*, with the subset being *positive numbers.*
- In Komatsu (2016) two learners conjectured that for the sum $ab + ba$ (where ab is any two-digit integer, and a and b are both integers between 0 and 9), the ones digit of the sum is always the same as the tens. The DoV is specified here as ab is any two-digit integer, and a and b are both integers between 0 and 9. The learners were given $85 + 58$ as a counterexample.
- In Zazkis and Chernoff (2008), a learner conjectured that any two prime numbers multiplied with each other is a prime number. The specific DoV here is *the set of any two prime numbers*. The learner was then given 3 and 5 as a counterexample.
- In Tupouniua (2022a) two learners conjectured that when trying to find the larger of two simple fractions, for each fraction, multiply the numerator and the denominator; the fraction that has the bigger product is the larger fraction. They proposed that this conjecture was true for all fractions that were not equal (i.e., they had ruled out fractions such as 1/2 vs 4/8). The learners were given 19/23 vs 23/29 as a counterexample.

In each of the instances above, an external source (e.g., teacher, researcher) is able to identify a test case that is by definition a counterexample to the learner's proposition because it exists within the specified DoV. However, this task of determining counterexamples becomes impossible the proposition is existentially quantified. As such, the question that we ask in this study is: how might test cases impact students' thinking and the development of their propositions when the proposition is existentially quantified?

We found in our survey a few studies that have touched on existentially quantified propositions, though their aims and focus differ from ours. Some studies have focused on propositions with multiple (both existential and universal) quantifiers. For instance, Sellers et al. (2021) explored how students generally interpreted (without interventions) quantified propositions and the variables within these statements. Vroom (2022) examined students' understanding of multiple quantified propositions and stressed the importance of developing students' fluency with formal mathematical language. Dawkins and Roh (2020) studied the roles of syntax, semantics, and pragmatics in university students' understanding of multiply quantified propositions. Piatek-Jimenez (2010) explored undergraduate mathematics students' interpretations of and argumentation related to mathematical propositions involving multiple quantifiers, comparing interpretations of statements of the form "There exists … for all …" with interpretations of "statements of the form "For all … there exists …". Ye's (2012) study included some questions that were existentially quantified with a single quantifier (e.g., "some athletes won gold medals"). But their focus was on how students transform logically quantified statements into equivalent statements (e.g., "some athletes won gold medals" to "it is not the case that all athletes won gold medals").

Despite the lack of literature focusing specifically on our phenomenon of interest – how students respond to test cases when they have incorrect mathematical propositions with a single existential quantifier, we were able to find and draw on a couple of studies and frameworks that we found particularly useful in theorizing about our study. Firstly, Buchbinder and Zaslavsky (2019) as discussed in the previous subsection introduced the Roles of Examples in Proving (REP) framework to explore students' understanding of examples in relation to proving mathematical statements. From this framework, we draw on the four types of test cases (see previous subsection) that are relevant to existentially quantified propositions. One of Buchbinder and Zaslavsky's main findings was that their participants' reasoning processes were weakest in relation to *non-confirming examples* (i.e., examples which were elements of the subject but did not satisfy the predicate). Secondly, we draw on Ellis et al.'s (2019) CAPS (criteria, affordance, purpose and strategies) framework which explains the different ways that learners use examples in proving-related activities. The framework also provides a way of examining how the learners' usage of examples evolves over time. The specific aspect of the CAPS framework that is most relevant to the present study is *affordances* which refers to "the gains or outcomes students experience from their example use, and address how working with examples can support students' understanding, conjecturing, and proving activities" (p. 6). The other three aspects are less relevant because they rely on understanding students' intentions and their reasons for choosing particular examples – which is beyond the scope of the present study.

3. Methods

The data on which this paper is based were collected as part of a study focused on generating plausible hypotheses about the influence of counterexamples on learners' mathematical propositions (Tupouniua, 2022a; Tupouniua, 2022b). It was motivated by an observation made while collecting data on another study which explored the role of counterexamples in the development of student-invented algorithms (see Tupouniua, 2022a). One of the participants in Tupouniua (2022a) had created an algorithm for expanding binomials and claimed that the expression $a^2 + b^2 + 2ab$ "will be negative for some numbers". When asked, "where do you think it will be negative?" the participant replied, "I don't know but I think it has to be negative in [sic] some numbers".

In the following subsections below, we describe the participants, data collection procedures, instruments, and data analysis methods for the present paper.

3.1. Participants and tasks

The participants whose work are analyzed in this paper were two in-service primary school teachers, who at the time of data collection were enrolled in a postgraduate diploma program[2] in mathematics education at a large public university in New Zealand. Fascinated by the observation made during Tupouniua (2022a) we wanted to conduct a small study to observe how learners respond to test cases when they claim that a normatively false existentially quantified proposition such as "$a^2 + b^2 + 2ab$ will be negative for some numbers" is true. Our intention was to conduct a very small exploratory, preliminary study that would lay the foundation for a larger study. One of the things we wanted to test in this preliminary study was whether the task, comprising multiple existentially quantified propositions (see below) would elicit instances in which the learners claim that a false existentially quantified proposition was true. The propositions in the task were all false existentially quantified propositions. They were also propositions that we, based on our teaching experiences, thought would not be too difficult given the mathematical background of the participants, but may still elicit some uncertainty from the participants regarding their truth values. For example, in the first author's experience teaching mathematics education courses with both pre-service and in-service primary school teachers, he has come across many students who were convinced that that $a^2 + 2ab + b^2 < 0$ must be true for some values of a and b. Moreover, given that the goal of our study was to explore how the students respond to test cases in light of their incorrect existentially quantified propositions, it was decided that in instances where the students claimed that a proposition was true, we would immediately give them test cases and then observe the impact that the test cases had on their thinking.

The two participants, Max and Leo (pseudonyms) worked alone on the task (see Fig. 1) at separate times with the researcher. The task consisted of five propositions to consider and to determine whether they were true or false (see Fig. 1 below).

All propositions in the task are false. That is, for all the propositions one cannot construct an element of the subject that will satisfy the predicate. The participants were given all five propositions one by one and were asked to determine whether they were true or false and to provide justifications where possible. The two participants, Max and Leo, were given 30 minutes to work on all the propositions, but they were not required to complete all five. The participants were also told that they could work on the propositions in any order they liked; however, both of them worked on the task in the order presented in Fig. 1. Both participants made it to the fourth proposition, but only Leo made it to the fifth (on which he spent about three minutes before the session ended). Both participants answered "false" for the first two propositions and were able to provide valid justifications. One participant (Leo) answered false for the third proposition and justified his answer. The other participant (Max) answered true for the third proposition, but he could not immediately provide a concrete example to justify his claim. Max answered true for the fourth proposition and began to work on it, but did not get far before the 30 minutes was complete. Leo answered true for the fourth proposition but was not able to initially provide a concrete example for the DoV.

3.2. Researcher-participant interactions

In this paper, we draw on the works of Steinbring, 1997; Steinbring, 2007 on teacher-learner interaction to frame our understanding of the data, which comprise interactions between a researcher and a research participant. Steinbring (2007) states that "education in school can be understood as an introduction of [learners] being guided [educators] into the culture of socially conveyed knowledge" (p. 95). Perhaps the most significant aspect of Steinbring's ideas that we draw on is the communicative funnel pattern, which refers to "routinized, question developing, step-by-step regulated, strictly goal-dependent interactions between teacher and children. The funnel is a narrowing down of the action caused by the expected answer" (Steinbring, 1997, p. 92). As such, we view our data as two sets of guided explorations, in which the researcher guides and funnels the participant's actions and thinking in the direction of a specific goal (our research goal). The researcher entered the data collection session with an interest mainly in situations in which the participants claimed that a proposition was true (which would make the existentially quantified proposition incorrect). The things that the researcher was able to notice during the session was very likely influenced by the aforementioned interest. Furthermore, when the participant claimed that a proposition was true, the researcher immediately swayed them in the direction of "let's find an example (that would prove the proposition)" by trying out different test cases (i.e., plugging in specific numbers and seeing whether it results in a correct mathematical statement). Subsequently, a repeated pattern of interaction took place in the session, whereby the researcher would give the participants a test case, a pair of numbers, the participant would plug the test case into the proposition and then observe the output (a true or false statement). If the output was false (as it always was, given the truth value of each the propositions), the researcher gave the participant another test case.

