Mathematics majors' perceptions of the admissibility

of graphical inferences in proofs

Bo Zhen   Juan Pablo Mejia-Ramos   Keith Weber

Rutgers, The State University of New Jersey

Abstract. In this paper, we investigate mathematics majors' perceptions of the admissibility of inferences based on graphical reasoning for calculus proofs. The main findings from our study is that the majority of mathematics majors did not think that graphical perceptual inferences (i.e., inferences based on the appearance of the graph) were permissible in a proof, but the majority of mathematics majors did believe that graphical deductive inferences (i.e., inferences based on what must necessarily be entailed by a graph having a certain property) were permissible.

Key words: Proof; Proof scheme; Undergraduate mathematics education; Visual reasoning

1. Introduction

As proof is a central activity in mathematical practice, a primary goal of the undergraduate mathematics curricula is to improve students' appreciation, understanding, and production of proof (Harel & Sowder, 1998). Unfortunately, numerous studies demonstrate that mathematics majors struggle both to write proofs (e.g., Hart, 1994; Iannone & Inglis, 2010; Ko & Knuth, 2009; Moore, 1994; Weber, 2001) and evaluate proofs for correctness (e.g., Alcock & Weber, 2005; Inglis & Alcock, 2012; Ko & Knuth, 2013; Selden & Selden, 2003; Weber, 2010), even after taking proof-oriented courses in mathematics.

There have been several accounts of mathematics majors' difficulties, many of which suggest that students fail to write or understand proofs because they lack certain competencies. For instance, some researchers have suggested that mathematics majors fail to construct proofs because they have trouble unpacking the meaning of complex logical assertions (e.g., Selden & Selden, 1995; Zandieh, Knapp, & Rho, 2014), they do not know what can be assumed and what should be concluded (e.g., Selden & Selden, 1995), and they lack proving strategies and heuristics (e.g., Weber, 2001), amongst other factors. Further, mathematics majors cannot distinguish between valid proofs and invalid arguments because they do not attend to the overarching logical structure of the purported proofs that they read (e.g., Inglis & Alcock, 2012; Selden & Selden, 2003) or infer the mathematical principles that were used to deduce new claims in a proof from previous assertions (e.g., Weber & Alcock, 2005). After identifying competencies that students lack, some researchers have designed instruction to help students develop these

2

competencies (e.g., Hodds, Alcock, & Inglis, 2014; Selden, & Selden, 2013; Weber, 2006) on the assumption that if students developed these competencies, their performance on proof-related tasks would improve.

There are also researchers who believe mathematics majors' difficulties with proof are more fundamental than lacking a set of cognitive skills. These researchers posit that a primary cause of mathematics majors' difficulties is that their standards for proof are different from those of mathematicians (Harel & Sowder, 1998, 2007). As a consequence, students will produce arguments that they find convincing but that mathematicians would not find acceptable. Hence, helping students write proofs successfully involves shifting students' perceptions about the goal of the activity of proving.

1. 1. Students' proof schemes

In a highly influential paper, Harel and Sowder (1998) defined an individual's *proof scheme* to be to the ways in which that individual attempts to convince himself or herself and persuade others about the truth of a mathematical assertion. While students' methods of obtaining personal conviction and persuading others are related, they are not identical (Segal, 2000). In particular, students might recognize that an empirical argument is not an acceptable proof, but still find it personally convincing (Brown, 2014; Healy & Hoyles, 2000; Segal, 2000) and they may believe that proofs must include symbols or be written in a two-column format, even if they find arguments lacking symbols or written as a narrative paragraph to be convincing (e.g., Healy & Hoyles, 2000; Martin & Harel, 1989). In this paper, we are

concerned with the persuasive aspect of mathematics majors' proof schemes-- what types of inferences do students think are acceptable in a mathematical proof?

There is a large body of research on students' proof schemes with students who do not have training in advanced mathematics, including middle school and high school students, and preservice elementary teachers. A key finding from this research is that a substantial number of students in these populations appear to believe that arguments based on empirical reasoning are an acceptable form of proof (e.g., Healy & Hoyles, 2000; Knuth, Choppin, & Bieda, 2009; Martin & Harel, 1989; Recio & Godino, 2001; for a review and critique of this literature, see Weber, 2010). There have been a smaller number of studies with undergraduates who have completed proof-oriented mathematics courses (from here on, *advanced mathematics majors*) and the situation appears different for this population of students. When given proving tasks, such students rarely submit empirical arguments (Iannone & Inglis, 2010) and if asked to evaluate such an argument, they will reject it as not meeting the standards of proof (e.g., Bleiler, Thompson, & Krajcevski, 2014; Pfeiffer, 2011; Segal, 2000; Weber, 2010). There is debate as to whether advanced mathematics majors do this because they recognize the limitations of empirical reasoning (Weber, 2010) or are merely following the social conventions of their classes or the directives of their instructors (Brown, 2014; Segal, 2000).

Although there has been much research on students' perceptions of empirical arguments, the literature on students' perceptions of visual arguments has been comparatively sparse. The goal of the current paper is to address this void in the

literature by describing the types of graphical arguments that advanced

mathematics majors consider acceptable within a proof. This issue is of theoretical

interest in its own right, but it also has an important consequence for the literature

on proof production. Mathematics educators have argued that diagrammatic

reasoning, specifically graphical reasoning, can and should form the basis for the

proofs that students write (e.g., Alcock, 2010; Gibson, 1998; Raman, 2003).

However, for this strategy to be viable, students will need to be able to translate

their graphical arguments into an argument that satisfies the standards of proof,

something that advanced mathematics majors find difficult (e.g., Alcock & Weber,

2010; Zazkis, Weber, & Mejia-Ramos, 2014). Crucial to understanding how advanced

mathematics majors translate arguments into proofs involves knowing *which*

inferences these students believe need to be translated.

1. 2. The role of diagrams in doing mathematics

Both mathematicians and mathematics educators consider diagrams[1] to be

an important component of doing and understanding mathematics (e.g., Hadamard

1945; Stylianou 2002). A critical benefit of diagrams is that they provide an

individual access to view, compare, and integrate simultaneous pieces of

information with little cognitive effort. Such reasoning is often difficult when the

same information is presented sequentially and symbolically (Dreyfus, 1991; Larkin

& Simon, 1987). As a result, certain properties of mathematical concepts that are

transparently obvious with a diagram would be difficult to discern with non-visual

representations of this concept (e.g., Piez & Voxman, 1996). For instance, one can

---

[1] For the purposes of this paper, by diagram, we are referring to any visual representation of a
mathematical situation where properties of mathematical objects or relationships between
mathematical objects are dependent on the physical location of these objects within the diagram in

frequently see that a function is increasing easier from its graph than deducing it from its formula[2].

Explanations involving diagrammatic reasoning often have different virtues than those provided in a verbal-syntactic representation system. Such explanations are often more accessible and more concrete, particularly to students of mathematics (e.g., Hersh, 1993), and they can provide students with different types of learning opportunities (Weber, 2005). These explanations can also highlight aesthetics or underlying mathematical principles, as is illustrated by Nelsen's (1993, 2000) *Proofs Without Words.*

Many mathematicians claim that diagrams play an irreplaceable role in their mathematical reasoning (e.g., Burton, 2004; Hadamard 1945). Research on mathematicians' problem-solving has revealed that the construction and consideration of diagrams is commonplace and has substantial benefits (e.g., Schoenfeld, 1985). For instance, diagrams allow problem solvers to infer consequences, elaborate on mathematical ideas, create sub-goals, and encourage metacognitive reasoning (Samkoff, Lai, & Weber, 2012; Stylianou, 2002).

The use of diagrams has also been shown to facilitate students' proof-writing. In an illustrative study, Gibson (1998) explored how students in a real analysis course overcame impasses when writing proofs. He noted "using diagrams helped students complete sub-tasks that they were not able to complete while working with verbal-symbolic representation systems alone" (p. 284) by facilitating understanding, evaluating the truth of statements, generating ideas, and expressing

---

[2] Such judgments, however, are fallible. For instance, graphs that are non-monotonic due to very small wrinkles may appear to be increasing.

ideas (see also Alcock & Simpson, 2004). Of course, not every inference drawn from a diagram would be permissible in a proof without a separate deductive justification. In the following sections, we discuss mathematicians' and students' views on what types of justifications are permissible within a proof.

1. 3. The normative view on the permissibility of diagrams in mathematical proof

Amongst mathematicians, the normative view is that while diagrams are useful for the construction of proofs, they are expected to play an ancillary role in the presentation of proof, such as helping the reader understand the proof. In principle, the validity of the proof should not be altered by the removal of the diagram. Without necessarily endorsing this viewpoint, Inglis and Mejia-Ramos (2009a) asserted that the common view on the relationship between diagrams and proof presentation is this: "Pictures may be useful heuristic tools which suggest ways of understanding proofs but that they are nevertheless inappropriate when it comes to providing unequivocal reliable evidence to support a mathematical claim, let alone providing a proof" (p. 100). There is suggestive empirical evidence to support Inglis and Mejia-Ramos' claim. For example, these authors (Inglis & Mejia-Ramos, 2009b) conducted a series of experiments in which mathematicians were asked how persuasive they found various mathematical arguments to be. The mathematicians judged the visual arguments that they evaluated as significantly less persuasive than the conventional symbolic proofs that they read (for further evidence, see Inglis & Mejia-Ramos, 2009a). This was the case even though the visual argument was written by a famous mathematician and is generally regarded as mathematically correct. Recently, several philosophers have argued that visual

7

arguments should be perfectly convincing and hence ought to be acceptable in a proof (e.g., Azzouni, 2013; Feferman, 2012; Kulpa, 2009). However, as these authors are arguing against the status quo-- i.e., that such arguments are usually regarded as unacceptable-- the existence of these essays in support of visual arguments offers further evidence for the viewpoint that diagrammatic reasoning is not currently permissible in a proof.

