Mathematicians' assessments of the explanatory value of proofs *

Juan Pablo Mejía Ramos
Rutgers University

Tanya Evans
University of Auckland

Colin Rittberg
Loughborough University &
Vrije Universiteit Brussel

Matthew Inglis
Loughborough University

Abstract. The literature on mathematical explanation contains numerous examples of explanatory, and not so explanatory proofs. In this paper we report results of an empirical study aimed at investigating mathematicians' notion of explanatoriness, and its relationship to accounts of mathematical explanation. Using a Comparative Judgement approach, we asked 38 mathematicians to assess the explanatory value of several proofs of the same proposition. We found an extremely high level of agreement among mathematicians, and some inconsistencies between their assessments and claims in the literature regarding the explanatoriness of certain types of proofs.

Keywords: Mathematical Explanation, Proof, Mathematical Practices, Comparative Judgment

1. Introduction

Outside of the philosophy of mathematics, mathematical explanation is frequently interpreted in a pedagogical sense: mathematical explanations are the kinds of things that someone may say or do to help someone else grasp a mathematical idea. In philosophy of mathematics this notion is often more closely related to scientific explanation, where scientific explanations are those things that account (or should account) for natural phenomena, and the philosophical study of scientific explanation is the characterization of the nature and structure of those explanations. Thus, just like we can think of science as offering explanations in its answers to the questions like "why does salt dissolve in water?" (with philosophers of science studying the nature of those explanations), some philosophers think of mathematics as offering explanations in its accounts of different types of phenomena (with philosophers of mathematics studying the nature of those accounts). Since Steiner (1978) offered a proposal for what constitutes an explanatory proof, the study of mathematical explanation has attracted modest but sustained interest. In the current study we set out to empirically investigate the notion of mathematical explanation held by mathematicians, and its relationship to philosophical accounts of mathematical explanation.

Before delving into the relevant literature, we specify the type of mathematical explanation that constitutes the focus of our study. Lyon and Coyvan (2008) wrote that, depending on the nature of the phenomena being explained, a mathematical explanation is either *extra-mathematical* (if the phenomena being explained is non-mathematical in nature) or *intra-mathematical* (if the explained phenomena itself is mathematical). Lyon and Coyvan were particularly interested in extra-mathematical explanations, those which help explain physical phenomena, such as why hive-bee honeycombs have a hexagonal structure. Here, we focus instead on intra-mathematical explanations, in which what is being explained (the *explanandum*) and what is doing the explaining (the *explanans*) are both within mathematics. More specifically, we focus on the idea that some

* This is a preprint of an article to appear in *Axiomathes*. To cite this article, please refer to the journal version.

mathematical proofs *explain why* a given mathematical theorem holds, while others merely *establish that* the theorem holds. Although our focus is on the explanatory value of mathematical *proofs*, we do not ascribe to the proof-chauvinism criticised by D'Alessandro (2020): the view that "all or most cases of mathematical explanation involve explanatory proofs in an essential way" (p.581). Like D'Alessandro, we believe that proofs do not have a monopoly of explanation in mathematics. Furthermore, the approach we take in this paper could easily be extended to study non-proof explanations.

We begin by describing and illustrating a distinction between two categories into which accounts of mathematical explanation fall, a distinction that was crucial in the design of our study. We then consider the relationship between accounts of mathematical explanation and mathematical practices and address the motivating question of why philosophers interested in mathematical explanation should be interested in the results of our investigations. We then review recent developments in the assessment of the explanatory value of proofs, describe our methodological approach, and present the results of our study. We end the paper by discussing our results in light of current accounts of mathematical explanation, and how the use of the method employed in this paper could help move the field forward.

2. Literature review

Ontic and epistemic accounts of explanation in mathematics

Based on a distinction made by Salmon (1984) in the context of scientific explanation, Delarivière, Frans, and Van Kerkhove (2017) distinguished between *ontic* and *epistemic* accounts of what it means for a proof to have explanatory value in mathematics:

> "An account of explanation is *ontic* if it states:
> Proof *P* of theorem *t* has explanatory value if and only if *P* itself is the explanans of *t* regardless of whether it gives understanding to any particular agent.
>
> An account of explanation is *epistemic* if it states:
> Proof *P* of theorem *t* has explanatory value if and only if the explanans consists of arguments (in the broad sense) including *P* that grants understanding of *t* for a particular agent S." (p. 311)

Whereas in ontic accounts the explanatory value of a proof relies on the extent to which the proof possesses certain characteristics (not necessarily related to understanding), in epistemic accounts the explanatory value of a proof relies on the extent to which the proof grants understanding to a particular agent. Thus, ontic accounts focus on specifying the kinds of (non-epistemic) characteristics that increase the explanatory value of a proof (e.g., in terms of certain mathematical properties), whereas epistemic accounts focus on specifying the type of understanding derived from proofs with higher explanatory value, and the conditions under which such understanding occurs. Crucially, in ontic accounts it is irrelevant (at least in principle) whether a proof with the appropriate characteristics is either understood or understandable by any one agent[1]. In contrast, in epistemic accounts, the assessment of the explanatory value of proofs varies depending on the agent. Proof *P* may have high explanatory value for agent *S*, but not for agent *S'*.

Steiner's (1978) account of what constitutes an explanatory proof in mathematics is an early exemplar of the ontic approach. For Steiner, a proof is explanatory if it deduces the theorem

[1] Delarivière et al. (2017) clarified that ontic accounts do not necessarily *deny* a possible relationship between explanation and understanding; ontic accounts simply do not use understanding as a *defining criterion* of the explanatory value of a proof (p. 312).

about a mathematical object by *evidently* relying on what he calls the *characterizing property* of that object:

> My view exploits the idea that to explain the behavior of an entity, one deduces the behavior from the essence or nature of the entity. Now the controversial concept of an essential property of x (a property x enjoys in all possible worlds) is of no use in mathematics, given the usual assumption that all truths of mathematics are necessary. Instead of 'essence', I shall speak of 'characterizing properties', by which I mean a property unique to a given entity or structure within a *family* or domain of such entities or structures. […]
>
> My proposal is that an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property. It must be evident, that is, that if we substitute in the proof a different object of the same domain, the theorem collapses; more, we should be able to see as we vary the object how the theorem changes in response (p. 143)

Steiner provided several examples of proofs meeting this criterion, including a proof that $\sqrt{2}$ is irrational that appeals to the number of powers of 2 in the prime factorization of a^2 and $2b^2$: since both a^2 and b^2 have an even number of 2s in their prime factorization, a^2 and $2b^2$ must have different prime factorizations, rendering $a^2 = 2b^2$ (and thus $\sqrt{2} = a/b$) impossible. For Steiner, this proof (unlike the classic proof showing it is impossible to express $\sqrt{2} = a/b$ in lowest terms) is explanatory because it relies on the prime power expansion of a number (a 'characterizing property' of number) in such a way that it is evident how the theorem collapses if we substitute the number 2 by any perfect square number, or how the theorem changes if we replace 2 with another non-perfect square number. Steiner's account has received a lot of attention and criticism, which have mainly illustrated its limitations as an account of *all* explanatory proofs in mathematics (e.g., Resnik & Kushner, 1987; Hafner & Mancosu, 2005; Lange, 2014). The crucial point here is that Steiner's account relies on the mathematical properties of the objects in the theorem and its proof, it is not concerned with whether an agent who reads the proof would or could gain increased mathematical understanding. In other words, Steiner's is an ontic, not an epistemic, account.