Viewing the study through this particular frame, it is clear here that the researcher has a significant influence on the participants' actions, and that any 'impact of the test cases on paticipants' thinking' was influenced to a great extent by the researcher.

3.3. Data analysis

The overall analysis was guided by our purpose, which was to explore how the learners responded to the test cases and how these test cases, in conjunction with the influence of the researcher/interviewer (as explained above), may have stimulated the development

[2] A postgraduate diploma program in New Zealand involves taking Master's level courses in a particular specialization (e.g., Mathematics education). Some students enroll in the program as a pathway to a Master's degree in the specialization. Other students enroll in the program as a means of advancing their knowledge within the specialized area.

1. For some prime numbers, a and b, a x b is also prime.
2. For some pairs of odd numbers a and b, $a + b$ is odd.
3. $a^2 + 2ab + b^2 < 0$ for some values of a and b
4. $(a + b)^2 - (a - b)^2 < 0$ for some negative numbers a and b
5. If a/b and c/d are both less than or equal to 1, there exist positive numbers a, b, c, d, and e such that $a/b > c/d$ and $(ac + e)/2 > (bd + e)/2$

NB: you can assume that we are working only within the set of real numbers

Fig. 1. Data collection propositions.

and refinement of their propositions. We acknowledge here that we did in fact give the students the proposition, which raises the question of whether or not the propositions were the learners' (as we have alluded to previously). However, we argue that once the participants make a claim about the truth value of the proposition, we can regard it as the participant's proposition. But of course this is to be understood as a nuance of this particular research setting, decided on by the researchers, and thus represents a situation that differs from say one in which the participant genuinely creates an incorrect existentially quantified proposition.

The data were analyzed via thematic analysis (Braun & Clarke, 2012). First, we read the interview transcripts from the video recordings several times and identified every moment in the transcript in which: a) the participant made a claim about the truth value of the proposition; b) we offered a test case, or the participant generated a test case; and c) the participants responded to the test case (irrespective of whether the participant modified their proposition or not). For clarity, some examples of the aforementioned moments that we identified are given in Table 2. Then, we examined these moments closely, looking for similarities and differences in how the students responded to the different test cases (e.g., what they did with the test case; what they said about the test case) and what they did or did not do to their proposition. This examination involved identifying changes (or no changes) in terms of the proposition's subject, DoV, predicate and quantification.

Given the specific aim of this study—to explore how learners respond to test cases when they hold incorrect existentially quantified propositions—we will present in the next section only Max's work on the third proposition and Leo's work on the fourth proposition. We chose to focus on these parts of the data that were most relevant to the focus of our study—incorrect existentially quantified propositions—which in our study are the instances in which the participants claimed that a proposition was true. This corresponds to Leo's work on the fourth and fifth propositions, and Max's work on the third and fourth propositions. That is, these were the instances in which the participants made incorrect existentially quantified propositions. However, we focus mainly on Leo's work on the fourth and Max's work on the third proposition because these were the cases in which the iterative processes of students' thinking and how they were responding to the test cases were most apparent.

Before proceeding to present Max and Leo's working on the third and fourth propositions respectively, an important point needs to be made about the instances in which Max and Leo claimed that the proposition was false (i.e., these are instances in which Max and Leo made correct existentially quantified propositions) and the impact of the researcher. This occurred in the first two propositions for Max, and the first three propositions for Leo. We present Leo's work on the first proposition and Max's work on the second propositions as examples of these situations in Figs. 2 and 3 below. The impact of the researcher in conjunction with the goals of the research and the a priori-determined pattern of interaction are apparent in these examples. For instance, the researcher could have, but did not give the participants test cases, in either of the examples below, most likely because they had answered it correctly, which meant that the participants had made correct existentially quantified propositions (i.e., not the phenomenon of interest). This is perhaps a limitation of the current study and thus presents an avenue for future research. How might have the researcher offering the participants test cases affected the participant's conviction in the validity of their claims? As we will see in the next section, the researcher immediately offered a test case in instances where the participant claimed that the proposition was true, which perhaps may have funneled the

Table 2
Examples of moments of interest in the data.

Type of moment	Example
The participant makes a claim about the truth value of the proposition	Max [6:33]: What am I saying? [short pause]. Yeah…yes, it will be negative in [sic] some numbers, and positive for some other ones, but I don't know which ones it'll be positive for…or negative for. So, it's true [pointing at $a^2 + 2ab + b^2 < 0$ for some values of a and b], but I just don't know…
We offer a test case, or the participant generates a test case	Researcher [7:50]: Exactly! So, what happens if you have a equals two and b equals three?
The participant responds to the test case	Max [8:07]: Umm…that'll be positive. Yeah, I think if you have both positive numbers, then it'll be positive. Actually, yes if they're both positive numbers, then it'll be positive…so they both have to be both not positive numbers.

Time	Speaker	
2:35	L (Leo)	Should I start with this one?
2:40	R (Researcher)	Up to you…any order you like.
2:42	L	Might as well.
2:50	R	And you don't have to finish one before moving to another…you can do what you like…so long as you can verbalize your thinking as much as possible, please.
2:54	L	OK. I think this is ["For some prime numbers, a and b, a x b is also prime"] false, right?
2:57	R	Can you explain?
3:10	L	Wait. [reads the statement again] Yeah, I think false…because if you have two and three…they are prime but if you multiply them, you get six, which is not prime because…the factors are 1, 2, 3 and 6.
3:42	R	So what makes it prime…not prime sorry?
3:45	L	Pardon?
3:48	R	So how do you know if a number is prime or not?
3:58	L	It has more than two…what's the word? Factors. More than two factors…one and the number.
4:04	R	OK.
4:07	L	Yeah, so if you take two prime number [sic] and multiply, then they will be factors of the number.

Fig. 2. Leo's work on proposition 1.

Time	Speaker	
3:25	M (Max)	I know this one ["For some pairs of odd numbers a and b, a + b is odd"] is not true because I can remember working on it [laughs]
3:30	R (Researcher)	Oh, okay. So do you think it's true or false?
3:45	M	Ah, false. If you add any two numbers, you will always get an even number. You add any two odd numbers, you are basically adding two even numbers plus two. You know what I mean?
4:09	R	Say that again.
4:14	M	Add two any two odd numbers, is adding two even numbers and two.
4:27	R	OK. I see what you're doing…you subtract one from each number.
4:30	M	Yes.

Fig. 3. Max's work on proposition 2.

participants' actions towards revision. In other words, giving them a test case may have signaled to the participants that they were incorrect. Similarly, perhaps giving them test cases in instances where they claimed the proposition was false may have also funneled the participants' actions toward revising. Perhaps not giving them a test case may have signaled to the participants that they had made a correct proposition.

4. Results

In this section, we present and analyze first Max's work on the proposition "$a^2 + 2ab + b^2 < 0$ for some real number values of a and b", and then Leo's work on the proposition "$(a + b)^2 - (a - b)^2 < 0$ for some negative numbers a and b". Each subsection contains transcripts from the task-based interviews followed by our analysis.

4.1. Max: episode 1

1. Max: Oh, that's the parabola one [pointing at $a^2 + 2ab + b^2$]! Yeah, so that's the one that's like a reverse U…like it goes from the positive side to the negative.
2. Researcher: What do you mean?
3. Max: It's an upside down U graph [gesturing an upside down U], starts in negative and then goes up into the positive, and then curves back into the negative, I think.