There is, however, reason to doubt the universality of this viewpoint. While diagrammatic proofs are unusual in many mathematical domains, they are commonplace in others, such as knot theory (e.g., Rav, 1999). Nelsen's (1993, 2000) publication of *Proofs without words*, which is comprised entirely of diagrammatic arguments, is not aimed at specialists in a particular sub-discipline of mathematics, but rather for a broad mathematical audience. This illustrates that diagrammatic arguments can be convincing proofs for the wider audience of mathematicians, at least in some sense of the word "proof". Even in conventional domains such as Euclidean geometry, some published proofs implicitly rely on perceptual reasoning to make inferences, although the authors of the proof might not be aware that they are doing so (e.g., Herbst, 2004). Aberdein (2009) described picture proofs as an instance of a "proof*", which he defined as "species of alleged 'proof' where there is no consensus that the method provides proof, or there is a broad consensus that it doesn't, but a vocal minority or an historical precedent point the other way" (p. 1). As Tall (2013) argued, the validity of reasoning from diagrammatic arguments sometimes has formal backing. Some theorems, which Tall calls *structure theorems*, can be interpreted as ensuring that diagrammatic reasoning will not lead to

fallacious inferences. For instance, a powerful structural theorem is that any structure satisfying Peano's postulates is isomorphic to the set of natural numbers (Tall, 2001). This permits mathematicians to model Peano's axiomatic system with the natural numbers on the number number line and use their associated intuition about this line to draw inferences about the axiomatic system. Whether mathematicians recognize visual reasoning justified by structure theorems as admissible in a proof is an open question. In summary, it is not clear how many mathematicians adopt or disagree with the normative position that visual reasoning is not permissible in a proof. Empirical studies addressing how mathematicians actually feel about visual arguments, as well as what types of visual arguments mathematicians find acceptable for proofs in the courses that they teach, are important avenues for future research.

1. 4. Mathematics majors' perceptions of diagrams in proofs

As we noted earlier, the literature on how students perceive diagrammatic evidence is sparse. Harel and Sowder (1998) observed that some mathematics majors held a perceptual proof scheme, which they defined as convincing oneself or persuading others by appealing to a diagram without regard to how that diagram can be transformed. In other words, students holding a perceptual proof scheme would draw conclusions solely by the appearance of the diagrams. However, Harel and Sowder did not specify how common this proof scheme is with mathematics majors.

Several empirical studies suggest that many undergraduates do not believe that diagrams are permissible in a proof. Inglis and Mejia-Ramos (2009b) asked

mathematics majors how persuasive they found a visual argument to be. They found that the mathematics majors largely dismissed the visual argument as unpersuasive; when asked why they did so, some students claimed they were taught that diagrams were not allowed in a proof. Raman (2003) also found some students in calculus would try to write a proof entirely by logical and algebraic manipulation; she argued that they did so because they did not perceive a connection between visual arguments and the proofs that their professors expect. Both the Inglis and Mejia-Ramos (2009b) and Raman (2003) studies suggest that some university students overgeneralized the maxim that "diagrams cannot prove" to believe that diagrams cannot provide conviction (Inglis & Mejia-Ramos, 2009b) or that diagrams are not useful in the construction of a proof (Raman, 2003).

Weber (2010) found that advanced mathematics majors were not consistent in their evaluations of visual arguments. Twenty-eight mathematics majors were presented with a diagrammatic argument that purported to establish the claim that $\int_0^\infty \frac{\sin x}{x} dx > 0$. The argument first presented a graph of the function $f(x) = \frac{\sin x}{x}$; the argument proceeded by observing the first positive area was larger than the first negative area, the second positive area was larger than the second negative area, and so on, which implied the improper integral would have positive area. Fourteen of the 28 participants evaluated this argument as not meeting the standards of proof, with nine citing the presence of a graph as a reason for their judgment. Weber (2010) also presented these students with a traditional area model to establish the claim that $(a+b)^2=a^2+2ab+b^2$ along with an explanation for why the diagram related

to the claim. No participant said this argument did not meet the standards of proof[3]

and 14 participants said it was their favorite argument of the ten arguments they

read in the study. One contribution of this current paper is to offer an explanation

for why the students in Weber's (2010) study evaluated these arguments

differently.

2. Study 1

2. 1. Rationale for study

The original intention of Study 1 was to identify the processes that students

used as they attempted to translate graphical arguments to (what they perceived to

be) deductive proofs. However, as the study progressed, we found the interesting

issue in our data was not *how* participants attempted to make this translation but *if

and when* they chose to do so. In many cases, participants would simply include the

graphical inference that they made without attempting to provide a non-graphical

justification for it.

The main point of presenting this study is to advance two ideas. First, we

noticed that participants made two different types of graphical inferences, which we

labeled *graphical perceptual inferences* and *graphical deductive inferences*. Second,

we found that participants usually would try to justify the graphical perceptual

inferences without reference to a graph while they would often acted as if they

believed the graphical deductive inferences were permissible within a proof.

2. 2. Methods

---

[3] Fifteen participants thought the argument was fully rigorous, nine thought the argument was a non-rigorous proof, and four participants said they were unsure if the argument met the standards of proof.

2. 2. 1. Corpus of data. Twelve mathematics majors agreed to participate in a study about their proof writing processes in exchange for a monetary fee. The participants consisted of students who had recently graduated with a degree in mathematics or who had completed their junior year. Each participant met individually with an interviewer for two 90-minute sessions. Participants were told that they would be asked to write proofs and to "think aloud" as they constructed the proofs. They were informed they would be given ten minutes to complete each proof and that they should write up their final proofs as if they were going to be graded in a mathematics exam.

Participants were asked to complete seven proving tasks from calculus and seven from linear algebra. The data presented in this paper is from the seven calculus tasks, which are given in the Appendix. The calculus tasks were chosen so that it would be feasible to approach them "semantically" (in the sense of Weber & Alcock, 2004)-- that is, participants could make progress on the tasks by considering informal representations of concepts such as graphs, diagrams, and prototypical examples. This choice of tasks was informed by a grant advisory board (which included a mathematician and two mathematics educators) and consultations with mathematics faculty members at the university where this study took place.

In the first interview session, each participant began by completing a practice problem to become accustomed to the interview format. The participant was then given one of the study tasks. The participant was permitted to work on a proof until he or she wrote a proof that he or she was satisfied with, the participant felt that he or she could not make any more progress, or ten minutes elapsed (whichever event

occurred first). The interviewer then asked the participant questions about their proving process, including a summary of what the participant did, what the main ideas of the proof were, and how the main ideas of the proof were generated. This process was repeated six more times for other tasks. In the second interview session, the participant attempted the remaining seven tasks using the same protocol as above.

At any point in the study, two resources were available to the participants. First, if participants could not recall the definition of a relevant concept, they could ask the interviewer for the definition. At that point, the interviewer would hand them a sheet of paper with the definition of the concept and an example of the concept. Second, participants had access to a computer with a graphing calculator application that enabled them to make basic calculations and view the graph of any function that they wished.

2. 2. 2. Identifying informal explanations. The purpose of this study was to investigate how students attempted to translate an informal explanation into a deductive proof. To identify informal explanations, we analyzed each of the participants' protocols on their proving tasks as follows: First, we flagged for every instance in which they represented a concept. Following Weber and Alcock (2009), the representation was coded as a *syntactic representation* if it consisted of the definition of a concept or a formula and a *semantic representation* if it was an informal representation of a concept such as a graph or a diagram. We then flagged each inference made by the participant (i.e., where a participant claimed a particular assertion was true, or likely to be true, that was not contained in the problem statement and had not been

stated previously by the participant). If the inference was drawn from a syntactic representation, we coded this as a *syntactic inference*[4]. If the inference had been drawn from a semantic representation, we coded this as a *semantic inference*. To illustrate, if a participant graphed a specific function (a semantic representation) and observed from the graph that the function appeared to be increasing, we coded this as a semantic inference. If the participant represented the function by its formula (a syntactic representation) and deduced that the function was increasing because its derivative was strictly positive, we coded this as a syntactic inference. An *informal explanation* was coded as occurring if there was a chain of inferences concluding with the statement to be proven that contained at least one semantic inference.

By following this procedure, we identified 16 informal explanations that contained a total of 38 semantic inferences. Each of the 38 semantic inferences involved drawing an inference from a graph.

2. 2. 3. Analyzing informal explanations. We analyzed each informal explanation as well as each proof that the participant submitted using the methodology of Pedemonte (2007). Pedemonte used a simplified Toulmin (2003) scheme where each inference was categorized in terms of a *claim* (the new statement being asserted), *data* (the facts that form the basis for the claim that are accepted as true), and a *warrant* (a general principle for why the data necessitates the claim). In some

---

[4] Note that we are not claiming that the student making a syntactic inference necessarily lacked the ability to interpret the statement in graphical terms, only that we saw no evidence that they did so in this case. A student producing a syntactic inference may have been following a procedure rather than focusing on logic, as discussed in Weber (2004), but as our primary concern is the student's semantic inferences, this distinction is not important for the purposes of this paper.

cases, the warrant was not explicitly stated. In these cases, the research team would infer the warrant if we perceived an obvious connection between the data and the claim. We classified the warrants for the semantic inferences into two types, which we termed graphical perceptual warrants and graphical deductive warrant. We define and illustrate both types of warrants shortly.