Delarivière, Frans, and Van Kerkhove's (2017) contextual account, and Inglis and Mejía-Ramos's (2019) functional account are two examples of epistemic accounts of mathematical explanation. While both define mathematical explanation in terms of an agent's understanding, they differ on the particular kind of understanding an explanatory proof may grant, and the types of factors that mediate this understanding. Delarivière et al.'s (2017) contextual account uses an abilities-based type of understanding (see also Avigad, 2008), in which "'Agent S understands X' corresponds to 'agent S possesses particular abilities related to X'" (p. 313). This account focuses on the background, skills, and the epistemic interests of the agent as the main contextual factors mediating such understanding. For instance, Delarivière et al. illustrated how the examples of explanatory proofs provided by Steiner (1978) could succeed or fail as explanations according to their own account. In Delarivière et al.'s account, the 'characterizing property' is no longer independent of an agent: an agent must use their background and skills to identify the property, and what ultimately makes it a 'characterizing' property depends on the agent's own epistemic interests. Similarly, making evident how the theorem depends on that property is no longer a job passively carried out by the proof, but by an agent who must use their background and skills to study how the theorem collapses or changes when different objects are considered.

While Inglis and Mejía-Ramos's (2019) approach is compatible with any cognitive, knowledge-based theory of understanding, the notion of understanding they favour relies on the psychological idea of schema. Schemas are cognitive structures, stored in long-term memory, which help a person integrate existing knowledge with new information observed in the environment. Understanding a mathematical object or phenomenon, on this account, involves constructing a "sufficiently well-organised schema" of that object or phenomenon, with a focus on the ways in which human cognitive architecture (involving sensory, working, and long-term memory) mediates such schema formation. According to Inglis and Mejía-Ramos (2019), the reason Steiner's (1978) criteria may lead to explanatory proofs is because the reference to a 'characterizing property' helps agents link relevant information in sensory, working, and long-term memory, which ultimately facilitates the encoding of new information into a sufficiently well-organized schema of the object or phenomenon. Thus, Steiner's explanatory proof that the $\sqrt{2}$ is irrational would be explanatory for an agent (on account of Steiner's criteria), because its reliance on the prime factorization of numbers would ultimately help that agent form more comprehensive linked schemas concerned with irrational numbers and $\sqrt{2}$. Clearly, referencing a 'characterizing property' is not the only way in which a particular proof may aid this cognitive process, which is how Inglis and Mejía-Ramos (2019) are able to incorporate other accounts (e.g., Kitcher, 1981; Lange, 2014).

Both Delarivière et al.'s (2017) and Inglis and Mejía-Ramos's (2019) approaches lead to a notion of explanatoriness that partly depends upon individual agents (as understanding is sensitive to individual differences in agents' abilities and knowledge), but also depends on factors which could be shared by larger groups of agents (e.g., common contextual factors in Delarivière et al.'s account) and a cognitive architecture shared by all humans (in the case of Inglis and Mejía-Ramos).

Mathematical explanation in mathematical practices

One issue that arises in debates about mathematical explanation concerns the relationship between the notion of explanation, as studied by philosophers of mathematics, and mathematical practices. This relationship is complex and studying it involves addressing general questions such as:

- To what extent do mathematicians describe themselves (or their mathematical work) as explaining mathematical phenomena?
- To what extent do mathematicians' assessments of what is (more) explanatory agree with those of philosophers?
- Are mathematicians concerned with the production of the kinds of explanation discussed in the philosophy of mathematics literature?

Weber and Frans (2017) argued that the role of explanation in mathematical practices affects the philosophical study of mathematical explanation differently depending on the specific aim of the philosophical project. Clearly, if the aim of the project is to describe and evaluate explanatory practices in mathematics (what Weber and Frans called *the analytical aim*), then the role of explanation in mathematical practices is crucial. On the other hand, if the aim of the project is to develop an ideal of the types of explanations that *should* be valued in mathematical practices (which they called *the reflective aim*), then the role that explanation *actually* plays in mathematical practice seems to be less important. However, Weber and Frans acknowledged that even when adopting the reflective aim, philosophers of mathematics should ultimately want to confront their developed ideal of explanation with actual mathematical practices. Without this, it would remain unclear how mathematicians' explanatory practices compare to the philosopher's ideal.

In the literature, questions about mathematical practices at large are often answered by referencing individual mathematicians' work, views, and beliefs (including the philosopher's own views and beliefs), without reference to complementary systematic analyses of the practices of the broader population of mathematicians (but see, e.g., Löwe & Van Kerkhove 2019). This has led to inconsistent claims in the literature regarding the explanatory practices of mathematicians. For example, some have suggested that mathematicians often describe themselves (or their work) as explaining mathematical phenomena (e.g., Hafner & Mancosu, 2005; Steiner, 1978), while others believe that mathematicians rarely do so (e.g., Avigad, 2006; Resnik & Kushner, 1987; Zelcer, 2013). In a recent systematic analysis of the use of explanatory language in a large sample of research papers (all papers uploaded to the ArXiv between January and August 2009), Mejía-Ramos et al. (2019) found no evidence of such extreme prevalences of explanatory talk in mathematical writing (at least when compared to the use of explanatory language in the writing in other scientific fields and in day-to-day discourse): while mathematicians do describe themselves (or their mathematical work) as explaining mathematics in their research papers, they do so around half as often as do physicists in their research papers, or does the general population in day-to-day English.

The general issue motivating the study reported in this paper is the relationship between mathematicians' assessments of explanatoriness and theories of mathematical explanation in the literature. To make progress on this general issue, we need methods to investigate mathematicians' assessments of explanatoriness at scale. In this paper, we focus on introducing comparative judgements as a method that allows one to measure mathematicians' assessments of the explanatory value *of proofs*. However, this method can be easily adapted to investigate other types of mathematical explanations (e.g., the explanatory value of definitions, diagrams, theorems and so on). We suggest that the ability to measure the perceived explanatoriness of mathematical proofs could be useful both for those who adopt an analytical aim and those who adopt a reflective aim, and we return to this issue later in the paper.

Assessing the explanatory value of a proof

To our knowledge, Inglis and Aberdein (2015) were the first researchers to collect a large dataset of mathematicians' assessments of the explanatory value of proofs. In their study, they asked 255 mathematicians to think of a proof that they had read recently and to rate, on a five-point Likert scale (from very inaccurate to very accurate), how well each one of 80 adjectives (including 'explanatory') described it. They then conducted an exploratory factor analysis, a statistical method that uses the strength of the correlation between observed variables (in this case the ratings of the 80 adjectives) to model their variability in terms of a lower number of unobserved variables (called factors). Inglis and Aberdein settled on a model with five factors, which they termed *Aesthetics, Non-Use, Intricacy, Utility,* and *Precision.* For instance, the factor termed *Aesthetics* captured high correlations between mathematicians' ratings for 24 of the adjectives, including 'striking', 'ingenious', 'inspired', 'profound', and 'creative'.

Inglis and Aberdein (2015) found that 'explanatory' had moderately positive loadings on the *Utility* and *Precision* factors, and a moderately negative loading on the *Intricacy* factor. In other words, proofs were likely to be rated as explanatory if they were seen as useful, precise and non-intricate. In a subsequent study, Inglis and Aberdein (2016) investigated mathematicians' levels of agreement on these kinds of judgements by asking 112 mathematicians to rate the same proof using a reduced instrument with only 20 adjectives (four per factor). They found no evidence of a high level of agreement in mathematicians' ratings and hypothesized that there could be large individual differences in how mathematicians evaluate proofs, including with respect to their explanatory value (although 'explanatory' was not itself included in the reduced set of 20 adjectives). The findings of the study presented in this paper challenge this hypothesis.