4. Researcher: Interesting! So, are you saying, if you plug in some numbers, you'll get…some negative and some positive, or all positive, or all negative?
5. Max: What am I saying? [short pause]. Yeah…yes, it will be negative in [sic] some numbers, and positive for some other ones, but I don't know which ones it'll be positive for…or negative for. So, it's true [pointing at $a^2 + 2ab + b^2 < 0$ for some values of a and b], but I just don't know…
6. Researcher: Ah, I see. Yep, so you're saying there are some values for a and b such that this [$a^2 + 2ab + b^2$] will give you positive values and then some will give you negative values.
7. Max: Yeah, I can just visualize it. It's a U, upside down one, so it will be like that [gestures an upside down U].
8. Researcher: OK, let's explore that a bit. Cool. So, let's try [and] determine the values that will give you the negative.
9. Max: Okay. Yeah, that makes sense…if I know which ones give me negative, then the rest will give me positive. Is that what we're trying to do?
10. Researcher: Exactly! So, what happens if you have a equals two and b equals three?
11. Max: Umm…that'll be positive. Yeah, I think if you have both positive numbers, then it'll be positive. Actually, yes if they're both positive numbers, then it'll be positive…so they both have to be both not positive numbers.
12. Researcher: Okay. Not positive. What about if they're both zero?
13. Max: Both zero…that will be…two…because that one [points at $2ab$] will be two.
14. Researcher: Which one will be two?
15. Max: That one [points at $2ab$]…two times…oh wow, that's zero. So zero, all zero [laughs].
16. Researcher: What does that mean in terms of the numbers that will make it negative?
17. Max: Oh yeah! Ummm so not positive…not both zero…Yeah umm…oh yeah negative numbers. Maybe it will be negative if the numbers are both negative…some pairs of negative numbers will make it negative.

In the excerpt above, Max begins by visualizing the expression $a^2 + 2ab + \mathrm{b}^2$, claiming that it is a parabola, part of which would take on positive values and part of it would take on negative values. Hearing this, the researcher (the first author) begins to funnel Max's actions (see Line 4) towards a particular method of plugging values of a and b into the expression, observing the output, and then acting accordingly, depending on the value of the output). The beginning of Max's response in Line 5 ("What am I saying?") suggests that the method proposed by researcher does not exactly align with Max's thinking. However, the subsequent interactions show that Max eventually realigns his thinking in the direction of the proposed method. Here, we can say that the researcher had communicated to Max what was an acceptable, expected approach and Max acts accordingly.

In line 7, Max reaffirms his claim that the proposition was true, but he could not specify any particular values of a and b that would make the proposition true (i.e., satisfy the predicate). At this point, the researcher's goal was to try and help Max recognize that the proposition was false, by repeatedly trying and failing to construct an element of the domain of validity (i.e., some pair of real numbers) that would satisfy predicate, hence the remark in Line 8: "So, let's try [and] determine the values that will give you the negative." Again, we see that the researcher not only reemphasizes the specific goal towards which Max is expected to work, but also (perhaps less explicitly) communicates the acceptable method that Max is expected to use as he works towards this goal. In Table 3 below, we summarize the development of Max's proposition and the affordances of the test cases offered by the researcher. We elaborate on this summary in the following paragraphs.

Upon being given the first case ($a = 2, b = 3$), Max tests and observes that this test case does not satisfy the predicate (Line 11). He then claims in Line 11 that the proposition will be false for any pair of positive numbers, thus ruling out this class of potential examples. As such this test case afforded, not just the ruling out of a specific case, but also a class of potential examples. That is, Max has categorized the single test case $\mathrm{a} = 2, \mathrm{b} = 3$ as a representative of all pairs of positive numbers. While we cannot ascertain why Max has done this, we hypothesize that Max may have recognized a specific property of this particular pair (i.e., they are both positive numbers). It is reasonable to suggest that Max could have just as easily categorized this case as "all pairs of real numbers (a, b) such that $\mathrm{b} = \mathrm{a} + 1$". The idiosyncratic nature of this categorization is apparent when Max considers the next test case.

When Max, in Line 12, is given the next test case, $a = 0$ and $b = 0$, which is an element of the current search class (i.e., some pairs of

Table 3
Development and refinement of Max's proposition in response to the test cases of episode 1.

Proposition	Search class	Test case(s)	Affordance(s) of the test case(s)
$\mathrm{a}^2 + 2\mathrm{ab} + \mathrm{b}^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b)	a = 2, b = 3	• Ruling out a potential concrete example: a = 2, b = 3 • Using a property of the test case (both positive) to create a class of potential examples:{(a, b) | a > 0 and b > 0} • Ruling out a potential class of examples:{(a, b) | a > 0 and b > 0} • Creating a new search class: some pairs of real numbers (a, b) where a and b are non-positive numbers.
$\mathrm{a}^2 + 2\mathrm{ab} + \mathrm{b}^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) where a and b are non-positive numbers	a = 0, b = 0	• Ruling out a potential concrete example: a = 0, b = 0 • Creating a new search class: some pairs of real numbers (a, b) where a and b are negative numbers.
$\mathrm{a}^2 + 2\mathrm{ab} + \mathrm{b}^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) where a and b are negative numbers	n/a	n/a

real numbers (a, b) where a and b are non-positive numbers) Max acknowledges that test case does not satisfy the predicate (Line 15). In Line 16, the researcher asks what seems to be a funneling question (trying to get Max to say something about the DoV of his proposition). As we interpret his response (Line 17), on the one hand one could argue that Max rules out another class of potential examples: all pairs (a, b) where $a = 0$ or $b = 0$ because of the search class that Max proposes next (i.e., pairs of negative numbers). However, on the other hand, we claim that Max only explicitly rules out the test case a = 0 and b = 0; and then immediately suggests searching for an example in the set of negative numbers (in Line 17). Note, Max's response here differs from his earlier response in which he explicitly ruled out a class of potential examples (see Line 11): "if they're both positive numbers, then it'll be positive". This is an important point to bear in mind for later in Episode 2 when Max claims that he has "tried everything". Moreover, the idiosyncratic nature of the categorization (or in this case, no categorization), and consequently the class of examples (or in this case, a single example) that is ruled out is evident. It is not unreasonable to think that Max could have just as easily categorized the case (a = 0 and b = 0) as the class of all pairs such that one of a or b is zero.

4.2. Max: episode 2

18. Researcher: Okay so if they're both negative…what about if a equals negative one and b is negative two?
19. Max: That will be … [writes out $-1^2 + (2 \times -1 \times -2) + \text{-}2^2 = 1 + 4 + 4 = 9$] nine! Nine? Positive nine? What? [re-examines what he had written]. Okay, so that's right, nine, positive nine. So even if they're negative, it's still positive? What?
20. Researcher: What about…?

Table 4
Development and refinement of Max's proposition in response to the test cases of episode 2.