While our original intention was to see how participants translated the graphical inferences into deductive inferences for the purposes of proving, we found that the more interesting issue was whether the participants expressed the need to do so. We coded a participant as *expressing the need to translate a graphical inference* if one of the three conditions occurred: (i) the warrant used in the proof to justify the graphical inference differed from the warrant in their informal argument, (ii) the participant attempted to construct a sub-proof using the data from the inference as an assumption and the claim as the conclusion but was unable to do so, or (iii) the participant submitted the proof but expressed doubt that the proof was correct because of the presence of this graphical inference. We coded a participant as *not expressing the need to translate a graphical inference* if both of the following conditions were met: (i) the warrant used to justify the graphical inference was the same warrant used to justify this step in the proof that they submitted and (ii) the participant expressed no indication that he or she had reason to doubt if this step was appropriate.

2. 3. Graphical perceptual and graphical deductive inferences

Definitions in advanced mathematics are usually expressed formally, using a combination of natural language and logical syntax and lending themselves to

syntactic manipulation by means of logic and algebra. While pictures are often important for motivating or comprehending the definition of a concept, the definition itself ordinarily avoids direct reference to a picture. Nonetheless, it is frequently the case that formal definitions have *graphical interpretations*. For instance, a strictly positive function can be interpreted as a function whose graph is strictly above the *x*-axis, an increasing function is a graph that is continually moving upward as the graph is read from left to right, and an even function can be interpreted as a graph that is symmetric across the *y*-axis.

We found that the participants in our study frequently took a graphical interpretation of a definition as a starting point in their reasoning. The participants used these graphical interpretations in two ways. For *graphical perceptual inferences*, participants would examine the graph of a specific function, noticed from the graph that the function satisfied the graphical interpretation of some property, and then inferred that the function had that property. Graphical perceptual inferences were of the form "A function *f* has property *P*" with the warrant that "the graph of *f* visually satisfies a graphical interpretation of *P*".

As an example of a graphical perceptual inference, consider P10's work on the task "prove that the only real solution to the equation $x^3+5x=3x^2+\sin(x)$ is $x=0$". Like many students, P10 reformulated the problem by defining the function *f(x)*=x^3+5*x*-3x^2-sin(*x*) and trying to show that *f(x)* only had a root at *x*=0. P10 sketched the graph of *f'(x)*=3x^2+5-6*x*-cos(*x*) and from the appearance of the graph, concluded "Alright, so this [the graph of *f'(x)*] doesn't hit zero at all". Note here that P10 is using a graphical interpretation of a strictly positive function-- that the

16

function is above the *x*-axis and never intersects it-- and observes that *f'(x)* satisfies this condition. This and only this is P10's grounds for claiming that *f'(x)* is a strictly positive function.

For *graphical deductive inferences*, participants would say that a function that satisfied the graphical interpretation of some properties would necessarily satisfy a graphical interpretation of another property. The justifications would depend upon what they perceived to be a common sense understanding of the nature of two-dimensional space. We chose the name graphical deductive inferences because to the individuals, these inferences are deductive in the sense that they view them as logically necessary consequences of how properties are conceptualized. They differ from conventional deductive inferences in traditional proofs in that the conceptualization of the concepts is based on graphical considerations rather than formal definitions and that the deduction itself involves spatial reasoning. Graphical deductive inferences are of the form, "Since a function *f* has property *P*, the function *f* must necessarily have property *Q*" with the warrant that "One cannot construct a graph of a function satisfying the graphical interpretation of *P* while not also satisfying a graphical interpretation of *Q* because this would violate principles of two dimensional space". Note that as opposed to graphical perceptual inferences, the warrants for graphical deductive inferences avoid direct reference to the specific graph of *f*.

For an example of a graphical deductive inference, after inferring that *f'(x)* was strictly positive (and hence *f(x)* was increasing) and then verifying that *f*(0)=0, P10 reasoned, "that means it [*f(x)*] will never cross over the x-axis again and it'll have to

decrease at some point, so it'll actually have to actually...the derivative will have to go under the x-axis for there to be another root". Our interpretation of this utterance is that P10 is describing a relationship between the graphical interpretation of *f(x)* being increasing (as you read from left to right, the height of the graph of *f(x)* will be increasing) and the graphical interpretation of *f(x)* having a root after *x*=0 (the graph of *f(x)* will have to "cross over the *x*-axis again")[5]. For the graph to have another root, the graph will have to decrease (or go down) at some point.

As a general heuristic, we can distinguish between graphical perceptual and graphical deductive inferences as follows. If there was an inaccuracy in the graph that was drawn, this could render the graphical perception to be invalid. With graphical deductive reasoning, the accuracy of the graph is less important as the graph can be viewed as a prototype of a graph satisfying certain hypotheses (e.g., being increasing) and hence the validity of such an inference is not dependent upon the graph.

Graphical perceptual inferences are related to Harel and Sowder's (1998) perceptual proof schemes, where an individual holding a perceptual proof scheme would believe a mathematical assertion was true based solely on the appearance of a visual representation of a mathematical object (rather than results anticipated by some transformation of this representation). The difference between the two is that we do not claim that participants making a graphical perceptual inference necessarily have absolute conviction that their inference is true or believe such

---

[5] Note P10's explanation is only sufficient to show that *f(x)* has no *positive* roots, not that it has no roots other than *x*=0. However, our intention is not to judge the quality or correctness of P10's reasoning.

inferences are appropriate in a proof. One may reinterpret our question of whether participants expressed a need to provide a different warrant for their graphical perceptual inferences as asking whether these participants are exhibiting perceptual proofs schemes.

Graphical deductive inferences are related to, although not identical with, Simon's (1996) transformational reasoning, which anticipates the results of performing a transformation on a set of objects (see also Harel and Sowder's (1998) transformational proof scheme). In the example in this section, one could argue that P10's graphical deductive inference involved imagining transforming the graph of a continuous function so that it had two *x*-intercepts and realizing this entailed the function must be decreasing at some point.

2. 4. Students' perceptions on the appropriateness of graphical perceptual inferences and graphical deductive inferences in a proof.

Table 1
Participants' Expressing a Need to Translate Their Graphical Inferences by Type of Inference

	Perceptual inferences	**Deductive inferences**
Expressed need to translate	13 (68%)	5 (26%)
Did not express need to translate	6 (32%)	14 (74%)

In Table 1, we present the number of times participants did or did not express a need to translate their graphical inferences as a function of the type of graphical inference that they drew. Table 1 indicates that when participants made a graphical perceptual inference in their informal explanation, in most cases (68%), participants expressed a need to translate this inference when writing up their proof. However, when participants made a graphical deductive inference,

participants did not express this need the majority of the time (74%). We illustrate

this with two examples.

2. 4. 1. A first example of providing justification for graphical perceptual inferences.

For the first example, P3 was attempting the following task:

> Suppose $f''(x)>0$ for all real numbers x. Suppose a and b are real numbers with $a < b$. Define $g(x)$ as the line through the points $(a, f(a))$ and $(b, f(b))$. Prove that for all x in $[a, b]$, $f(x) \leq g(x)$.

P3 attempted to prove this statement by contradiction. To do so, P3 first drew a

graph where there was an x between a and b such that $f(x)>g(x)$. This graph is

presented in Figure 1. (The entirety of P3's written work and the proof that she

submitted is presented in the Appendix). Note that in Figure 1, the variables a and b

are represented as an interval on the x axis, $g(x)$ is the line connecting $(a, f(a))$ and

$(b, f(b))$, and $f(x)$ lies above $g(x)$ on the interval (a, b). P3 then presented the

following argument based on the picture.

[1] So my pretend point is up here [P3 draws a point whose x-coordinate was in the interior of the interval (a,b) with the point lying above the graph of $g(x)$. Note that the existence of such a point is implied by assuming the statement to be proven is false].
[2] And I think the way I'm going to do this is by saying that if my function looks like this [pause] then it has.....so I can use the Intermediate Value Theorem I think, to do this...
[3] So this slope will be greater than 0, [draws a line between $(a, f(a))$ and the point plotted in step [1]] and this one [draws a line between the point plotted in [1] and $(b, f(b))$] will be less than 0.
[4] So there are values...in fact for all of the values here, $f'>0$ [draws a parabolic curve and writes $f'>0$] and $f'<0$ [writes $f'<0$ to the right of $(x, f(x))$]. Which means that somewhere on this interval that f'' needs to be 0 […] I need to contradict that $f''>0$ so if I can show that if this is the case [points at the graph] than f'' has to equal 0 somewhere […] in this interval […] and then that contradicts my hypothesis."
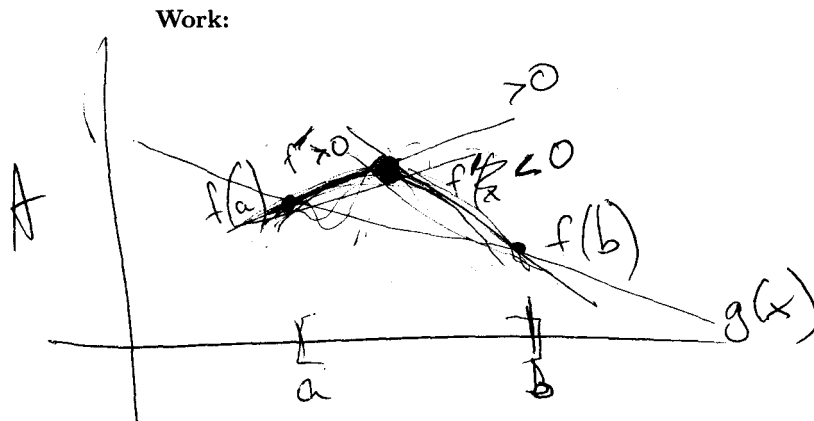
Figure 1. P3's graph for her informal argument.