While Inglis and Aberdein's (2015, 2016) approach provides some information about mathematicians' assessments of the explanatory value of proofs (and a hypothesis about their level of agreement on these assessments), their method was not designed to investigate this particular type of assessment. Using a five-point scale to rate how accurately the adjective 'explanatory' describes a given proof, and modelling mathematicians' appraisals of explanatoriness as a linear combination of more general factors provides us with only a rough approximation of mathematicians' assessments of the explanatory value of proofs. We will return to this point in the Discussion section.

If there are disagreements between different mathematicians' assessments of the explanatory value of the same proof, there are at least two different explanations. They may represent a disagreement about the appropriate answer to a question which has been interpreted in roughly the same manner by everyone. However, such disagreements could also be the product of different interpretations of the same question. For example, if we simply asked mathematicians "is proof *P* explanatory?", disagreements could certainly emerge from mathematicians who interpret the question as "does *P* have mathematical property *E* that makes it inherently explanatory?" (consistent with an ontic account of mathematical explanation), and those who interpret it as "would *P* grant agent *S* type of understanding *U*?" (consistent with an epistemic account of mathematical explanation). Clearly, even among mathematicians who interpret the question in the epistemic sense, we could still have disagreements based on the specific agent and the specific type of understanding considered by individual mathematicians.

In the current study, we wanted to clarify to mathematicians the kind of assessment we were interested in (i.e., narrowing down the notion of explanation of interest). We asked mathematicians to conduct paired comparisons of purported explanations of the same theorem, with each comparison asking them to select the best explanation. With the intention of guiding mathematicians to interpret the explanatory value of a proof in a way that was consistent with ontic accounts of mathematical explanation, we instructed them to focus their assessments on how well the proofs themselves accounted for why the theorem holds, without regard to whether those proofs would provide understanding to any particular agent. We now discuss our method in detail.

3. Method

Approach

Rather than asking mathematicians to judge the explanatoriness of individual explanations, we adopted a comparative judgement approach. Comparative Judgement (CJ) approaches to understanding human judgement exploit the finding that people are better at comparing two objects against each other than at evaluating one object against specific criteria (Thurstone, 1927). For example, people are more consistent when judging whether one room is hotter than another, than when judging the temperature of a single object in degrees Celsius. Thurstone (1927) harnessed this finding to assign temperatures to objects, based on participants 'pairwise judgements of which object was hotter. He also adopted the same technique to construct scales for other physical phenomena, such as weight. Subsequently, and most importantly for our purposes, Thurstone applied CJ techniques to construct scales of subjective phenomena such as social attitudes (Thurstone, 1954).

The CJ approach relies upon the Bradley-Terry model (Bradley & Terry, 1952), which assumes that each explanation i has a parameter β_i which captures its explanatoriness. Given two explanations, i and j, then the probability that i is judged to be more explanatory than j is given by $P(i > j) = \frac{e^{\beta_i}}{e^{\beta_i} + e^{\beta_j}}$. By recording the results of repeated paired comparisons, empirical estimates of β_i and β_j can be obtained. Jones, Bisson, Gilmore & Inglis (2019) suggested that, in their experience, an average of 10 judgements per item (explanations in this case) usually suffices to provide a reliable estimate of the βs.

CJ methods have since been applied to measurement in a variety of contexts, notably education. For example, CJ has been used to assess the quality of students 'essays (Heldsinger & Humphry, 2013) and laboratory reports (McMahon & Jones, 2015). It has also been applied in mathematics, to assess students' understanding of calculus, statistics, and algebra (Bisson, Gilmore, Inglis & Jones, 2016; Jones et al., 2019) and their problem-solving skills (Jones & Inglis, 2015). CJ methods have even been successfully used to assess nebulous constructs such as who is "the better mathematician", as part of a project to track examination standards across time (Jones, Wheadon, Humphries & Inglis, 2016). The commonality across such studies is using CJ to assess constructs – such as explanatoriness – about which experts are expected to have an intuitive understanding, but which they may not be able to fully articulate, or use to make reliable absolute judgements (Pollitt, 2012).

One strength of CJ is that it permits empirical investigation of the extent to which the judges agree about the construct they are asked to judge. For instance, if teachers are asked to repeatedly select which of two students is "the better mathematician" on the basis of their written work, we can quantify the extent to which they agree with each other by calculating an appropriate reliability coefficient. Such a coefficient represents the extent to which the judges agree about what constitutes a good mathematician.

Following a CJ approach, we asked research-active mathematicians to select the best explanation in a series of pairs of mathematical explanations of the same statement, while interpreting the explanatory value of a proof in a way that was consistent with ontic accounts of mathematical explanation. Our primary goal was to investigate whether our participants' judgements cohered with each other's. Our secondary goal was to begin to explore the relationship between mathematicians' assessments of what is explanatory in mathematics and the corresponding assessments made by philosophers in the literature.

Materials

The nine proofs we used in the study were all taken from Ording's (2019) *99 Variations on a Proof* of the proposition:

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

The full set of explanations, typeset as seen by participants is given in the Appendix, and summarized in Table 1.

Explanation Name	Explanation Description
One-line (p. 3)	A one-line argument which asserts the factorisation of the equation's standard form.
Two-column (p. 5)	A two-column proof which reorganises the equation into standard form, then solves it.
Elementary (p. 9)	A narrative version of the same underlying argument as presented in the two-column proof.
Visual (p. 23)	A visual 'proof' which deforms a cube of side length x into a cuboid with volume equal to the equation in standard form, and with the side lengths equal to x - 1, x - 1 and x - 4.
Contradiction (p. 29)	Verification that x = 1 and x = 4 are solutions, followed by a demonstration that the existence of a third solution would imply that 1 = 0.
Contrapositive (p. 31)	A demonstration that if x were neither 1 nor 4, then the LHS of the equation would not equal the RHS.
Substitution (p. 49)	An argument which substitutes $x = y + 1$ into the standard form of the equation and shows that y must be 0 or 3.
Taylor series (p. 75)	Uses the Taylor series expansion of the function represented by the equation's standard form and shows the roots must be 1 or 4.
Experimental (p. 183)	Employs Newton's method with a computer algebra system to show that the equation has roots very near if not equal to 1 and 4.

Table 1: A description of each of the explanations used in the study. Full versions are given in the Appendix. Note that neither the explanations' names nor descriptions were presented to participants. Page references are to Ording's (2019) *99 Variations on a Proof.*

In the selection of these nine proofs, we wanted to end up with a diverse set of proofs containing some of the types of proofs that had been discussed in the mathematics explanation literature. This was done to address our secondary, more exploratory goal: to illustrate the relationship between mathematicians' assessments of what is explanatory in mathematics and the corresponding assessments made by philosophers in the literature.

- The 'Elementary' and 'Two-column' proofs, shown in Figure 1, were of particular importance to explore whether mathematicians are influenced by epistemic factors when asked to judge ontic explanatoriness. We considered these two explanations to be mathematically equivalent in the sense that they both reorganized the original equation into the standard form $ax^3 + bx^2 + cx + d = 0$, then split $9x$ into $5x + 4x$, which permitted the equation to be factorized as $(x^2 - 5x)(x - 1) + 4(x - 1) = 0$ and then $(x^2 - 5x + 4)(x - 1) = 0$. Importantly, the only difference between the two explanations was the level of detail provided (substantially higher in the 'Two-column' proof). If mathematicians were eschewing all epistemic considerations when judging the ontic explanatoriness of these explanations, we would expect that these two explanations would be judged to be similarly explanatory.
- The 'One-line', 'Substitution', and 'Taylor series' proofs were chosen as three direct proofs of different length, and with varying degrees of complexity and generality of approach.