Proposition	Search class	Test case and realization	Affordance(s)
$a^2 + 2ab + b^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) where a and b are negative numbers	$a = -1, b = -2$	• Ruling out a potential concrete example: $a = -1, b = -2$ • Attending to the individual components (a^2, b^2, 2ab, and +) of the expression and gaining an insight into the collective effect of the components on the output • Ruling out a potential class of examples: $\{(a, b) \mid a < 0$ and $b < 0\}$ • Creating a new search class: some pairs of real numbers (a, b) where one is positive and the other is negative
$a^2 + 2ab + b^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) where one is positive and the other is negative	$a = 1, b = -1$	• Ruling out a potential concrete example $a = 1, b = -1$ • Attending to the individual components of the expression and gaining an insight into the conditions that the example (which he is trying to identify) must satisfy: $a^2 + 2ab + b^2 < 0$ if one of a and b is positive, the other is negative, and $a^2 + b^2 < \lvert 2ab\rvert$ • Creating *a* new search class: some pairs of real numbers (a, b) where one number is between 0 and 1, and the other is a large negative number
$a^2 + 2ab + b^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) where one number is between 0 and 1, and the other is a large negative number	$a = 0.5, b = -100$	• Ruling out a potential concrete example $a = 0.5, b = -100$ • Noticing a particular property of the test case ("maybe because *b* was big, that made *b* square big") and ruling out a potential class of examples {(a, b) | one is positive, the other is negative, and a and b are too far apart in absolute value} • Gaining insight into other conditions that might need to be satisfied by the variables in order to produce the desired result ("maybe… maybe because *b* was big, that made *b* square big…so maybe *b* shouldn't be too big? Maybe they should actually be close to each other, but one is negative and the other is positive") • Creating a new search class: some pairs of real numbers (a, b) such that one is positive, the other is negative, and $0 < \lvert a\rvert, \lvert b\rvert < 1$
$a^2 + 2ab + b^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) such that one is positive, the other is negative, and $0 < \lvert a\rvert, \lvert b\rvert < 1$	$a = 0.5, b = -0.5$	• Ruling out a potential concrete example $a = 0.5, b = -0.5$ • Ruling out a potential class of examples {(a, b) | one is positive, the other is negative, and $0 < \lvert a\rvert = \lvert b\rvert < 1$} • Producing a particular output ("0" – not positive) that motivates a feeling of "getting close" and thus further exploration • Creating a new search class: some pairs of real numbers (a, b) such that one is positive, the other is negative, and 0 < |positive number| < |negative number| < 1
$a^2 + 2ab + b^2 < 0$ for some pairs of real numbers (a, b)	Some pairs of real numbers (a, b) such that one is positive, the other is negative, and 0 < |positive number| < |negative number| < 1	$a = 0.5, b = -0.75$	• Ruling out a potential concrete example $a = 0.5, b = -0.75$ • Sensing that "[the output] is always positive" and that he [Max] has "tried everything"

21. Max: Oh I get it. If they're both positive…I mean both negative… If they're both negative then… it'll be just like they're positive…because you're squaring them there [points at a^2 and b^2] and [pointing at $2ab$]…so I think one has to be positive and the other one is negative!
22. Researcher: Okay, I see! So how about if a is negative one and b is one?
23. Max: That will be [writes out $-1^2 + (2 \times -1 \times 1) + 1^2 = 0$]…zero. So it's still not negative.
24. Researcher: What are your thoughts?
25. Max: I'm thinking there has to be one negative and one positive to make it negative… not that one [a = -1, b = 1] obviously… not one and negative one… but there has to be other number pairs that will make it negative.
26. Researcher: What if you try another pair?
27. Max: Oh, I think… I think a must… a can be positive… and b can be negative… so that will make that (pointing at $2ab$) negative, and it must be bigger… like it must have more weight than a square and b square…since these squares have positive weight.
28. Researcher: Hmmm…okay.
29. Max: Yup…I think a…or b… it doesn't matter…b can be, needs to be a big negative and a is a small positive like a decimal.
30. Researcher: Sweet…so how about if b is negative one hundred and a is positive zero point five? Is that an example of what you're thinking?
31. Max: Yeah, that's the one! [inputs the values] Far out…it's still negative…I mean positive…it's still positive. So the big negative and small positive combo doesn't work. Maybe…maybe because b was big, that made b square big…so maybe b shouldn't be too big? Maybe they should actually be close to each other, but one is negative and the other is positive.
32. Researcher: OK.
33. Max: Yeah [writes out "one number post, other number neg, a close to b"]… so we could make b less than one so that b is more than b square…and then make a less than negative one…like not less than negative one…but…you know what I mean? So that a is greater than a square.
34. Researcher: Nice…so how about if one, if b is negative half and the other, a, is positive half?
35. Max: [writes out 0.5^2 + -0.5^2 + 2 × -0.5 × 0.5]…that's quarter plus quarter, plus minus half…which is zero! So still not negative but I think I'm getting close…I think this was zero because they were the same. Like negative and positive of the same number. I think the negative needs to be a little bigger than the positive aye? So if I make b three quarters…negative three quarters and a positive half. That'll give me… [writes out calculations and uses calculator] … one over sixteen… [short pause] [laughing]. This is frustrating… [short pause]. I'm starting to think it's never going to be negative! Mmmm, yeah… I think it's always positive…I've tried everything.
36. Researcher: Okay…so as I said before, you decide…if you're happy…you can move on…or work on this some more. I think we have a couple of minutes left.
37. Max: I'm going to move on to this and try this one before I run out of time.
38. Researcher: All good!
39. Max: Can I, or do you want me to keep working on that one?
40. Researcher: Whatever you feel comfortable with.
41. Max: Yeah, I'm gonna switch to this…I think I've done enough on that one.

We summarize the development of Max's proposition and the affordances of the test cases offered by the researcher in Table 4 below. We elaborate on this summary in the following paragraphs.

The episode begins with Max considering the test case a = -1 and b = -2, upon which he determines that it does not satisfy the predicate and then in Line 19, the case affords a generalization–"even if they're (the numbers in the pair are) negative, it's (the output's) still positive?"

In line 21, we see that this test case also affords Max an insight into, and justification for, why negative pairs always would yield a positive output (i.e., they are effectively equivalent to positive pairs because of the variables that make up the left-hand side of the expression (i.e., a^2, b^2 and 2ab) and the operation (i.e., addition). Max here seems to gain this insight through what seems to be a consideration of not a concrete test case, but rather a generalized case (in which both numbers in the pair are negative). We point this out because it seems to differ from the first time Max ruled out a general class of potential examples in the first episode (see row 1 of Table 4). In that first iteration, unlike what occurs in this Episode, Max does not explicitly consider, or test, a generalized case. Furthermore, this moment Max's work in Line 21, also differs from his work in Episode 1 in the sense that Line 21 is the first time Max explicitly attends to the individual components of the expression and the effect they each have on the output.

At the end of Line 21, Max suggests the next search class: "one has to be positive and the other one is negative!" and the researcher responds by offering Max another test case a = - 1 and b = 1. Max tests the test case. Upon realizing that this test case does not satisfy the predicate, he rules it out of a = -1 and b = 1, but he does not immediately rule out a general class of potential examples, as he is still convinced that a pair in the current search class would satisfy the predicate: "I'm thinking there has to be one negative and one positive to make it negative".

In Line 27, Max attends to the components of the expression and using not a concrete test case in his consideration (like $a = -1$ and $b = 1$) but rather a generalized case in which a is positive and b is negative to gain insight into the structure of the expression and the effect that the components have on the output. This affords Max an insight into the conditions that the example (which he is trying to identify) must satisfy: one of a and b needs to be positive, the other negative, and $a^2 + b^2 < |2ab|$.

Though attending to the individual components of the expression is similar to what happened earlier in this episode (Line 21), it differs in a couple of ways. In line 21, Max tests a generalized case ($a < 0$, $b < 0$) and concludes that it will always be positive for the general class {(a, b) | $a < 0$ and $b < 0$} under consideration. He rules out this general class and then suggests a different search class. However in Line 27, Max begins to test a generalized case (i.e., one number is positive and the other is negative) but he does not complete the testing. He notices the effect that the generalized case has on 2ab and the individual components but he does not reach a conclusion about a generalized output (as he did in Line 21). Instead, Max specifies a general condition that would make the output negative (i.e., $a^2 + b^2 < |2ab|$).