In [2], P3 makes two perceptual inferences-- namely that *f* is increasing between *a* and *x* and *f* is decreasing between *x* and *b*. (P3's decision to label the point where *f(x)>g(x)* as *x* makes it difficult for us to smoothly describe her work). As P3 observed later, these graphical perceptual inferences were not valid as her hypothetical point *(x, f(x))* need not appear with a higher altitude than the point (*a, f(a))*. With these assumptions, P3 incorrectly uses the Intermediate Value Theorem[6] to conclude that there must be a value where *f''(x)* = 0. When P3 attempted to write up the proof, she began by trying to provide algebraic backing for these two perceptual inferences. She wrote that *f(x)-f(a)>0* and *x-a>0* and used this to infer that $\frac{f(x)-f(a)}{x-a} > 0$. Upon making this inference, P3 then noticed a flaw in her reasoning, stating:

> [6] So then *f(x)* minus *f(a)*...no that's not true. I think I've drawn a picture in a way that tricks me actually. Because it's not true that *f(x)*...I wanted to write that *f(x)-f(a)* is greater than 0, but actually the picture could've been something like this [draws graph in Figure 2, a generic concave down arc such that the first intersection of the arc and the line is higher up, or has a greater *y*-value, than the

---

[6] Actually, the Intermediate Value Theorem would imply that there is a *c* on the interval [*a,b*] where *f'(c)*=0, which would not contradict any of the hypotheses in the proposition to be proved. However, it would be sufficient to show that *f'(x)* could not have a positive value and then a negative value, as this would imply *f'(x)* was not strictly increasing, as asserted by the assumption that *f''(x)>0*.

second intersection] and you know, *f(x)* would be somewhere out here. Sorry I need to think about this for a second because I need to set it up correctly and it should just fall out but I actually have to set it up correctly.

P3's final proof, presented in Figure 3, is based on a faulty use of "without loss of generality" to retain her original argument. What is clear from this episode is that P3 recognized that her two perceptual inferences required algebraic backing and this backing is provided in the proof that she submitted.

Figure 2. P3's second graph for how the conclusion of the proposition can be contradicted.

Figure 3. P3's submitted proof of the proposition.

2. 4. 2. A second example where P5 expresses the need to provide algebraic support for graphical perceptual inferences, but not graphical deductive inferences. For the second example, we describe P5's attempt to prove that the only real solution to the equation $x^3+5x=3x^2+\sin(x)$ is $x=0$. P5 initially attacked this problem by expressing sin x as its Maclaurin series, but quickly abandoned this approach when he saw its algebraic complexity. P5 then transformed the problem into showing that $f(x)=x^3+5x-3x^2-\sin(x)$ had a root at $x=0$. He used the graphing software to graph $f(x)$ and said:

> [1] Oh God there is a much easier way to do that. So this function becomes the function $x^3-3x^2+5x-\sin(x)$ ...we know that's equal to zero.
> [2] But if this is our function, and if we take the derivative of this function...the derivative is...$3x^2-6x+5$...derivative of sine is cosine, so minus $\cos(x)$.
> [3] So the most $\cos(x)$ can take away from this is...the absolute value of $\cos(x)$ is at most 1, so this is less than or equal to...I want to take absolute value. So we want to prove that this is greater than 0 for all x, and…
> [4] if we can do that then we can prove what the graph shows us, that the function is increasing...and if the function is increasing everywhere then...I think that's what the graph shows us.
> [5] If we can prove that the function is increasing everywhere, then we've done enough to show that 0 is the only real solution, because if the function is increasing everywhere and it crosses at zero then it can't go back.

After making these comments, he manipulated $f'(x)$ and (erroneously) believed it was sufficient to prove that $3x^2+5>6x$ so he graphed both sides of the inequality and said:

> [6] We want $3x^2+5$ greater than $6x$ [P5 graphs $3x^2+5$ and $6x$ on the same screen]. Hey look it's everywhere, which is exactly what we needed it. Can I use that as part of the proof? Or should I 'prove it' prove it?

Here, the interviewer replied, "I mean however you would write it in an exam". P5

spent the remainder of his time attempting to prove that *f'(x)>0*, but was not

successful. For the proof, P5 presented the following proof:

> Let $y = x^3 - 3x^2 + 5x - \sin x$. Observe that y = 0 when x = 0, and that y'$(x) = 3x^2 - 6x + 5 - \cos(x)$, which is positive for all x, so the function is always increasing & cannot therefore cross the x-axis anywhere else, so 0 is the only solution.

In this protocol, P5 made two graphical perceptual inferences. The first was that *f(x)*

was increasing. In [4], P5 expresses a need to provide a non-graphical justification

for this, saying, "if we can prove what the graph shows us [that *f(x)* is increasing]".

The proof itself does contain such a justification, that *f'(x)* is strictly positive. The

claim that *f'(x)* is strictly positive was also a graphical perceptual inference, but

again P5 expresses doubt about its appropriateness in [6], where he says, "Can I use

that as part of the proof? Or should I 'prove it' prove it?" When told to proceed as if

he were answering an exam question, he continued to search for an algebraic

justification. He was unable to find one and the statement appears unjustified in the

proof, but P5 clearly demonstrated a need to do so.

P5 also made a graphical deductive inference in [5], noting that *f(x)* cannot

have two roots because "if it is increasing and it crosses at zero then it can't go

back". However, he does not express a need to justify this in a non-graphical

manner, saying in [5] that if he can establish that *f(x)* is increasing, "then we've done

enough to show that 0 is the only solution". No further justification is presented in

the proof.

2. 5. Summary

From this study, we introduced the constructs of graphical perceptual inference and graphical deductive inference. By analyzing students' actions after making each inference, we have the hypotheses that the participants perceive the validity of each inference differently. In particular, participants demonstrated a strong propensity to believe that graphical perceptual inferences, but not graphical deductive inferences, required non-graphical justifications in a proof.

There are three limitations to this study that prevent us from making broad conclusions. First, the results from this study were based on only 16 informal explanations. A larger sample is needed before attempting to generalize the findings of this study to the larger population of advanced mathematics majors. Second, in other publications, we cautioned researchers not to infer what proof schemes that students possess based on the justifications that they submit for credit (e.g., Weber, 2010; Weber, Inglis, & Mejia-Ramos, 2014), in part because students may realize that they are handing in a flawed product (see also Stylianides & Stylianides, 2009). Perhaps students did not attempt to justify graphical deductive inferences in a non-graphical manner for reasons that they did not state orally (e.g., time constraints or they perceived such a justification to be too difficult to construct). Third, while students were asked to write up proofs as if they were completing an examination, the course in which this exam was given was not specified to students. It may have been the case that students thought such an explanation was appropriate on a first-semester calculus exam, but not a real analysis exam. We conducted a confirmatory quantitative study that addresses each of these three concerns.

3. Study 2

3. 1. Rationale for Study 2

The goal for Study 2 is to explicitly test the main hypotheses generated from Study 1-- that mathematics majors believe graphical deductive inferences are permissible in a proof but graphical perceptual inferences are not-- while addressing the limitations from Study 1. Participants in this study viewed three proofs, each of which contained a graphical perceptual and a graphical deductive inference. They were asked to judge whether each inference was appropriate for a proof and if their professor would take off points if that inference appeared in a proof.

This study addresses the limitations in Study 1 in the following ways: First, 90 mathematics majors participated in the study, limiting the possibility that the findings from Study 1 were an artifact of having a small sample. Second, participants were explicitly asked which graphical inferences would be permissible in a proof, a more direct and transparently valid way to address the research questions in this paper. Third, participants were told these proofs were given in a specific class (half were told real analysis and half were told introductory calculus), eliminating ambiguity about the context in which the proofs were couched.

3. 2. Methods

3. 2. 1. The use of an Internet study. Following the methodology employed by Inglis and Mejia-Ramos (2009b), we collected data through the Internet in order to maximize our sample size. Recent studies have examined the validity of Internet-based experiments by comparing this type of studies with their laboratory equivalents (e.g. Kranz & Dalal, 2000; Gosling et al., 2004). The notable degree of

congruence between the two methodologies suggests that, by following simple

guidelines, Internet data has comparable validity to more traditional data. We

adopted the measures described in Inglis and Mejia-Ramos (2009b) to ensure the

validity of our data.

3. 2. 2. Participants. We recruited mathematics majors to participate in this study as

follows. Twenty-four secretaries from top-ranked mathematics departments in the

United States[7] were contacted and asked to distribute an email to the mathematics

majors at their university. The email invited mathematics majors who had

completed a course in real analysis to participate in our study. Mathematics majors

who agreed to participate in our study could click on a hyperlink that directed them

to the website of the study. When they clicked on the page, the study began by

asking for demographic information. One question asked participants if they had

taken a course in real analysis. The data for students who answered no to this

question was not included in our study[8]. Through this process, we recruited 90

mathematics majors who claimed to have taken a course in real analysis and

completed our experiment.