- The ‘Visual’ and ‘Experimental’ explanations both have debatable status as proofs. Furthermore, some have observed that visual proofs are often seen as being explanatory (e.g., Hanna, 2000; Steiner, 1978) and the notion of an explanatory proof is commonly illustrated with visual proofs in the literature (most often with examples in Euclidean geometry, but also with dot-diagrams in elementary number theory).
- The ‘Contradiction’ and ‘Contrapositive’ proofs were chosen as examples of indirect proofs. Traditionally, proofs by contradiction have been discussed in the literature as being generally non-explanatory (Lange, 2016, and Mancosu, 2018, offer some historical examples). However, some authors have offered examples of proofs by contradiction that they deemed to be explanatory (Steiner, 1978; Colyvan, 2012; Hanna, 2018).

Two-Column

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation.

STATEMENT	REASON
1. $x^3 - 6x^2 + 11x - 6 = 2x - 2$	Given.
2. $x^3 - 6x^2 + 11x - 6 + 2 = 2x - 2 + 2$	Addition property of equations.
3. $x^3 - 6x^2 + 11x - 4 = 2x$	Addition.
4. $x^3 - 6x^2 + 11x - 4 - 2x = 2x - 2x$	Subtraction property of equations.
5. $x^3 - 6x^2 + 9x - 4 = 0$	Subtraction.
6. $x^3 - (1+5)x^2 + (5+4)x - 4 = 0$	Addition.
7. $x^3 - x^2 - 5x^2 + 5x + 4x - 4 = 0$	Distributive property.
8. $x^2(x-1) - 5x(x-1) + 4(x-1) = 0$	Factoring.
9. $(x^2 - 5x + 4)(x-1) = 0$	Factoring.
10. $[x^2 - (1+4)x + 4](x-1) = 0$	Addition.
11. $(x^2 - x - 4x + 4)(x-1) = 0$	Distributive property.
12. $[x(x-1) - 4(x-1)](x-1) = 0$	Factoring.
13. $[(x-4)(x-1)](x-1) = 0$	Factoring.
14. $x - 1 = 0$ or $x - 4 = 0$	Zero product property.
15. $x - 1 + 1 = 1$ or $x - 4 + 4 = 4$	Addition property of equations.
16. $x = 1$ or $x = 4$	Addition.

Elementary

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. The degree three equation admits the standard form $x^3 - 6x^2 + 9x - 4 = 0$. Expand the linear term $9x$ as the sum $5x + 4x$ so that a common factor is apparent on the left hand side.

$$(x^3 - 6x^2 + 5x) + (4x - 4) = (x^2 - 5x)(x-1) + 4(x-1)$$

Factoring out $x - 1$ leaves a quadratic, which also factors easily.

$$(x^2 - 5x + 4)(x-1) = (x-4)(x-1)(x-1)$$

Since the right side of the equation is zero, one of the factors $(x-1)$ or $(x-4)$ must be zero. Thus $x = 1$ or $x = 4$.

Figure 1. The ‘Two-column ’and ‘Elementary ’explanations as presented to mathematicians.

Procedure and Participants

Participants were recruited by email. Once they had read information about the study contained in the invitation email, if they wished to participate then they visited a website which explained the purpose of the study and asked them to state their research area, by selecting which category of the Mathematics Subject Classification most of their research fell into. They then read detailed instructions about the study:

> Our aim is to study mathematicians' sense of what makes a good explanation in mathematics. To this end we will ask you to conduct a series of paired comparisons of mathematical explanations. In each comparison you will be asked to read two explanations of a given proposition in mathematics and to choose the one which you think best explains why the proposition holds. […] All arguments you will read in this study come from Philip Ording's book "99 Variations on a Proof". ***In each paired comparison, we want you to think about which argument best explains why the proposition holds, and not to focus on how it might be received by a particular audience.*** (Emphasis in the original.)

Our intention with these instructions was to prompt participants to focus on the explanatory value of a proof in a way that was consistent with ontic, rather than epistemic accounts of mathematical explanation. The instructions also explained how the paired comparison process worked, and asked participants to complete a total of twenty judgements.

Once participants had read the instructions, they clicked through to the first paired comparison which was presented on the No More Marking platform[2]. Participants saw two explanations side-by-side and were asked to click "left" or "right" based on which they thought was the better explanation of why the proposition holds. Once participants had made their selection, another two explanations were presented. Each pairing was selected randomly from the set of possible pairs (the order of explanations in each pairing was also randomized). After participants had completed twenty judgements, their participation in the study finished.

Our data collection proceeded in two stages. Participants in the first stage were research-active mathematicians affiliated with the Department of Mathematics at the University of Auckland. We continued recruiting participants until we had collected a complete dataset, which consisted of twenty comparisons from each of 16 mathematicians. After analysing the data from the first stage we decided to attempt to replicate the study in a new context, and so attempted to recruit 16 further mathematicians, this time affiliated with the Department of Mathematics at Rutgers University. Although we planned to collect data from 16 Rutgers-based mathematicians, we had already obtained 22 complete datasets before we were able to stop data collection. Thus, the final sample consisted of a total of 38 mathematicians and 760 judgements, meaning that we had an average of 84 judgements per explanation, well above Jones et al.'s (2019) recommendation of 10.

The participants researched a wide range of mathematical topics, the most common being combinatorics ($N = 4$), partial differential equations ($N = 4$), K-theory ($N = 3$), and group theory ($N = 3$).

4. Results

The paired comparison data were fitted to the Bradley-Terry model (Bradley & Terry, 1952; Hunter, 2004) using the sirt package in R.[3] The Bradley-Terry model used the set of paired comparison judgements to produce estimates of the 'explanatoriness' of each proof. This was captured with a quality parameter and associated standard error (a measure of the precision with which the parameter was estimated). These parameters were used to explore the mathematicians' judgements further.

[2] www.nomoremarking.com

[3] Data and code available at http://dx.doi.org/10.17028/rd.lboro.12458486

Mathematicians' agreement

To address our main question – do mathematicians agree about the criteria that make explanations explanatory? – we used an inter-rater reliability coefficient based on Bisson, Gilmore, Inglis and Jones's (2016) split-half technique. Specifically, we randomly split the group of Auckland judges into two equal subgroups (each with eight participants), used the Bradley-Terry model to produce parameter estimates separately from the judgements from each group, and then correlated the resulting parameter values. We repeated this process 1000 times (with a new random split in each case) and calculated the average correlation coefficient across the 1000 iterations. If this so-called split-half inter-rater reliability coefficient were close to 1, it would indicate that the mathematicians in our sample completely agreed with each other about which explanations were most explanatory. However, if the coefficient was close to zero, then this would indicate that the mathematicians had completely different conceptions of explanatoriness.

In the Auckland sample the split-half inter-rater reliability coefficient was very high, at .882, indicating that the mathematicians largely had a shared conceptualisation of explanation. We repeated this analysis in the Rutgers sample, finding that the parameter values derived from the Rutgers participants correlated very highly with those derived from the Auckland participants, r = .909. The results from the two samples are shown in Figure 2. To ensure that the reliability coefficient in the Rutgers sample was comparable to the Auckland sample, we used randomly created groups of the same size (i.e., 8 participants) to calculate the split-half inter-rater reliability coefficient. This yielded a value of .910.