In line 29, Max goes on to suggest another search class "[one number needs] to be a big negative and [the other] is a small positive like a decimal". In response, the researcher offers a test case a = 0.5, b = -100. Max tests the case and notices that the output is "still positive" and immediately declares that the "big negative and small positive combo doesn't work". So here, the test case affords the ruling out of a potential concrete example and also the immediate ruling out of a potential class of examples{(a, b) | one is positive, the other is negative, and a and b are too far apart in absolute value. We also see here (latter half of Line 31) that the test case affords Max some insight into other conditions that might need to be satisfied by the variables in order to produce the desired result. In considering the concrete test case ($a = 0.5$, $b = -100$) Max attends to a property of this particular case ("Maybe…maybe because *b* was big, that made *b* square big") that he thinks contributed to the undesirable result, hence he rules out a generalized class of potential example and then proposes "maybe *b* shouldn't be too big…Maybe they should actually be close to each other, but one is negative and the other is positive."

In line 33, without the researcher offering Max a test case. Max further narrows down the search class to "Some pairs of real numbers (a, b) such that one is positive, the other is negative, and $0 < |a|, |b| < 1$". Max arrives at this search class by attending to the individual components of the expression and gaining an insight into the conditions that the example (which he is trying to identify) must satisfy: "so we could make *b* less than one so that *b* is more than *b* square…and then make *a* less than negative one…like not less than negative one…but…you know what I mean? So that *a* is greater than *a* square." Following this, the interviewer suggests a test case: a = 0.5, b = -0.5. Max tests this case and observes that the output is zero. The fact that the output is zero leads to a remark: "So still not negative but I think I'm getting close". As such, this particular test case seems to afford Max a feeling that he is "getting close". As he did in the previous iteration, Max uses a property of the test case to give him insight into why the output was zero – he attributes this to the fact that $|a| = |b|$. At this point, he implicitly rules out *a* general class and proceeds to suggest another search class, when he says, "think the negative needs to be a little bigger than the positive". This is the first time Max generates his own test case, a = 0.5, b = -0.75, rather than being given one by the researcher. It seems that Max has fully taken on board the particular method that the researcher intended for him (as the beginning of the task session) to use.

Max proceeds to test his case ($a = 0.5$, $b = -0.75$), observes that it does not satisfy the predicate, and it is at this point that he says "I'm starting to think it's never going to be negative! Mmmm, yeah… I think it's always positive…I've tried everything." It is particularly interesting to analyze the point at which Max declares, "I've tried everything". If we consider all the cases that Max tested, and explicitly ruled out, we find that he has not actually "tried everything" (i.e., comprehensively ruled out all classes of potential examples). For example, recall from the previous episode that Max did not explicitly rule out the class of examples where either a or b was zero; he only ruled out $a = 0$ and $b = 0$. However, what is true is that Max has considered and ruled out a small group of test cases, and in some situations, he categorized these individual test cases and ruled out the corresponding general classes. So when Max says, "I've tried everything" we interpret it as some sort of feeling about what he has generally done. That is, he has tried different alternatives to support his initial claim, none of which have succeeded, at which point he considers the possibility that his initial claim was incorrect.

Furthermore, in lines 36–41, after Max declares that he has "tried everything" we get a sense of how the setting (the task etc.) and the researcher impacted the direction of Max's activity. The researcher, in line 36 tells Max that he can stop here if he wishes. This is the first time during Max's work on Proposition 3 that the researcher declares this as an option for Max. Is it a coincidence that it happens just as Max first explicitly begins to wonder about whether the output is "always positive"? We wonder what would have happened if the researcher had urged Max to continue, or if the researcher had pointed out the cases that Max had not considered. These interactions provide further evidence of the funneling that took place in this session. The researcher contributes to the decision as to whether or not Max has done enough, and whether it is acceptable to move on. What would have happened if the researcher had reminded Max of this (option to try the next proposition) prior to Max reaching this point? The limited time (30 min) also likely contributes to Max's decision to move on (see line 37) – "move on to this [the next proposition] and try this one before I run out of time". Perhaps if there were no other propositions to work on, Max would have continued. We also acknowledge that it can be misleading to interpret what is said in Line 36 as an indication that Max had total agency over their work and the direction his work takes. Instead, we again highlight the role of the researcher here—notice how the researcher chose to make this option (in Line 36) available at a specific time. Overall, the affordances of the test cases on the participant's decisions must be understood in relation to the researcher's influence and decisions.

5. Leo

Leo observes the proposition "$(a + b)^2 - (a - b)^2 < 0$ for some negative values of *a* and *b*", then says:

1. Leo: That's true. I think you can make this part [points to $(a - b)^2$] really large, then the whole thing will be negative, because you have to subtract the very large quantity from a smaller positive number. So, I'm pretty sure it's true. Like, you can locate a couple of negative numbers that will make it true.
2. Researcher: OK. Let's try and locate some numbers then. So how about *a* equals negative one and *b* equals negative one?

3. Leo: Ummmm [writes on paper]…I got four minus…oh that's zero [pointing at (*a* -*b*)]…so just four! And that is clearly not less than zero, hmmmm. Yeah okay, so we can maybe eliminate those negative numbers that are equal like negative one and negative one. But I still think you can find some negatives that will make the whole thing negative.
4. Researcher: So you are saying…correct me if I'm wrong please…that this [*pointing at the conjecture*] is going to be true for some pair of negative numbers, but you are ruling out the pairs that are equal? Am I understanding your…am I understanding you correctly?
5. Leo: Yes. Yes, that's correct.
6. Researcher: OK. Good. So, if we try…how about if we try, say…hmmm…[*writes* $a = -1$, $b = -2$].
7. Leo: They are not equal and they are both negative… I'll get [writes $-3^2 - 1^2 = 8$]. I got eight! What? Woah, I didn't expect that.
8. Researcher: Okay, what were you expecting?
9. Leo: Negative. Well, *a*… *a* minus *b* is going to be a positive number if *a* is bigger than *b*, both negative…So maybe it will work if the two negative numbers have *a* being less than *b*… because we want a negative number.
10. Researcher: Can you explain that again, please?
11. Leo: Yeah, that was confusing, eh? [laughs]. I was thinking if *b* is less than *a*, than *a* subtracting *b* is going to be positive because we are subtracting a negative…so we get a positive number. So, I think if we make *a* less than *b*, then we get a negative number.
12. Researcher: OK [*in response to Leo's claim in Line 11*], so something like a = -9 and b = -1?
13. Leo: Yes.
14. Researcher: Can you try that?
15. Leo: Ten square minus negative eight square…thirty six. Still no. And it's actually getting more positive. Oh, you're kidding me! [*Small period of silence ensues. Leo is seen writing down* a = -0.9, b = -0.1].
16. Researcher: Are you trying another example there?
17. Leo: Yeah…that doesn't work either [pointing at *a* = -0.9, *b* = -0.1]…because you get one minus point six four…and, that makes it zero point three six [0.36]. So, it doesn't work if *a* is less than *b*. What about if *b* is less than *a*? That doesn't work…we already tried that before…and, when they are equal negative numbers. Why? Maybe it's always positive.
18. Researcher: Interesting.
19. Leo: Anytime *a* and *b*…Wait a minute, I don't think it'll…I think it…It will never be negative if *a* and *b* are both negative because *a* plus *b* looks like it will be always be larger than *a* minus *b*. Oh, maybe it will work for if *a* and *b* that [sic] are both positive…but you will need the difference between *a* and *b* to be big so that [writes -(*a* - *b*)] is a big number…. so *b* should be much larger than *a*.
20. Researcher: So how about $a = 5$ and $b = 100$?
21. Leo: I need a calculator. Can I use my phone?
22. Researcher: Of course!
23. Leo: Oh, I don't need it because 105 squared is bigger than 95 squared. That was silly…*a* plus *b* is going to be larger than *a* minus *b* if they are both positive! Maybe if one is negative and the other is positive?
24. Leo: What's this one? [points at fifth proposition]. This looks interesting…and confusing [laughs…then reads out the proposition].
25. Researcher: Want to have a go? Or you can continue working on the fourth one
26. Leo: I've had enough of that [points at fourth proposition].
27. Researcher: Fair enough.