3. 2. 3. Procedure. Upon participating in the experiment, participants were randomly

assigned to the real analysis or introductory calculus group. Participants in the real

analysis group received the following instructions:

> We will ask you to read three mathematical statements and proofs. After carefully
> reading, please respond to the questions that follow as if the proofs are items on
> an exam in a **real analysis** class in your university. Two questions will be asked
> following each proof. The first question will ask you whether you think a step in
> the proof is sufficiently justified for an exam in a **real analysis** class. The second

---

[7] As ranked by the USNews.com "Best Graduate Schools" list of "top mathematics programs".

[8] Only five of the 95 students who completed the study answered no to this question.

question will ask you whether you think your class' professor would take points off of the exam for the step.

We will first provide you with an annotated and answered sample item to clarify the questions we are asking. Each proof will also be separated into steps so the reasoning is easier to follow. (The phrase "real analysis" was given in bold font and underlined on the actual webpage).

The text for the calculus group was identical, except the "real analysis class" phrases were substituted with "first year calculus class".

Next, participants were shown a worked example to illustrate the ideas of the experiment. They were shown a sample proof of the claim that "$\frac{1}{x^4+x^2+2x+1} + 1$ was positive for all real-valued x". Step 1 of the proof claimed that "my roommate told me that $\frac{1}{x^4+x^2+2x+1}$ was positive". In the worked example, step 1 was evaluated as not being an adequate justification in a calculus/real analysis class and that the professor would take points off, because even though the claim in Step 1 is true, appealing to one's roommate is presumably regarded as an impermissible justification. Step 2 declared that since $\frac{1}{x^4+x^2+2x+1}$ was positive and 1 was positive, $\frac{1}{x^4+x^2+2x+1} + 1$ was positive. In the worked example, this step was considered acceptable, since if one assumed step 1 as correct, step 2 is only using the accepted fact that the sum of two positive numbers is positive. One reason for presenting this worked example is to make participants aware that one could accept a step in the proof as permissible, even if the step is logically building on a previous step that was problematic.

From here, participants saw three proofs in a randomized order, where each proof contained a graphical perceptual inference and a graphical deductive

inference (the proofs are presented in Appendix and discussed shortly). They were first shown the theorem statement and asked to read the proposition. They were then shown a proof of the proposition with the graphical perceptual inference shown in red and were asked the following two questions: "Do you think the argument in step *x* is an adequate justification for the claim that [claim made in step *x*] if the proof was written on an exam for a **real analysis** student". They were also asked, "do you think the professor would take points off for the justification of the step highlighted in red if the proof was written on an exam for a **real analysis** class?" (In each case, "real analysis" appeared in bold in the instructions. The calculus group had "first year calculus" printed in place of "real analysis". The variable "x" represented the step number where the graphical perceptual inference was highlighted in red). This was then repeated with the graphical deductive inference highlighted in red with the same two questions being asked.

3. 2. 4. Materials. There were three proofs used in this study, each of which contained a graphical perceptual and a graphical deductive inference. The complete tasks are given in the Appendix. We illustrate our task with Proof 2, which purports to establish that the derivative of e^{-x^2} is odd. (This proof is adapted from the task in Raman, 2003). The proof begins by presenting a graph of $f(x)=e^{-x^2}$. Step 1 in the proof is the graphical perceptual inference, stating "we can see from the graph that $f(x)$ is symmetric across the *y*-axis". Step 3 in the proof is a graphical deductive inference that builds upon Step 1, claiming, "Thus for any point *a*, the tangent line of *f* at *a* and the tangent line of *f* at -*a* will be mirror images of each other. Thus the slopes of these tangent lines will have the same magnitude but opposite signs".

In addition to the three proofs found in the Appendix, we added two other tasks, one involving an inference that we believed was clearly valid and acceptable and a second that we believed was clearly invalid and unacceptable. In Proof 3, we highlighted an inference that we thought was clearly justified in an adequate manner (an algebraic demonstration that the solutions to $12x^2 - 4x^3 = 0$ are $x=0$ and $x=3$). When participants evaluated Proof 3, in addition to the graphical perceptual inference and the graphical deductive inference, they were also asked if this transparently good inference was appropriate and whether the professor would take points off for this step in the proof. We presumed that if participants were taking our tasks seriously, the answer the participant would judge this as appropriate and would claim a professor would not take points off for this.

For a transparently bad inference, we created an alternative proof to Proof 1 that we believed was clearly inadequate (claiming $x=0$ was the only solution to an equation by verifying that $x=0$ was a solution). Prior to reading Proof 1, the participants read a proof consisting entirely of the transparently bad inference and were asked if this inference was appropriate and whether the professor would take points off for this step in the proof. We presumed that if participants were taking our tasks seriously, the participant should claim this inference was not appropriate and a professor would take points off for it. We included these additional inferences to be sure that participants were not saying that every inference was acceptable or that no inference was acceptable.

3. 2. 5. Planned comparisons. We planned to test two specific hypotheses in this study. *Hypothesis 1* is that participants will find graphical deductive inferences as

more permissible in a proof than a graphical perceptual inference. *Hypothesis 2* is

that participants in the calculus condition will be more likely to find an inference as

acceptable than participants in the real analysis condition.

3. 3. Results

3. 3. 1. What inferences are acceptable within a proof? In Table 2, we present the

percentage of participants who thought each of the eight inferences in the study was

acceptable for a proof on an exam. In Table 3, we aggregate participants' judgments

across the three graphical perceptual inferences and the three graphical deductive

inferences. We first note that the transparently good and bad justifications had their

desired effects. From Table 2, we see that 98% of the participants thought the

transparently good justification was acceptable in a proof and less than 10% of the

participants judged the transparently bad inference to be acceptable.

Table 2
Participants' Judgments of Acceptability of an Inference by Inference

			Perceptual inferences			*Deductive inferences*		
Condition	Good Inf.	Bad Inf.	Proof1	Proof2	Proof3		Proof1	Proof2	Proof3
Real analysis (*N*=40)	98%	10%	18%	23%	25%		68%	60%	80%
Calculus (*N*=50)	98%	6%	30%	30%	32%		76%	66%	76%

Table 3
Participants' Aggregate Judgments of the Acceptability of Graphical Perceptual Inferences and Graphical Deductive Inferences

Experimental condition	Graphical perceptual inferences	Graphical deductive inferences
Real analysis (*N*=40)	22%	69%
Calculus (*N*=50)	31%	73%

Related-samples Wilcoxon signed-rank tests reveal that for both the Real Analysis

group and the Calculus group, participants found more graphical deductive

inferences than graphical perceptual inferences within a proof to be acceptable
($p<.001$ for both comparisons), confirming Hypothesis 1. However, we found no
statistically reliable difference between the Real Analysis and Calculus participants
regarding their judgments of either the graphical perceptual inferences (Mann-
Whitney, $U=1136.5$, $p=.217$) or the graphical deductive inferences (Mann-Whitney,
$U=1024$, $p=.827$). Hence, this data does not support Hypothesis 2, that participants
would find more inferences acceptable within a calculus context than a real analysis
context.

In Table 4, we present the percentage of participants who were completely
consistent in their evaluations of both the graphical perceptual and graphical
deductive inferences on the three proofs in this study (i.e., participants who judged
all three of a type of inference as acceptable or all three as unacceptable). As Table 4
illustrates, the majority of participants in both the calculus and real analysis
conditions thought no perceptual inferences were acceptable and the majority
judged all three graphical deductive inferences to be acceptable.

Table 4
Consistency of Participants' Judgments Across the Graphical Perceptual Inferences and the Graphical Deductive Inferences

Experimental condition	Judged all three perceptual inf. as acceptable	Judged no perceptual inf. as acceptable	Judged all three deductive inf. as acceptable	Judged no deductive inf. as acceptable
Real analysis	5%	63%	55%	8%
Calculus	10%	52%	60%	8%

3. 3. 2. For what inferences would a professor take points off?

In Table 5, we present the percentage of participants who thought a
professor would take off points for each of the eight inferences in this study. In

Table 6, we aggregate participants' judgments across the three graphical perceptual

inferences and the three graphical deductive inferences. Tables 5 and 6 show trends

similar to Tables 2 and 3. Most participants believed a professor would not take off

points for the transparently good inference but would do so for the transparently

bad inference. Related-samples Wilcoxon signed rank tests demonstrated that both

the Real Analysis participants and the Calculus participants judged professors more

likely to take off points for the graphical perceptual inferences than the graphical

deductive inferences ($p<.001$ in each case), thereby confirming Hypothesis 1.

Table 5
Participants' Judgments of Whether Professor Would Take Off Points by Inference

Condition			**Perceptual inferences**			**Deductive inferences**		
	Good Inf.	Bad Inf.	Proof1	Proof2	Proof3	Proof1	Proof2	Proof3
Real analysis (*N*=40)	3%	93%	83%	73%	78%	40%	48%	33%
Calculus (*N*=50)	4%	94%	66%	54%	60%	20%	36%	20%

Table 6
Participants' Aggregate Judgments of Graphical Perceptual Proofs and Graphical Deductive Proofs

Experimental condition	Graphical perceptual inferences	Graphical deductive inferences
Real analysis (*N*=40)	78%	40%
Calculus (*N*=50)	60%	25%

As opposed to their judgments on whether an inference was appropriate, the Real

Analysis participants were more likely than the Calculus participants to believe the

professor would take points off, both for the graphical perceptual inferences (Mann-

Whitney, *U*=760.5, *p*=.034) and the graphical deductive inferences (Mann-Whitney,

U=730, *p*=.015), confirming Hypothesis 2.