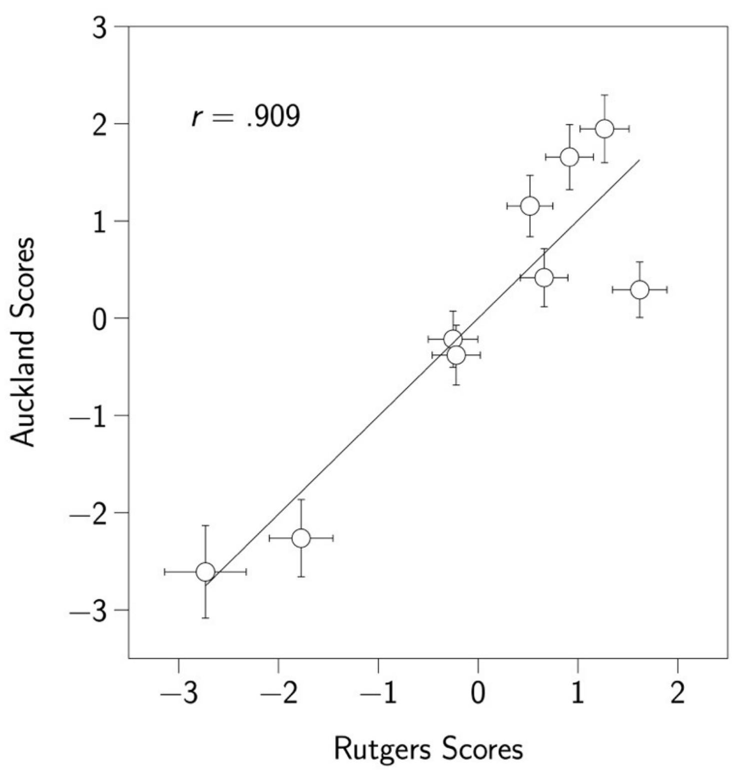


Figure 2. Relationship of parameters obtained from the Auckland and Rutgers samples (the largest disagreement was the 'Substitution' proof). Error bars show ± 1 SE of the mean.

Given that the results derived from the Auckland and Rutgers samples were extremely similar, we refitted the Bradley-Terry model to the combined set of judgements. This again yielded an extremely high split-half inter-rater reliability coefficient of .947. Finally, given that previous research on mathematicians 'evaluations of proofs had found differences between pure and applied mathematicians (e.g., Inglis et al., 2013; Inglis & Aberdein, 2020), we first used our participants ' area of research (as self-reported using the Mathematics Subject Classification) to classify them as either pure or applied mathematicians, and we then fitted the Bradley-Terry model to each one of these two groups. The reliability coefficient for both groups was very high (.929 for pure and .803 for applied mathematicians), as was the correlation between the two groups' parameters (r = .957). In sum, the research mathematicians in our sample tended to agree with each other about which of the proofs best explained why the proposition holds.

Exploratory results

To address our second goal – to explore the relationship between mathematicians 'assessments of what is explanatory in mathematics and the corresponding assessments made by philosophers in the literature – we explored the parameters associated with each of the proofs. These parameters are shown in Figure 3. Note that these values are only meaningful in relation to each other (they are on an arbitrary scale) but that, nevertheless, the numbers are interpretable as a scale (i.e., the gap in explanatoriness between explanations with parameters 0 and 0.5 is the same as the gap between explanations with parameters 0.5 and 1).

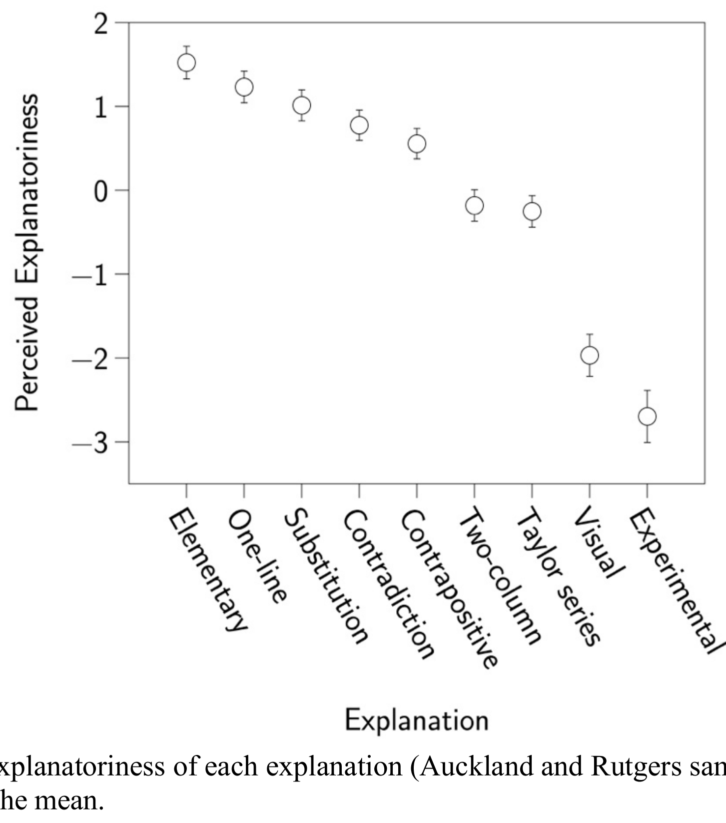


Figure 3. Perceived explanatoriness of each explanation (Auckland and Rutgers samples combined). Error bars show ± 1 SE of the mean.

Below we briefly summarize the explanation parameters by their selection criteria:

- The 'Elementary' and 'Two-column' proofs: The 'Elementary' proof was deemed to be most explanatory by the mathematicians in the study. Critically, the 'Elementary' and 'Two-column' explanations had substantially different parameters, 1.52 and -0.18 respectively. In other words, despite the two proofs presenting what we see as being the same underlying mathematical argument, our participants perceived them to have very different explanatory values. Given this result, and despite the clear experimental instructions, it seems unlikely that our participants' judgements were solely influenced by ontic explanatoriness. We return to this issue in the Discussion.
- The 'One-line', 'Substitution', and 'Taylor series' proofs: With parameters of 1.23, 1.01 and -0.252, respectively, these three direct proofs had substantially different parameters. In particular, the extremely brief 'One-line' proof was deemed the second most explanatory by the mathematicians.
- The 'Visual' and 'Experimental' proofs: With parameters of -1.97 and -2.70 respectively, these two proofs were deemed the least explanatory by the mathematicians in the study.
- The 'Contradiction' and 'Contrapositive' proofs: With parameters of 0.78 and 0.56, respectively, the perceived levels of these indirect proofs were perhaps surprisingly high given suggestions in the literature that indirect proofs are rarely explanatory.

5. Discussion

We see the main contribution of this paper to be methodological. We have successfully used the Comparative Judgement (CJ) approach to measure mathematicians' notion of the ontic explanatory value of nine proofs. This success is partly manifested in the high-level of agreement between mathematicians regarding which of the proofs best explained why the given proposition holds. While it remains unclear the extent to which mathematicians can articulate this notion (or use it to make absolute judgements), or where it comes from (e.g., whether these judgements are the product of enculturation, as internalized norms and values of mathematical practices), we believe that theories of mathematical explanation must be able to account for mathematicians' judgements of explanatoriness of the type we have recorded.