As summarized in Table 5, right from the beginning of this episode in Line 1 Leo attends to the components, $(a + b)^2$ and $(a - b)^2$ that make up the expression, which he uses to establish the claim that a negative output is possible (i.e., it is possible to make $(a - b)^2$ larger than $(a + b)^2$). In line 2, the researcher funnels Leo's work in the direction of trying to "locate some numbers" and offers the first test case a = -1 and b = -1. Leo tests the first case and subsequently rules out both the test case and a general class of potential examples: "negative numbers that are equal" (Line 3). The ruling out of the general class of potential examples here is done without actually explicitly testing a generalized case (like Max did in some instance in his work). In lines 4–6, the researcher further funnels Leo's activity, establishing among other things, what one is expected to do if a test case does not satisfy the predicate, but one feels that there is a test case that satisfies the predicate—i.e., try another pair of numbers that does not belong to the class of potential examples previously ruled out. An interesting point to note here that differs from most our observations in Max's work, is that after Leo (in Line 3) rules out the first general class of potential examples, he does not immediately suggest a different search class. As such, the test case offered by the researcher in line 6 ($a = -1$, $b = -2$) is not an element of a search class suggested by Leo, but rather simply not an element of the class most-recently ruled out.

After noticing that the test case ($a = -1$ and $b = -2$) does not satisfy the predicate in Line 7, Leo attends to the components of the expression, and realizes that (a-b) is going to be positive if a is greater than b and they are both negative (see beginning of Line 11). He goes on to rule out $\{(a, b) \mid b < a < 0$. In the second half of line 11, motivated by the class he has just ruled out, Leo explicitly states the next search class: some negative real number pairs (a, b) where a is less than b. Then, the researcher offers Leo a test case a = -9, b = -1 in line 12, which Leo tests, and rules out. Leo then (in Line 15) immediately tries another test case: a = -0.9 and b = -0.1, which again does not satisfy the predicate. In line 17, Leo rules out the set of negative real number pairs (a, b) where a is less than b without testing the generalized case. He immediately wonders about looking within another search class: $\{(a, b) \mid b < a < 0\}$, but then he quickly realizes that he had already considered it, and also recalls that he had ruled out$\{(a, b) \mid a < 0$ and $b < 0$; and $a = b$.

At this point (end of Line 17) he wonders, "maybe it's always positive". Here we see that these two test cases, a = -9, b = -1; a = -0.9

and b = -0 afford the ruling out of the general class of potential examples which these two cases represent—{(a, b) | a < 0 and b < 0; and a = b}. The ruling out of this general class, in conjunction with the realization that two other classes have been ruled out, which collectively make up the set of negative pairs affords the feeling that "maybe it's always positive".

In line 19, Leo attends to the components of the proposition, and tests the generalized case (a and b are both negative), upon which he realizes that if a and b are negative, a + b will always be greater (in absolute value) than a – b, which means that $(a+b)^2 - (a-b)^2$ will never be negative. In essence, here Leo provides a justification for why he feels that the proposition is false (i.e., $(a+b)^2$ will always be greater than $(a-b)^2$ when a and b are negative). It is interesting to note here that this justification comes after Leo had exhausted and ruled out all classes of potential examples within the set of negative number pairs. However, at the end of Line 19 that Leo does not stop working, as he wonders "maybe it will work…if a and b…are both positive", which in effect is a revised proposition. The new proposition contains the same predicate as the initial one, but the subject has changed from "some pairs (a, b) of negative numbers" to "some pairs of positive numbers (a, b)" and then subsequently to "some pairs of positive numbers (a, b) where b is much larger than a" which he arrives at by (see end of Line 17) attending to the components of the proposition and making explicit a condition that he claims needs to be satisfied (i.e., -(a - b) needs to be a big number). While we cannot ascertain whether Leo is aware that he has changed the proposition, we interpret that he is trying to locate a domain of validity within which he can construct a concrete example that satisfies the predicate. This differs from what happens in Max's work, because Max never leaves the confines of the initial subject. This may be somewhat unsurprising given that in Max's work, the initial subject of the propositions is the set of all real number pairs, while for Leo the initial class is a subset of real number pairs. So when Max feels that the initial proposition is false (and that he has "tried everything"), it effectively means that he feels he has ruled out the entire set of real number pairs. However, for Leo feeling that the initial proposition is false equates to ruling out only a subset of the set of real number pairs, thus allowing him to explore other subsets of real number pairs in which an example (that satisfies the predicate) could potentially exist.

In line 20, the researcher gives Leo a test case a = 5 and b = 100 which he notices (line 23) does not satisfy the predicate. Leo then attends to the components of the proposition, and tests the generalized class (where a and b are both positive), upon which he realizes that if a and b are positive, a + b is going to be larger than a – b, which means that $(a+b)^2 - (a-b)^2$ will never be negative. At the end of Line 23, Leo once considers another proposition, by again changing the subject (from pairs of positive numbers to pairs of numbers in which one is negative and the other is positive). This is ultimately the point at which Leo stops working on proposition 4.

The stopping point (see lines 24–27) in Leo's work is again interesting to reflect on in relation to the overall interactions between Leo and the researcher within the context of this particular research setting. The fact that there was another proposition (proposition 5) that Leo had not reached perhaps meant that he felt he needed to at least attempt all of the propositions. The researcher does not explicitly push Leo in either direction (continue working on Proposition 4 or to move on to Proposition 5). However, it is interesting to note that Line 25 is the first time the researcher explicitly gives Leo the option of continuing or moving on, which just so happens to be the moment at which Leo considers a normatively correct proposition: $a+b)^2 - (a-b)^2 < 0$ for some pairs of numbers in which one is positive and the other is negative. We wonder what would have happened if the researcher had provided this option (to the next

Table 5
Development and refinement of Leo's proposition in response to the test cases.

Proposition	Search class	Test case	Affordances
$(a+b)^2 - (a-b)^2 < 0$ for some pairs of negative numbers (a, b)	Some pairs of negative numbers (a, b)	a = −1, b = −1	• Ruling out a potential concrete example: a = −1, b = −1 • Using a property of the test case (equal negative numbers) to create a class of potential examples • Ruling out a potential class of examples: {(a, b) | a < 0 and b < 0; and a = b}
$(a+b)^2 - (a-b)^2 < 0$ for some pairs of negative numbers (a, b)	Some pairs of negative numbers (a, b) where a ≠ b.	a = −1, b = −2	• Ruling out a potential concrete example: a = −1, b = −2 • Attending to the components of the expression, testing of a generalized case ((a, b) 0 > a > b) • Ruling out a class of potential examples: {(a, b) | b < a < 0} • Creating a new search class: {(a, b) | a < b < 0}
$(a+b)^2 - (a-b)^2 < 0$ for some pairs of negative numbers, b)	Some negative real number pairs (a, b) where a is less than b	a = −9, b = −1 a = −0.9 and b = −0.1	• Ruling out a potential concrete example: a = −9, b = −1 • Ruling out a potential concrete example a = −0.9 and b = −0.1 • Ruling out a class of potential examples {(a, b) | b < a < 0} without testing the general case. • Wondering whether $(a+b)^2 - (a-b)^2$ is "always positive" for pairs of negative numbers (a, b) • Attending to the components of the expression, testing of a generalized class (a and b are both negative), and ruling out a general class of potential examples: {(a, b)| a < 0 and b < 0)} • Revising the subject of the initial proposition
$(a+b)^2 - (a-b)^2 < 0$ for some pairs of positive numbers (a, b) where b is much larger than a	Some pairs of positive numbers (a, b) where b is much larger than a	a = 5 and b = 100	• Ruling out a potential concrete example a = 5 and b = 100 • Attending to the components of the expression, testing of a generalized case (a and b are both negative), and ruling out a class of potential examples {(a, b)| a > 0 and b > 0)} • Revising the subject of the current proposition
$(a+b)^2 - (a-b)^2 < 0$ for some pairs numbers (a, b) where one is positive and the other is negative	n/a	n/a	n/a

propositions) earlier? Perhaps Leo would have moved on to Proposition 5 without reaching this point. Or what would have happened if the researcher encouraged Leo to continue with this proposition? Perhaps Leo reaches a proof or maybe even return to an incorrect proposition.