*3. 3. 3. Accounting for the differences between validity and the professor taking points

off.* For Hypothesis 2, we predicted that (a) participants were more likely to judge

graphical perceptual and graphical deductive as valid in a calculus setting than a real analysis setting and (b) participants were more likely to believe a professor would take points off in a real analysis setting than a calculus setting. Our data confirmed (b) but not (a). We did not anticipate this result. We performed the following post-hoc analysis in an attempt to explain this discrepancy.

To see if there was an interaction between the experimental condition (Real Analysis vs. Calculus) and the type of evaluation of being made (appropriate vs. whether a professor would take points off)[9], we tabulated the cases where there was an inconsistency between a participant's judgment on a graphical perceptual or graphical deductive inference. That is, we counted the cases where (i) the participant indicated an inference would be appropriate for a proof but also that the professor would take points off and (ii) the participant indicated that an inference would be unacceptable for a proof but the professor would not take points off.

We found that situation (i) occurred 12 times in the Calculus condition (.24 times per participant) and 17 times in the Real Analysis condition (.43 times per participant). A post-hoc Mann-Whitney test did not find a significant difference in the occurrences of situation (i) between the two experimental conditions (U=867.5, p=.123).

Situation (ii) occurred 28 times in the Calculus condition (.54 per participant) and seven times in the Real Analysis condition (.18 per participant). A post-hoc Mann-Whitney test found a significant difference between the two groups in this respect (U=1222.5, p=.016), suggesting participants in the Calculus condition

---

[9] Of course, obtaining statistical significance in one comparison but not another comparison does not necessarily imply that the size of the effects in the two comparisons differed significantly.

were more likely to believe calculus professors would be lenient in their grading of inferences that they felt were invalid. As these tests were not planned comparisons, we treat this account as speculative and recommend verifying this trend in a future study.

4. Discussion

4. 1. Summary of main results

In this paper, we distinguished between two types of graphical inferences: graphical perceptual inferences and graphical deductive inferences. We investigated advanced mathematics majors' perceptions of the appropriateness of both types of inference within a proof after having taken a course in real analysis. Our main findings are that: (i) Most advanced mathematics majors indicated that they did not believe graphical perceptual inferences were appropriate in a calculus proof. (ii) Most advanced mathematics majors indicated that they believed graphical deductive inferences were appropriate in a calculus proof. (iii) Whether a proof was given in a calculus or real analysis context did not significantly alter participants' judgments about the appropriateness of an inference in a proof, but it did influence their judgments of whether a professor would take off points for including such an inference. Our tentative account of this finding is that participants believed that professors in a calculus course were less likely to penalize an invalid inference. (iv) As Study 1 illustrates, these perceptions affected the way that advanced mathematics majors attempt to translate informal arguments into proofs. The participants in Study 1 generally expressed a need to justify a graphical perceptual

35

inference via conventional deduction but believe a graphical deductive inference can be written in a proof without translation.

4. 2. Caveats and limitations

There are three important limitations of our study. First, although Study 2 examined mathematics majors' perceptions of the appropriateness of graphical arguments with a reasonably large number of participants, the number of tasks used was relatively small (three graphical perceptual inferences and three graphical deductive inferences) and all in the same context (elementary calculus). It is possible that using a wider range of tasks, including tasks in a domain other than calculus, may have elicited different responses from these mathematics majors. Also, the consistency that many participants demonstrated with their evaluation of the three graphical perceptual inferences and three graphical deductive inferences may have been an artifact of using only three tasks.

Second, as we indicated in our opening section, we are unsure of how *mathematicians* would judge the appropriateness of the inferences that we presented to students. It is plausible that mathematicians might have considered the graphical deductive inferences to be acceptable in a proof, perhaps because they were permitted by the existence of a structure theorem (Tall, 2013), or there might not be a consensus amongst mathematicians on the appropriateness of some inferences. For this reason, we deliberately refrained from making normative judgments on participants' evaluations. We believe more research on mathematicians' practice is needed to address these issues.

Third, it is plausible that advanced mathematics majors from top ranked universities who agreed to participate for free on a survey about real analysis had a better understanding of real analysis than a typical student. This is a common bias in empirical studies that recruit mathematics majors from advanced mathematics courses; students who perform poorly in these courses likely exhibit a greater reluctance to participate in such studies. Hence, the findings may not necessarily generalize to all mathematics majors who completed real analysis.

4. 3. Relationship to the mathematics education literature

Weber (2010) found that some mathematics majors were inconsistent in their evaluation of visual arguments. The participants in his study all believed that a justification using an area model of multiplication to show that $(a+b)^2 = a^2+2ab+b^2$ constituted a proof, but many did not think a graphical argument showing that $\int_0^\infty \frac{\sin x}{x} dx > 0$ was a proof. The results reported in this paper both corroborate and explain these findings. The first argument in Weber's (2010) study used a graphical deductive inference while the second employed a graphical perceptual inference. In this study, we find that many advanced mathematics majors are accepting of the former but not the latter in a proof.

Harel and Sowder (1998) noted that some mathematics majors held perceptual proof schemes, meaning that they would convince themselves and persuade others by the appearance of a graph or diagram. Our data support Harel and Sowder's claim in that these data verify that some mathematics majors were willing to accept some of the graphical perceptual inferences as valid. However, the data also suggest such perceptions are uncommon with advanced mathematics

majors, at least with respect to the persuasive aspect of this proof scheme. In Study 1, participants usually expressed a need to justify their graphical perceptual inferences. In Study 2, the majority of the participants claimed that all three graphical perceptual inferences would not be appropriate in a proof.

Raman (2003) conveyed concern that some mathematics students would be reluctant to base their proofs off of informal personal arguments. She exemplified this by showing a graphical argument that employed a graphical deductive inference and claiming that students would not use such a graphical argument when writing a proof. Raman conjectured that this is due to calculus students (not necessarily advanced mathematics majors) having an undesirable epistemological belief that there is no connection between the formal (non-graphical) proofs that one produces and the informal (possibly graphical) arguments that one uses to understand why something is true. Our data suggest that this does not appear to be a significant concern with most advanced mathematics majors. Not only did the participants in Study 1 and Study 2 believe graphical arguments can form the basis for a formal proof, they also believe that graphical deductive inferences are *appropriate* for a proof and require no translation. The majority of participants in Study 2 judged all three graphical deductive inferences to be appropriate in a proof and 92% of participants found at least one of the three inferences of this type to be appropriate.

4. 4. Suggestions for future research

We suggest two avenues for future research. First, it is important to investigate mathematicians' viewpoints on the validity of graphical inferences in a proof and their appropriateness in proof-oriented mathematics courses. This

research would be important for determining whether the viewpoints expressed by some students in this study were normatively correct and for setting instructional goals for the beliefs about proof that we want students to develop.

Second, it would be worthwhile to conduct qualitative studies on *why* students hold the beliefs that they do about graphical deductive inferences and graphical perceptual inferences. Our data suggest that a substantial number of advanced mathematics majors accept the former as valid but reject the latter as invalid. Are these students aware that they are treating these two types of inferences differently? Do they perceive graphical deductive inferences as being in a separate category than graphical perceptual ones? If the answer to the previous two questions is yes, what rationale do they provide for why the former are permissible while the latter are not?

Acknowledgments

References

Aberdein, A. (2009). Mathematics and argumentation. *Foundations of Science*, 14, 1-8.

Alcock, L. (2010). Interactions between teaching and research: Developing pedagogical content knowledge for real analysis. In R. Leikin & R. Zazkis (Eds.) *Learning through teaching mathematics*. Springer: Dordrecht.

Alcock, L. J. & Simpson, A. P. (2004). Convergence of sequences and series: Interactions between visual reasoning and the learner's beliefs about their own role. *Educational Studies in Mathematics, 57*(1),1-32.

Alcock, L., & Weber, K. (2005). Proof validation in real analysis: Inferring and checking warrants. *The Journal of Mathematical Behavior*, *24*(2), 125-134.

Alcock, L. & Weber, K. (2010). Referential and syntactic approaches to proving: Case studies from a transition-to-proof course. *Research in Collegiate Mathematics Education*, 7, 101-123.

Azzouni, J. (2013). The relationship of derivations in artificial languages to ordinary rigorous mathematical proof. *Philosophia Mathematica*, 21, 247-254.

Bleiler, S. K., Thompson, D. R., & Krajčevski, M. (2014). Providing written feedback on students' mathematical arguments: proof validations of prospective secondary mathematics teachers. *Journal of Mathematics Teacher Education*, 17, 105-127.

Brown, S. A. (2014). On skepticism and its role in the development of proof in the classroom. *Educational Studies in Mathematics*, *86*(3), 311-335.

Burton, L. (2004). *Mathematicians as Enquirers: Learning about Learning Mathematics*. Berlin: Springer.

Dreyfus, T. (1991). On the status of visual reasoning in mathematics and mathematics education. In F. Furinghetti (Ed.), *Proceedings of the 15th International Conference on the Psychology of Mathematics Education,* Vol. I (pp. 33 - 48). Assisi, Italy.

Feferman, S. (2012). And so on…: reasoning with infinite diagrams. *Synthese*, 186, 371-386.

Gibson, D. (1998). Students' use of diagrams to develop proofs in an introductory real analysis. *Research in Collegiate Mathematics Education,* 2, 284-307.