One possible challenge to the claimed success of this approach is that the explanations we used were extremely simple. We note that the mathematician Ording (2019) called these explanations proofs, and that these proofs are not much simpler than the kind of toy examples commonly used in the literature to illustrate the notion of explanatory proof. However, we agree that the CJ method should be tested with more complex results and proofs. On the other hand, the fact that mathematicians displayed such high level of agreement when making these comparative judgements about the explanatoriness of these simple proofs, could present a challenge to the following prediction by Lange (2014):

> My proposal predicts that if the result exhibits no noteworthy feature, then to demand an explanation of why it holds, not merely a proof that it holds, makes no sense. There is nothing that its explanation over and above its proof would amount to until some feature of the result becomes salient.
>
> This prediction is borne out. For example, there is nothing that it would be for some proof to explain why, not merely to prove that, $\int_1^3 (x^3 - 5x + 2)dx = 4$. Nothing about this equation calls for explanation. (p. 507)

It is unclear whether our participants would have been able to make absolute judgements about the explanatoriness of these simple proofs (i.e., to decide that any of these individual proofs explained the given proposition), or whether Lange (2014) would see anything worth explaining in the proposition we used, but our results suggest that mathematicians were able to make sense of our instruction to compare which of our simple proofs best explained why that proposition holds.

Another challenge could come from the comparison of our findings with those of Inglis and Aberdein (2016): while we found mathematicians in our sample tended to agree with each other about the explanatory value of these nine proofs, Inglis and Aberdein found evidence against a high level of agreement in mathematicians' more general proof appraisal, hypothesizing that there could be large individual differences in how mathematicians evaluate proofs (including with respect to their explanatory value). As suggested earlier in the paper, we believe these discrepancies are mainly due to differences in methodological approach: Inglis and Aberdein's dimension reduction approach (which approximates mathematicians' judgements of explanatoriness through Likert-scale ratings of other adjectives that load onto factors with known correlation coefficients with 'explanatoriness') is not as well suited for the study of mathematicians' judgements with respect to individual criteria as our more direct CJ approach. We hypothesize that the CJ approach will in general be a better approach to study mathematicians' proof appraisals with respect to specific criteria. For instance, Inglis and Aberdein found that mathematicians tend to disagree about the aesthetic quality of proofs when asked to make absolute judgements about a single proof in isolation. Future research could productively test whether mathematicians agree about mathematical aesthetics in relative terms, using a similar CJ method as deployed here.

Finally, the high level of agreement in mathematicians' judgments of explanatoriness would seem to pose a challenge for epistemic accounts of mathematical explanation, particularly for philosophers studying explanation with an analytical aim: if the proposal is to use an epistemic account of mathematical explanation to describe explanatory practices in mathematics one would have to justify how the lack of specification of an agent could lead to similar judgements from different mathematicians. However, this challenge can be easily addressed: we would expect very few individual differences in mathematicians' understanding of polynomials and their views of the 'generic student' which they may (despite our instructions) be considering when evaluating these explanations epistemically. In this sense, an epistemic account would predict a lower level of agreement in mathematicians' assessment of the explanatoriness of more complex proofs, or of proofs in more specialized topics. This is a hypothesis that could be tested in future research.

With respect to our more exploratory results, it is worth noting that each of the following observations requires its own study (or sequence of studies). This is partly because claims of the form "Mathematicians find X (non) explanatory", or "Mathematicians find X to be more explanatory that Y", where X and Y represent large categories of proofs, clearly require testing that goes beyond the use of single instances of X and Y. Nevertheless, we believe the following exploratory results constitute promising avenues for future research.

Despite explicitly asking mathematicians to focus on the proofs themselves (and not on how they might be received by a particular audience), the perceived explanatory value of two proofs presenting the same underlying argument ('Elementary 'and 'Two-column') was substantially different. Given that these parameters were obtained from all paired comparisons made by mathematicians (not a single comparison of these two proofs), these substantially different assessments cannot be explained by claiming that mathematicians could have been forced to consider other factors when shown two proofs that they considered to be equally explanatory.

Thus, even when guiding mathematicians to interpret the explanatory value of a proof in a way that was consistent with ontic accounts of mathematical explanation, they seem to have considered factors other than the mathematical reasons offered by the proofs for why the proposition holds.

The challenge for an ontic account is to produce some characteristic of the 'Elementary' proof (other than its underlying mathematical argument) that is not only missing from the 'Two-column' proof but can also be disassociated (at least in principle) from an agent's understanding. On the other hand, while we do not have data on the specific factors that played a role in mathematicians' assessments of the explanatory value of these proofs, those factors could be related to how easy it is to *understand them.* We suspect that the extra level of detail contained in the 'Two-column' explanation was perceived to be excessive, and that this would constitute an obstacle to gaining understanding for readers with an undergraduate (or higher) level of mathematical knowledge. This belief is supported by the educational psychology literature. A level of instructional guidance that is suitable for low-knowledge learners has been found to be disruptive for high-knowledge learners, a result known as the 'expertise-reversal effect' (Kalyuga, 2007). This also suggests that certain types of gaps in proofs (Fallis, 2003; Andersen, 2020) could increase the explanatory value of a proof in mathematical practices. Future research could test this hypothesis.

Finally, in contrast to claims made in the literature regarding the explanatory value of different types of proofs, mathematicians in our study did not seem to judge visual proofs as particularly explanatory, or proofs by contradiction as particularly non-explanatory. Indeed, the fact that our proof by contradiction was deemed to be substantially more explanatory than our visual proof (with more standard, direct proofs being rated somewhere in between) provides an interesting counterexample to general claims about the explanatory value of these two types of proof (at least with respect to their description of explanatory practices of mathematicians). The CJ method could be used to investigate mathematicians' judgements of the explanatoriness in more controversial cases, such as proofs by induction, which have been described in the literature as being both generally explanatory (e.g., Brown, 1999), and generally not explanatory (Lange, 2009). Moreover, this method could be used to investigate the factors that influence mathematicians' judgements of the explanatoriness of these different types of proofs.

References

Andersen, L.E. (2020) Acceptable gaps in mathematical proofs. *Synthese*, ***197***, 233–247.

Avigad, J. (2006). Mathematical method and proof. *Synthese, 153*, 105–159.

Avigad, J. (2008). Understanding proof. In P. Mancosu (Ed.), *The Philosophy of Mathematical Practice* (pp. 317–353). New York: Oxford University Press.

Bisson, M.-J., Gilmore, C., Inglis, M., & Jones, I. (2016). Measuring conceptual understanding using comparative judgement. *International Journal of Research in Undergraduate Mathematics Education, 2*, 141–164.

Bradley, R. & Terry, M. (1952) Rank analysis of incomplete block designs the method of paired comparisons, *Biometrika, 39*, 324–345.

Brown, J.R. (1999). *Philosophy of mathematics: An introduction to a world of proofs and pictures*. London: Routledge.

Colyvan, M. (2012). *An Introduction to the Philosophy of Mathematics.* Cambridge: Cambridge University Press.

D'Alessandro, W. (2020). Mathematical explanation beyond explanatory proof. *British Journal for the Philosophy of Science, 71,* 581–603. https://doi.org/10.1093/bjps/axy009.

Delarivière, S., Frans, J., & Van Kerkhove, B. (2017). Mathematical explanation: A contextual approach. *Journal of Indian Council of Philosophical Research, 34*, 309–329.

Fallis, D. (2003). Intentional gaps in mathematical proofs. *Synthese, 134*, 45–69.

Hafner, J., & Mancosu, P. (2005). The varieties of mathematical explanation. In P. Mancosu et al. (Ed.), *Visualization, Explanation and Reasoning Styles in Mathematics* (pp. 215–250). Berlin: Springer.

Hanna, G. (2000). Proof, explanation and exploration: An overview. *Educational Studies in Mathematics, 44*(1–2), 5–23.

Hanna G. (2018) Reflections on proof as explanation. In A. Stylianides, and G. Harel (Eds.), *Advances in Mathematics Education Research on Proof and Proving* (pp. 3-18). ICME-13 Monographs. Cham, Switzerland: Springer.