6. Summary and discussion of findings

Towards the goal of extending the applicability of test cases to the context of existentially quantified propositions, we explored how test cases might support learners with refuting their incorrect existentially quantified propositions. We presented and analyzed data from two separate instances in which two participants initially made incorrect existentially quantified propositions and then were asked to find a valid example of their respective propositions (i.e., an element of the subject that satisfies the predicate). The participants were given, and sometimes generated their own, test cases which led to an iterative process of ruling out potential examples and classes of potential examples. Our analysis of this iterative process as it emerged within our specific research setting, comprising among other aspects, particular researcher-participant interactions, sheds light on the affordances of test cases with respect to the development, refinement, and evolution of the learners' existentially quantified propositions. We summarize the main findings below and discuss them with respect to relevant literature.

Firstly, the list below is a summary of the affordances (Ellis et al., 2019) of the test cases that we identified in Max and Leo's work:

- Afforded the ruling out of a potential concrete example (i.e., the test case)
- Afforded the ruling out of a class of potential examples. This did not always follow the ruling out of the test case because sometimes only a concrete example (the test case) was ruled out. Sometimes this affordance immediately followed the ruling out of the test case; and we inferred that the class of potential examples that was ruled out consisted of all potential examples that shared a specific property with the test case—a property that the participant was attending to in that particular moment. At other times, this affordance did not immediately follow the ruling out of the test case. Rather, after the ruling out of the test case, the learner explicitly tested a generalized case (e.g., if the test case was $a = 1$, $b = -1$; and the class was "the set of all pairs of numbers such that the two numbers are equal in absolute value"; the generalized case was (a, b) where $|a| = |b|$), ultimately concluding that this generalized case did not satisfy the predicate and thus ruling out the class of potential examples represented by the test case and the generalized case. In all of the situations in which a class of potential examples was ruled out, the idiosyncratic nature of the categorization (of the test case) – the properties that the participant attends to is evident.
- Afforded the creation of a new search class. Sometimes the ruling out of a test case led to the ruling out of a class (of potential examples) and the explicit suggestion of a new search class. At other times, there was no explicit suggestion of a new search class, but it was implied that a potential example may exist in the class(es) that have not been ruled out.
- Afforded the production of a particular output that motivated a feeling of "getting close" and thus further exploration (trying other test cases). For instance, if the goal was to find an example (input) that produced a positive output, getting "0" as the output after previously getting negative outputs consistently gives one the feeling of "getting close".
- Afforded attending to the individual components of the expression and gaining an insight into conditions that might need to be satisfied by the test case in order to produce the desired result or conditions that particular (prior) test cases did not meet and led to them being ruled out. This affordance was not a direct affordance of the test case. Rather, in certain situations the test case produced an undesirable output, which then motivated the learner to examine the individual components of the proposition in order to understand better why that test case did not work.
- Afforded the feeling that "[the output] is always positive" and that one has "tried everything". This was not a direct affordance of a particular test case, but rather of a collection of test cases and classes (of potential examples) that were ruled out. Even though the test cases in this present study did not lead to a proof, we argue that they were still useful because they provided an experience of failure; multiple failures likely motivated the learner to reconsider their initial conviction and consider the possibility that their proposition was incorrect.
- Afforded a revised proposition, specifically by changing the subject of the initial proposition. This was not the affordance of a single test case, but rather in conjunction with other factors (e.g., multiple test cases and classes of potential examples had been ruled out).

As evident in the list above, apart from the first affordance, all the other affordances are not affordances of a single test case. This observation perhaps suggests a different way of thinking about affordances, in the sense that an affordance that becomes apparent of a specific test case can rarely be attributed entirely to that test case—there tends to be other factors at play. As such we can think of test cases as part of a collection of factors that support the eventual refutation of an incorrect proposition.

In the CAPS framework (Ellis et al., 2019) there is a category under justification support called "revised conjecture", which refers to how an example sometimes affords a narrowing of the domain of applicability (which seems to be equivalent to what we are referring to in our study as the domain of validity—the set of all elements that satisfy the predicate). On the one hand, we could say that the test cases in the present study led to the narrowing of the domain of applicability. On the other hand, however, we could also argue that these test cases did not narrow the domain of applicability because the domain of applicability is unspecified. Instead, what we can say is that the test cases narrowed the set within which the DoV could exist (i.e., the set within which potential examples could reside is iteratively narrowed). This iterative narrowing process bears similarities with Lakatos' notion of *exception barring*, "reformulating [a] conjecture by restricting its domain to exclude the counterexample" (Larsen & Zandieh, 2008, p. 208). In the process of exception barring, the domain of validity (DoV) of a proposition is iteratively reduced by removing counterexamples and classes of counterexamples, in the search for a set all of whose elements satisfy the predicate. Similarly, for existentially quantified propositions, as the learner rules out classes of potential examples, the area of the subject within which a potential example might exist also shrinks. However note, the iterative reduction process in exception barring is well-defined for universally quantified incorrect propositions, because for example a learner might begin with the DoV of "all real numbers" and then through exception barring reduces the DoV to "all positive numbers" or the empty set (for propositions where no element of the subject satisfies the predicate). However, for existentially quantified propositions we cannot ascertain a clear reduction of the DoV in each iteration. For instance, moving from "some real numbers" to "some positive numbers" does not give sufficient information about the actual sizes of the two DoVs. The only apparent reduction that takes place in the context of incorrect existentially quantified propositions is the transition from "some elements of set X satisfy the predicate" to "no elements of set X satisfy the predicate".

Using Stylianides' framework (2008) in conjunction with Dogan and Williams-Pierce's (2021) interpretation of generic examples, one could say that the participants in the present study were trying to find a *generic example* of their respective propositions (i.e., an element of the subject that satisfies the predicate thus representing the general class of "some number pairs (which satisfy the predicate)"). In each iteration of the revision process that Max and Leo undertook, the researcher offered a test case to the participant, the participant checked and acknowledged that the test case does not satisfy the predicate, and more often than not the ruling out of the single test case was followed by the ruling out of a class (of potential examples) of which the test case is an element. In this ruling out, we can also say that some test cases served as generic examples (of the respective classes that were ruled out). While we are not saying here that the participants were aware of looking for a generic example, it is not unreasonable to say that the way in which the test cases were ruled out and categorized seems to suggest that the participants were treating these test cases as generic examples (i.e. "a particular case which represents a general case"). The similarities between some of the test cases (in the present study) with generic examples, can also be drawn with Lannin's (2005) notion of "empirical evidence", which refers to justification that is provided through the correctness of particular examples. For instance, in order to prove that a rule is true within a particular domain, particular examples from that domain are chosen and are shown to satisfy the rule. As such, we could say that the participants in the present study had assumed that the proposition was true and thus were trying to find an empirical example to prove that it is true. We could also say that in some instances in which the participants ruled out classes of potential examples, the participants were making a claim that for a particular class of examples, no element of the class would satisfy the predicate of the proposition. The test cases which were ruled out as such served as empirical examples because these concrete cases were used to justify the truth of their claim.