Gosling, S.D., Vazire, S., Srivastava, S., & John, O.P. (2004). Should we trust web-based studies? A comparative analysis of six studies about internet questionnaires. *American Psychologist*, *59*, 93-104.

Hadamard, J. (1945). *The Psychology of Invention in the Mathematical Field*. Princeton: Princeton University Press.

Harel, G., & Sowder, L. (1998). Students' proof schemes: Results from exploratory studies. In A. H. Schoenfeld, J. Kaput, & E. Dubinsky (Eds.), *Research in collegiate mathematics education III* (pp. 234–283). Providence, RI: American Mathematical Society.

Harel, G., & Sowder, L. (2007). Toward comprehensive perspectives on the learning and teaching of proof. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 805–842). Greenwich, CT: Information Age Publishing.

Hart, E. (1994). A conceptual analysis of the proof writing performance of expert and novice students in elementary group theory. In J. Kaput, and Dubinsky, E. (Ed.), *Research issues in mathematics learning: Preliminary analyses and results* (pp. 49-62). Washington: Mathematical Association of America.

Healy, L., & Hoyles, C. (2000). Proof conceptions in algebra. *Journal for Research in Mathematics Education, 31*(4), 396-428.

Herbst, P. (2004). Interactions with diagrams and the making of reasoned conjectures in geometry. *ZDM- The international journal of mathematics education*, 36, 129-139.

Hersh, R. (1993). Proving is convincing and explaining. *Educational Studies in Mathematics*, 24(4), 389-399.

Hodds, M., Alcock, L. ,& Inglis, M. (2014). Self-explanation training improves proof comprehension. *Journal for Research in Mathematics Education*, 45, 62-101.

Iannone, P. & Inglis, M. (2010). Self-efficacy and mathematical proof: Are undergraduates good at assessing their own proof production ability? In *Proceedings of the 13th Conference for Research in Undergraduate Mathematics Education*. Raleigh, North Carolina.

Inglis, M., & Alcock, L. (2012). Expert and novice approaches to reading mathematical proofs. *Journal for Research in Mathematics Education*, *43*(4), 358-390.

Inglis, M. & Mejia-Ramos, P. (2009a). On the persuasiveness of visual arguments in mathematics. *Foundations of Science*, 14, 97-110.

Inglis, M. & Mejia-Ramos, P. (2009b). The effect of authority on the persuasiveness of mathematical arguments. *Cognition and Instruction*, 27, 25-50.

Ko, Y. & Knuth, E. (2009). Undergraduate mathematics majors' writing performance producing proofs and counterexamples about continuous functions. *Journal of Mathematical Behavior*, 28, 68-77.

Knuth, E., Choppin, J., & Bieda, K. (2009). Middle school students' productions of mathematical justification. In M. Blanton, D. Stylianou, & E. Knuth (Eds.) *Teaching and Learning Proof Across the Grades: A K-16 Perspective*. Routledge: New York.

Kranz, J. and Dalal, R. (2000). Validity of web-based psychological research. In M. Birnbaum (Ed.) *Psychological experiments on the internet*. (pp. 35-60). Academic Press: San Diego, CA.

Kupla, Z. (2009). Main problems with diagrammatic reasoning. Part 1: The generalization problem. *Foundations of Science*, 14, 75-96.

Larkin, J. & Simon, H.A. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science*, 11, 65-99.

Martin, W. G., & Harel, G. (1989). Proof frames of preservice elementary teachers. *Journal for Research in Mathematics Education, 20*(1), 41-51.

Moore, R. C. (1994). Making the transition to formal proof. *Educational Studies in Mathematics, 27*(3), 249-266.

Nelsen, R.B. (1993). *Proofs without words*. MAA: Washington, DC.

Nelsen R. B. (2000). *Proofs Without Words II*. MAA: Washington, DC.

Pedemonte, B. (2007). How can the relationship between argumentation and proof be analysed? *Educational Studies in Mathematics*, 66, 23-41.

Pfeiffer, K. (2011). *Features and purposes of proofs in the view of novice students: Observations from proof validation and evaluation performances*. Unpublished doctoral dissertation. National University Ireland.

Piez, C. M., & Voxman, M. H. (1997). Multiple representations: Using different perspectives to form a clearer picture. *The Mathematics Teacher, 90*, 164-166.

Raman, M. (2003). Key ideas: What are they and how can they help us understand how people view proof? *Educational Studies in Mathematics, 52,* 319-325.

Rav, Y. (1999). Why do we prove theorems? Philosophia Mathematica, 7(3), 5-41.

Recio, A. M., & Godino, J. D. (2001). Institutional and personal meanings of mathematical proof. *Educational Studies in Mathematics, 48*, 83-99.

Samkoff, A., Lai, Y., & Weber, K. (2012). On the different ways that mathematicians use diagrams in proof construction. *Research in Mathematics Education*, *14*(1), 49-67.

Schoenfeld, A. H. (1985). *Mathematical Problem Solving*. Orlando, FL: Academic Press.

Segal, J. (2000). Learning about mathematical proof: Conviction and validity. *Journal of Mathematical Behavior*, 18, 191-210.

Selden, A. & Selden, J. (2003). Validations of proofs written as texts: Can undergraduates tell whether an argument proves a theorem? *Journal for Research in Mathematics Education*, 36(1), 4-36.

Selden, A., & Selden, J. (2013). Proof and Problem Solving at University Level. *Montana Mathematics Enthusiast*, *10*, 303-334.

Selden, J., & Selden, A. (1995). Unpacking the logic of mathematical statements. *Educational Studies in Mathematics*, *29*(2), 123-151.

Stylianides, A. J., & Stylianides, G. J. (2009). Proof constructions and evaluations. *Educational Studies in Mathematics*, *72*(2), 237-253.

Stylianou, D. A. (2002). On the interaction of visualization and analysis - The negotiation of a visual representation in problem solving. *Journal of Mathematical Behavior*, 21, 3, 303-317.

Tall, D. (2001). Natural and formal infinities. *Educational Studies in Mathematics*, *48*(2-3), 199-238.

Tall, D. (2013). *How humans learn to think mathematically: exploring the three worlds of mathematics*. Cambridge: Cambridge University Press.

Toulmin, S. E. (2003). *The uses of argument*. Cambridge: Cambridge University Press.

Weber, K. (2001). Student difficulty in constructing proofs: The need for strategic knowledge. *Educational Studies in Mathematics*, *48*(1), 101-119.

Weber, K. (2004). A framework for describing the processes that undergraduates use to construct proofs. In M.J. Hoines & A. B. Fuglestad (Eds.) *Proceedings of the 28th conference of the International group for the Psychology of Mathematics Education*, 4, 425-432. Bergen, Norway: PME.

Weber, K. (2005). Problem-solving, proving, and learning: The relationship between problem-solving processes and learning opportunities in the activity of proof construction. *The Journal of Mathematical Behavior*, *24*(3), 351-360.

Weber, K. (2006). Investigating and teaching the processes used to construct proofs. *Research in Collegiate Mathematics Education*, 6, 197-232.

Weber, K. (2010). Mathematics majors' perceptions of conviction, validity, and proof. *Mathematical Thinking and Learning*, *12*(4), 306-336.

Weber, K., & Alcock, L. (2004). Semantic and syntactic proof productions. *Educational Studies in Mathematics*, *56*(2-3), 209-234.

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

Weber, K. & Alcock, L. (2009). Proof in advanced mathematics classes: Semantic and syntactic reasoning in the representation system of proof. In D.A. Stylianou, M.L. Blanton & E. Knuth (Eds.), *Teaching and learning proof across the grades: A K-16 perspective* (pp. 323-338). New York: Routledge.

Weber, K., Inglis, M., & Mejía-Ramos, J.P. (2014). How mathematicians obtain conviction: Implications for mathematics instruction and research on epistemic cognition. *Educational Psychologist*, *49*, 36-58.

Zandieh, M., Roh, K. H., & Knapp, J. (2014). Conceptual blending: Student reasoning when proving "conditional implies conditional" statements. *The Journal of Mathematical Behavior*, *33*, 209-229.

Zazkis, D., Weber, K., & Mejia-Ramos, J.P. (2014). Activities that mathematics majors use to bridge the gap between informal arguments and proofs. In P. Liljedahl, C. Nicol, S. Oesterle, & D. Allan (Eds.) *Proceedings of the 38th Conference for the International Group of the Psychology of Mathematics Education*, 4, 417-424. Vancouver, Canada: PME.

Appendix- Calculus tasks used in Study 1

1. Suppose $f(0) = f'(0) = 1$. Suppose $f''(x) > 0$ for all positive *x*. Prove that $f(2) > 2$

2. Prove that the only real solution to the equation $x^3 + 5x = 3x^2 + \sin x$ is *x*=0.

3. Suppose $f(x)$ is a differentiable even function. Prove that $f'(x)$ is an odd function.

4. Prove that $a^2 + ab + b^2 \geq 0$ for all real numbers *a* and *b*.

5. Suppose $f''(x) > 0$ for all real numbers *x*. Suppose *a* and *b* are real numbers with $a < b$. Define $g(x)$ as the line through the points $(a, f(a))$ and $(b, f(b))$. Prove that for all $x \in [a,b]$, $f(x) \leq g(x)$.

6. Prove that $\int_{-a}^{a} \sin^3(x)\,dx = 0$ for any real number *a*.

7. Let *f* be differentiable on $[0,1]$, and suppose that $f(0) = 0$ and f' is increasing on $[0,1]$. Prove that $g(x) = \dfrac{f(x)}{x}$ is increasing on $(0,1)$.