Heldsinger, S., Humphry, S. (2013) Using calibrated exemplars in the teacher-assessment of writing: an empirical study. *Educational Research, 55*(3), 219–235.

Hunter, D. R. (2004). MM algorithms for generalized Bradley-Terry models. *Annals of Statistics, 32*, 384–406.

Inglis, M., & Aberdein, A. (2015). Beauty is not simplicity: An analysis of mathematicians' proof appraisals. *Philosophia Mathematica, 23*, 87–109.

Inglis, M., & Aberdein, A. (2016). Diversity in proof appraisal. In B. Larvor (Ed.), *Mathematical cultures* (pp. 163–179). Switzerland: Birkhäuser.

Inglis, M., & Aberdein, A. (2020). Are aesthetic judgements purely aesthetic? Testing the social conformity account. *ZDM Mathematics Education, 52*(6), 1127–1136. https://doi.org/10.1007/s11858-020-01156-8

Inglis, M., Mejía-Ramos, J. P., Weber, K., & Alcock, L. (2013). On mathematicians' different standards when evaluating elementary proofs. *Topics in Cognitive Science, 5*(2), 270–282. https://doi.org/10.1111/tops.12019

Inglis, M., & Mejía-Ramos, J. P. (2019). Functional explanation in mathematics. *Synthese.* https://doi.org/10.1007/s11229-019-02234-5

Jones, I., Bisson, M., Gilmore, C., & Inglis, M. (2019). Measuring conceptual understanding in randomised controlled trials: Can comparative judgement help? *British Educational Research Journal, 45,* 662–680.

Jones, I., & Inglis, M. (2015). The problem of assessing problem solving: Can comparative judgement help? *Educational Studies in Mathematics, 89*, 337–355.

Jones, I., Wheadon, C., Humphries, S., & Inglis, M. (2016). Fifty years of A-level mathematics: Have standards changed? *British Educational Research Journal, 42*, 543–560.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509–539.

Kitcher, P. (1981). Explanatory unification. *Philosophy of Science, 48*, 507–531.

Lange, M. (2009). Why proofs by mathematical induction are generally not explanatory. *Analysis, 69*, 203–211.

Lange, M. (2014). Aspects of mathematical explanation: Symmetry, unity, and salience. *Philosophical Review, 123*, 485–531.

Lange, M. (2016). Explanatory proofs and beautiful proofs. *Journal of Humanistic Mathematics, 6*(1), 8-51. https://doi.org/10.5642/jhummath.201601.04.

Löwe, B., Van Kerkhove, B. (2019). Methodological triangulation in empirical philosophy (of Mathematics). In A. Aberdein, and M. Inglis (Eds), *Advances in Experimental Philosophy of Logic and Mathematics* (pp. 15-38). London: Bloomsbury.

Lyon, A., & Colyvan, M. (2008). The explanatory power of phase spaces. *Philosophia Mathematica, 16*(2), 227–243.

Mancosu, P. (2018). Explanation in Mathematics. In E. N. Zalta (Ed.) *The Stanford Encyclopedia of Philosophy* (Summer 2018 Edition). https://plato.stanford.edu/archives/sum2018/entries/mathematics-explanation/.

McMahon, S., & Jones, I. (2015). A comparative judgement approach to teacher assessment. *Assessment in Education: Principles, Policy & Practice, 22*(3), 368-389. https://doi.org/10.1080/0969594X.2014.978839

Mejía Ramos, J. P., Alcock, L., Lew, K., Rago, P., Sangwin, C., & Inglis, M. (2019). Using corpus linguistics to investigate mathematical explanation. In F. Eugen & C. Mark (Eds.) *Methodological Advances in Experimental Philosophy* (pp. 239–263). London: Bloomsbury.

Ording, P. (2019). *99 Variations on a Proof*. Princeton: Princeton University Press.

Pollitt, A. (2012) The method of adaptive comparative judgement, *Assessment in Education: Principles, Policy & Practice, 19*, 281–300.

Resnik, M. D., & Kushner, D. (1987). Explanation, independence and realism in mathematics. *The British Journal for the Philosophy of Science, 38*(2), 141–158.

Salmon, W. C. (1984). *Scientific explanation and the causal structure of the world.* Princeton, NJ: Princeton University Press.

Steiner, M. (1978). Mathematical explanation. *Philosophical Studies, 34,* 135–151.

Thurstone, L. L. (1927). A law of comparative judgement. *Psychological Review, 34*, 273– 286.

Thurstone, L. L. (1954). The measurement of values. *Psychological Review, 61*, 47–58.

Weber, E., & Frans, J. (2017). Is mathematics a domain for philosophers of explanation? *Journal for General Philosophy of Science, 48*(1), 125–142.

Zelcer, M. (2013). Against mathematical explanation. *Journal for General Philosophy of Science, 44,* 173–192.

Appendix

One-line

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. By subtraction, $x^3 - 6x^2 + 9x - 4 = 0$, which factors as $(x-1)^2(x-4) = 0$.

Two-column

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation.

STATEMENT	REASON
1. $x^3 - 6x^2 + 11x - 6 = 2x - 2$	Given.
2. $x^3 - 6x^2 + 11x - 6 + 2 = 2x - 2 + 2$	Addition property of equations.
3. $x^3 - 6x^2 + 11x - 4 = 2x$	Addition.
4. $x^3 - 6x^2 + 11x - 4 - 2x = 2x - 2x$	Subtraction property of equations.
5. $x^3 - 6x^2 + 9x - 4 = 0$	Subtraction.
6. $x^3 - (1+5)x^2 + (5+4)x - 4 = 0$	Addition.
7. $x^3 - x^2 - 5x^2 + 5x + 4x - 4 = 0$	Distributive property.
8. $x^2(x-1) - 5x(x-1) + 4(x-1) = 0$	Factoring.
9. $(x^2 - 5x + 4)(x-1) = 0$	Factoring.
10. $[x^2 - (1+4)x + 4](x-1) = 0$	Addition.
11. $(x^2 - x - 4x + 4)(x-1) = 0$	Distributive property.
12. $[x(x-1) - 4(x-1)](x-1) = 0$	Factoring.
13. $[(x-4)(x-1)](x-1) = 0$	Factoring.
14. $x - 1 = 0$ or $x - 4 = 0$	Zero product property.
15. $x - 1 + 1 = 1$ or $x - 4 + 4 = 4$	Addition property of equations.
16. $x = 1$ or $x = 4$	Addition.

Elementary

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. The degree three equation admits the standard form $x^3 - 6x^2 + 9x - 4 = 0$. Expand the linear term $9x$ as the sum $5x + 4x$ so that a common factor is apparent on the left hand side.

$$(x^3 - 6x^2 + 5x) + (4x - 4) = (x^2 - 5x)(x - 1) + 4(x - 1)$$

Factoring out $x - 1$ leaves a quadratic, which also factors easily.