We conclude this paper with some remarks about the limitations of the present study and some potential directions for further study. Further research with larger numbers of participants and from a wider range of educational settings are needed to provide more in-depth insights into the iterative process that we identified in this present study and the particular affordances that emerged. Furthermore, it is worth exploring whether more nuanced choices can be made a priori regarding the test cases that are offered to learners when they make incorrect existentially quantified propositions. The test cases used in this study were merely examples of the current search class (subsets of the subject that the learner was currently considering). It would be worth exploring how test cases that were non-examples of the subject would impact students' propositions. It may also be worth exploring, in a manner similar to that in the research carried out by Zazkis and Chernoff (2008) (albeit in the context of existentially quantified propositions) whether test cases can be differentiated in terms of *how* they may motivate learners to revise their propositions, and also provide more nuanced explanations for how these test cases contribute to the classes that are ruled out in any revision iteration. It is also potentially useful to explore the other facets of the CAPS (criteria, affordances, purposes, strategies) framework, beyond merely affordances, with respect to incorrect existentially quantified propositions. For instance, what criteria do students use to choose or create particular test cases as they try to (dis)prove the validity of an existentially quantified proposition?

An obvious limitation of the present study was the massive influence that the researcher had on the participants' work. As such, further research could explore what impact test cases have in research settings where the researcher has less influence, for example by allowing the participants to generate their own test cases, and to examine whether the affordances of student-generated test cases differ from those found in the present study. Moreover, another avenue for further study is perhaps to also allow for alternative approaches (that are not based on test cases) and to explore how these alternative approaches can be used together with test cases to support learners in refuting their incorrect existentially quantified propositions. For example, Max could have re-presented $a^2 + 2ab + b^2$ as $(a + b)^2$, which may (or may not) have led to a realization that it could never be negative for any pair of real numbers. Similarly, in the second case, $(a + b)^2 - (a - b)^2$ could have been represented as $(a + b - a + b)$ $(a + b + a - b)$ which equates to $(2b)(2a)$. Speculating about why the learners did not try to re-express the two propositions is beyond the scope of this paper. This could be another avenue for further research.

CRediT authorship contribution statement

John Smith: Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Conceptualization. **John Griffith Tupouniua:** Writing – review & editing, Writing – original draft, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

Ethics approval

The research was approved by the Massey University Human Ethics Committee (Ethics Notification Number 4000025375).

Sources of funding

The research reported in this study was supported by Research Enhancement Funding from Massey University.

Financial or non-financial interests

The authors have no financial or non-financial interests related to work submitted for publication.

Declaration of Competing Interest

The authors have no financial or non-financial interests related to work submitted for publication.

Data Availability

The authors do not have permission to share data.

References

Balacheff, N. (1991). Treatment of refutations: Aspects of the complexity of a constructivist approach to mathematics learning. In E. von Glasersfeld (Ed.), *Radical constructivism in mathematics education* (pp. 89–110). Kluwer Academic Publishers.

Barkai, R., Tsamir, P., Tirosh, D., Dreyfus, T., & Tabach, M. (2008). Inservice teachers' judgement of proof in ENT. *In Proceedings of the Joint Meeting of the 32nd Conference of the International Group for the Psychology of Mathematics Education, and the XX North American Chapter, 4*, 345–353.

Braun, V., & Clarke, V. (2012). *Thematic analysis*. American Psychological Association.

Buchbinder, O., & Zaslavsky, O. (2019). Strengths and inconsistencies in students' understanding of the roles of examples in proving. *The Journal of Mathematical Behavior, 53*, 129–147.

Dawkins, P. C., & Roh, K. H. (2020). Assessing the influence of syntax, semantics, and pragmatics in student interpretation of multiply quantified statements in mathematics. *International Journal of Research in Undergraduate Mathematics Education, 6*, 1–22.

Dogan, M. F., & Williams-Pierce, C. (2021). The role of generic examples in teachers' proving activities. *Educational Studies in Mathematics, 106*(1), 133–150.

De Bock, D., Van Dooren, W., Janssens, D., & Verschaffel, L. (2002). Improper use of linear reasoning: an in-depth study of the nature and the irresistibility of secondary school learners' errors. *Educational Studies in Mathematics, 50*(3), 311–334.

Ellis, A. B., Ozgur, Z., Vinsonhaler, R., Dogan, M. F., Carolan, T., Lockwood, E., & Zaslavsky, O. (2019). Student thinking with examples: the criteria-affordances-purposes-strategies framework. *The Journal of Mathematical Behavior, 53*, 263–283.

Klymchuk, S. (2010). *Counterexamples in calculus*. American Mathematical Society.

Komatsu, K. (2016). A framework for proofs and refutations in school mathematics: increasing content by deductive guessing. *Educational Studies in Mathematics, 92*, 147–162.

Lakatos, I. (1976). *Proofs and refutation: the logic of mathematical discovery*. Cambridge University Press.

Lannin, J. K. (2005). Generalization and justification: The challenge of introducing algebraic reasoning through patterning activities. *Mathematical Thinking and Learning, 7*(3), 231–258.

Larsen, S., & Zandieh, M. (2008). Proofs and refutation in the undergraduate mathematics classroom. *Educational Studies in Mathematics, 67*(3), 205–216.

Mason, J., & Pimm, D. (1984). Generic examples: Seeing the general in the particular. *Educational Studies in Mathematics, 15*, 227–289.

Moala, J. G., Yoon, C., & Kontorovich, I. (2019). Localized considerations and patching: Accounting for persistent attributes of an algorithm on a contextualized graph theory task. *The Journal of Mathematical Behavior, 55*, Article 100704. https://doi.org/10.1016/j.jmathb.2019.04.003

Piatek-Jimenez, K. (2010). Students' interpretations of mathematical statements involving quantification. *Mathematics Education Research Journal, 22*(3), 41–56.

Peled, I., & Zaslavsky, O. (1997). Counter-examples that (only) prove counter-examples that (also) explain. *Focus on Learning Problems in Mathematics, 19*(3), 49–61.

Sellers, M. E., Roh, K. H., & Parr, E. D. (2021). Student quantifications as meanings for quantified variables in complex mathematical statements. *The Journal of Mathematical Behavior, 61*, Article 100802.

Steinbring, H. (1997). Epistemological investigation of classroom interaction in elementary mathematics teaching. *Educational Studies in Mathematics, 32*(1), 49–92.

Steinbring, H. (2007). Epistemology of mathematical knowledge and teacher–learner interaction. *ZDM, 39*, 95–106.

Stylianides, G. J. (2008). An analytic framework of reasoning-and-proving. *For the Learning of Mathematics, 28*(1), 9–16.

Tabach, M., Levenson, E., Barkai, R., Tsamir, P., Tirosh, D., & Dreyfus, T. (2012). An organizer of mathematical statements for teachers: The six-cell matrix. *International Journal of Mathematical Education in Science and Technology, 43*(6), 765–777.

Tupouniua. (2022a). Differentiating between counterexamples for supporting students' algorithmic thinking. *Asian Journal for Mathematics Education, 1*(4), 475–493.

Tupouniua. (2022b). What challenges emerge when students engage with algorithmatizing tasks? *Journal of Pedagogical Research, 7*(2), 93–107.

Vroom, K. (2022). A functional perspective on student thinking about the grammar of multiply quantified statements. *The Journal of Mathematical Behavior, 67*, Article 100992.

Ye, R. (2012). Universal and existential quantifiers revisited. In Y. Tso,T. (Ed.), *Proceedings of the 36th Conference of the International Group for the Psychology of Mathematics Education, 4* pp. 235–242). Taipei, Taiwan: PME.

Zazkis, R., Liljedahl, P., & Chernoff, E. J. (2008). The role of examples in forming and refuting generalizations. *ZDM - Mathematics Education, 40*, 131–141.

Zazkis, R., & Chernoff, E. J. (2008). What makes a counterexample exemplary? *Educational Studies in Mathematics, 68*(3), 195–208.