Appendix- P3's scratch work and proof

Suppose $f''(x) > 0$ for all real numbers x. Suppose a and b are real numbers with $a < b$. Define $g(x)$ as the line through the points $(a, f(a))$ and $(b, f(b))$. Prove that for all $x \in [a, b]$, $f(x) \leq g(x)$.

Work:

$$f(a) = g(a)$$
$$f(b) = g(b)$$

$$f''(x) > 0$$
$$f'(x) \text{ is increasing}$$
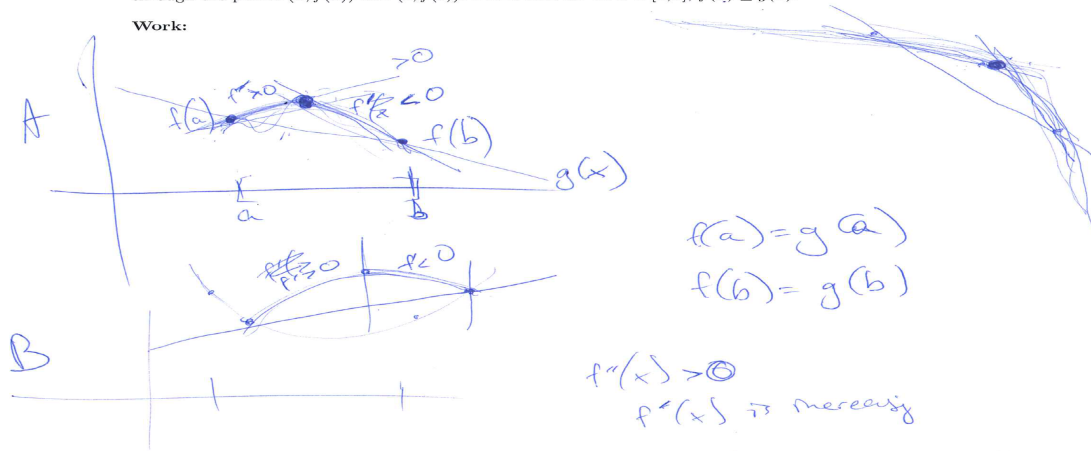
Final proof: Suppose to the contrary that $\exists x \in (a, b)$ such that $f(x) > g(x)$. wlog, suppose $g' \leq 0$

$f(x) - f(a) > 0 \qquad$ so $\quad \dfrac{f(x) - f(a)}{x - a} > 0$

$f(x) - f(b) > 0 \qquad$ so $\quad \dfrac{f(b) - f(x)}{b - x} \leq 0$

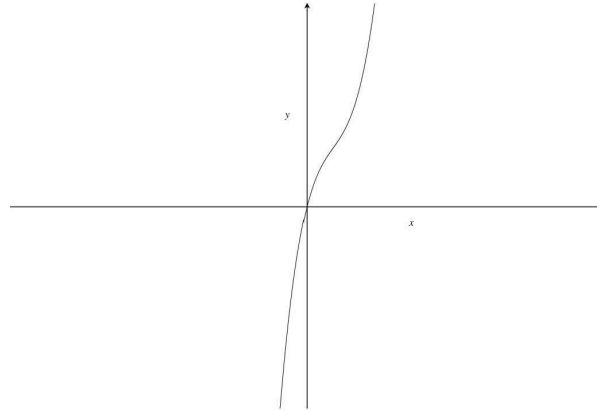so by IVT, I have a point where $f' > 0$ and a point where $f' \leq 0$, ~~then must be a point where f''~~ f'' would be less that 0 (contradiction, contradiction.

Appendix- Tasks used in Study

Proposition 1. *Prove the only real solution to the equation $x^3 + 5x = 3x^2 + sin(x)$ is $x = 0$.*

$$f(x) = x^3 - 3x^2 + 5x - sin(x)$$

Proof.

Step 1. We can rewrite the given equation as $x^3 - 3x^2 + 5x - sin(x) = 0$. Then it suffices to show that $f(x) = x^3 - 3x^2 + 5x - sin(x) = 0$ if and only if $x = 0$.

Step 2. First, we know that $x = 0$ is a solution since $f(0) = 0^3 - 3*0^2 + 5*0 - sin(0) = 0$. So we show that $x = 0$ is the only solution.

Step 3. Given the graph of $f(x)$ above, we see that $f(x)$ is strictly increasing.

Step 4. Because $f(x)$ is strictly increasing, $f(x)$ does not have a positive root, since it will have to come back down or remain flat which contradicts the fact that $f(x)$ is strictly increasing.

Step 5. Similarly, f(x) does not have a negative root. Thus $x = 0$ is the only solution. $\square$

STEP 3 is the graphical perceptual inference

STEP 4 is the graphical deductive inference

Proposition 1. *Prove that the derivative of* $f(x) = e^{-x^2}$ *is odd.*

$$f(x) = e^{-x^2}$$
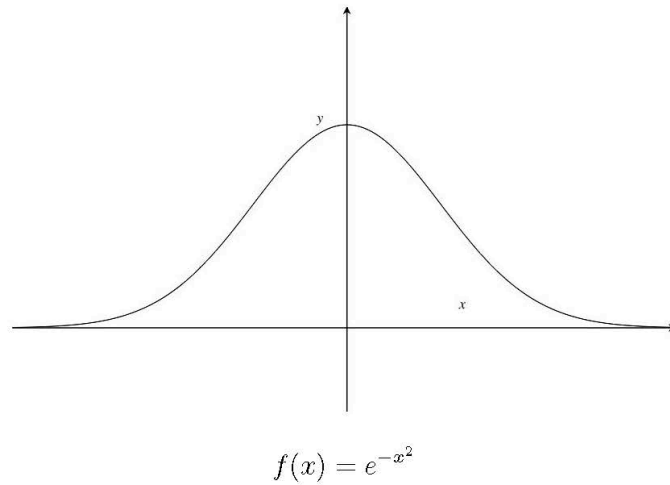
Proof.

Step 1. From the graph, we see that $f(x)$ is symmetric across the y-axis.

**Step 2.** For any a, the slope of the tangent line of f at a is $f'(a)$.

Step 3. By step 1, $f(x)$ is symmetric across the y-axis. Thus, for any point a, the tangent line of f at a and the tangent line of f at $-a$ will be mirror images of each other. Thus the slopes of these tangent lines will have the same magnitude but opposite signs.

Step 4. Thus, from step 2 and step 3, we get $f'(a) = -f'(-a)$ for any a and hence f' is odd, as desired. $\square$

STEP 1 is the graphical perceptual inference

STEP 3 is the graphical deductive inference

Proposition 1. *Prove that the equation $4x^3 - x^4 = 30$ has no real solutions.*
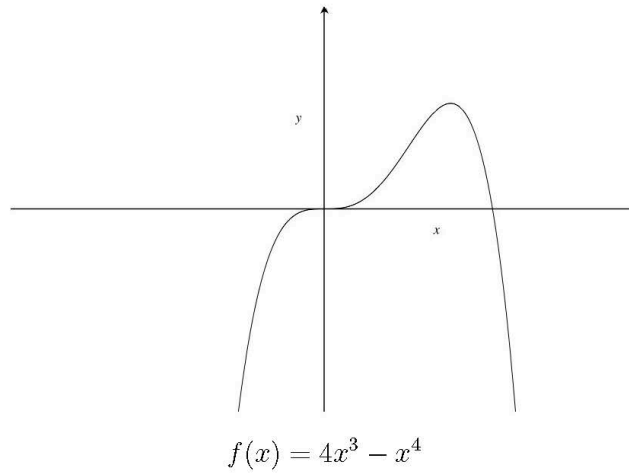
$$f(x) = 4x^3 - x^4$$

Proof.

Step 1. Let $f(x) = 4x^3 - x^4$. Since $f(x)$ is polynomial function, $f(x)$ is continuous.

Step 2. From the graph above, we see that $f(x)$ tends to $-\infty$ as x tends to $-\infty$.

Step 3. Similarly, $f(x)$ tends to $-\infty$ as x tends to ∞.

Step 4. By steps 2 and 3, $f(x)$ tends to $-\infty$ as x tends to $-\infty$ and x tends to ∞. Therefore, $f(x)$ must be bounded. If not, $f(x)$ must have a vertical asymptote, which implies $f(x)$ is not continuous.

Step 5. Taking the derivative of $f(x)$ we get $f'(x) = 12x^2 - 4x^3$.

Step 6. Setting $f'(x) = 0$ gives $12x^2 - 4x^3 = 0$. This can be simplified to $4x^2(3 - x) = 0$. Hence, a solution to this equation occurs when $4x^2 = 0$ or $3 - x = 0$ and so $x = 0$ or $x = 3$. Thus $f'(x) = 0$ when $x = 0$ and $x = 3$. Thus $f(x)$ has critical points $x = 0$ and $x = 3$.

Step 7. Since $f(x)$ is bounded above, one of these critical points must be a global maximum. Since $f(0) = 0$ and $f(3) = 27$, $f(0) < f(3)$, and thus $f(3)$ is a global maximum. So $f(x) < f(3) < 30$ for all $x \in \mathbb{R}$, the equation $4x^3 - x^4 = 30$ has no real solutions. $\qquad\square$

STEP 2 is the graphical perceptual inference

STEP 4 is the graphical deductive inference

STEP 6 is the transparently good inference