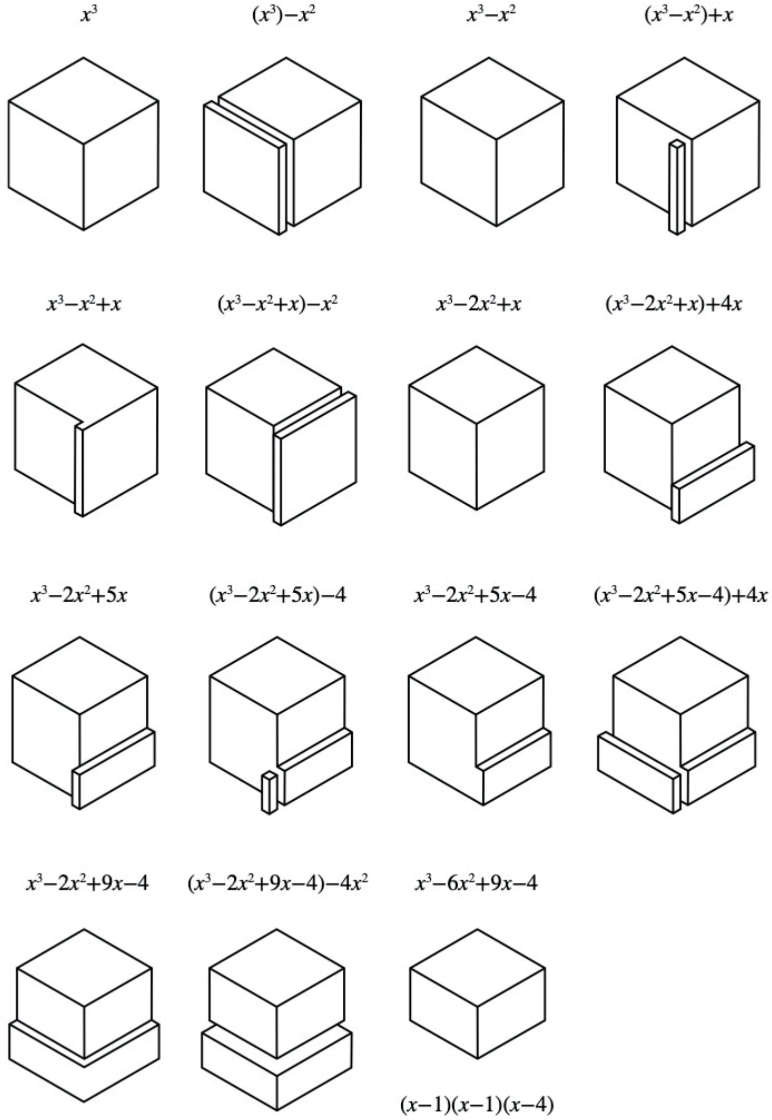
$$(x^2 - 5x + 4)(x - 1) = (x - 4)(x - 1)(x - 1)$$

Since the right side of the equation is zero, one of the factors $(x-1)$ or $(x - 4)$ must be zero. Thus $x = 1$ or $x = 4$.

Visual

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation.



Contradiction

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. That $x = 1$ and $x = 4$ are solutions is readily verified. Suppose there is a third solution x such that $x \neq 1$ and $x \neq 4$. Then we may divide $x^3 - 6x^2 + 11x - 6 = 2x - 2$ by $x - 1$ to obtain $x^2 - 5x + 6 = 2$, that is, $x^2 - 5x + 4 = 0$. Now dividing by $x - 4$ leaves $x - 1 = 0$. Dividing again by $x - 1$ we conclude $1 = 0$, which is absurd. Hence $x = 1$ or $x = 4$ as required.

Contrapositive

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. Suppose $x \neq 1$ and $x \neq 4$. Then

$$(x-1)(x-1)(x-4) \neq 0,$$

since the factors of $(x - 1)$ and $(x - 4)$ are nonzero. Now

$$\begin{aligned}(x-1)(x-1)(x-4) &= x^3 - 6x^2 + 9x - 4 \\ &= (x^3 - 6x^2 + 11x - 6) - (2x - 2),\end{aligned}$$

which must also be different to zero. Hence

$$x^3 - 6x^2 + 11x - 6 \neq 2x - 2,$$

and we have proved by contraposition that the only possible roots are 1 and 4. That these are indeed roots is easily verified.

Substitution

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. By subtracting $2x - 2$ from both sides of the given equation, we have the cubic equation $x^3 - 6x^2 + 9x - 4 = 0$. This can be simplified by the substitution $x = y + 1$:

$$\begin{aligned} 0 &= (y+1)^3 - 6(y+1)^2 + 9(y+1) - 4 \\ &= (y^3 + 3y^2 + 3y + 1) - (6y^2 + 12y + 6) + (9y + 9) - 4 \\ &= y^3 - 3y^2 \\ &= y^2(y - 3). \end{aligned}$$

Therefore y is 0 or 3. It follows that x is 1 or 4, as was to be shown.

Taylor series

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. Define $f : \mathbb{R} \to \mathbb{R}$ to be the function

$$f(x) = x^3 - 6x^2 + 9x - 4.$$

To prove the proposition we need to show that if $f(x) = 0$, then $x = 1$ or $x = 4$. Let's construct the Taylor series of f about the first of the two purported roots, $x = 1$. Compute the derivatives of f: $f'(x) = 3x^2 - 12x + 9$, $f''(x) = 6x - 12$, $f'''(x) = 6$, and $f^{(n)} = 0$ for all $n \geq 4$. Hence

$$\begin{aligned} f(x) &= f(1) + \frac{f'(1)}{1!}(x-1) + \frac{f''(1)}{2!}(x-1)^2 \\ &\quad + \frac{f'''(1)}{3!}(x-1)^3 + \dots \\ &= 0 + 0 - 3(x-1)^2 + (x-1)^3 + 0 \\ &= (x-1)^2(-3 + x - 1) \\ &= (x-1)^2(x-4). \end{aligned}$$

Thus, the roots of f are 1 and 4, as claimed.

Experimental

Proposition. Let $x \in \mathbb{R}$. If $x^3 - 6x^2 + 11x - 6 = 2x - 2$, then $x = 1$ or $x = 4$.

Explanation. To solve the equation $x^3 - 6x + 11x - 6 = 2x - 2$, we employ Newton's method to find the roots of the polynomial $p(x) = x^3 - 6x + 9x - 4$.

That is, we begin with the initial root estimate r_0, and we obtain each subsequent estimate r_{i+1} using linear approximation:

$$r_{i+1} = r_i - \frac{p(r_i)}{p'(r_i)}.$$

This root-finding scheme was implemented using a computer algebra system with initial points $r_0 = 0, 10$. The resulting data suggests that at least two of the roots are very near if not equal to 1 and 4. (See the following table.)

i	r_i	$p(r_i)$	r_i	$p(r_i)$
0	0	-4	10	486
1	0.44444444	-1.09739369	7.42857143	141.6909621
2	0.7020934	-0.29268375	5.76958525	40.25619493
3	0.8446098	-0.07619041	4.75376676	10.62115027
4	0.92043779	-0.01949408	4.21597868	2.23376352
5	0.95971156	-0.00493487	4.02557435	0.23411012
6	0.97972319	-0.00124178	4.00042516	0.00382751
7	0.98982768	-0.00031148	4.00000012	0.00000109
8	0.99490526	-0.00007800	4.00000000	$\approx 10^{-13}$
9	0.99745047	-0.00001952	4.00000000	$\approx 10^{-27}$
10	0.9987247	-0.00000488	4.00000000	$\approx 10^{-56}$
11	0.99936221	-0.00000122	4.00000000	$\approx 10^{-113}$
12	0.99968107	-0.00000031	4.00000000	$\approx 10^{-226}$
13	0.99984053	-0.00000008	4.00000000	$\approx 10^{-453}$
14	0.99992026	-0.00000002	4.00000000	$\approx 10^{-907}$
15	0.99996013	$\approx 10^{-9}$	4.00000000	$\approx 10^{-1815}$

We applied Newton's method to one thousand additional choices of initial point r_0, sampled at random in the range $-10^9 \leq r_0 \leq 10^9$. The results in each case displayed the behaviour of one or the other of the two patterns shown in the table.