Proof, explanation, and justification in mathematical practice

Moti Mizrahi, Florida Institute of Technology

Abstract: In this paper, I propose that applying the methods of data science to "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018) might be a fruitful way to shed new light on the problem. By carefully selecting indicator words for explanation and justification, and then systematically searching for these indicators in databases of scholarly works in mathematics, we can get an idea of how mathematicians use these terms in mathematical practice and with what frequency. The results of this empirical study suggest that mathematical explanations do occur in research articles published in mathematics journals, as indicated by the occurrence of explanation indicators. When compared with the use of justification indicators, however, the data suggest that justifications occur much more frequently than explanations in scholarly mathematical practice. The results also suggest that justificatory proofs occur much more frequently than explanatory proofs, thus suggesting that proof may be playing a larger justificatory role than an explanatory role in scholarly mathematical practice.

1. Introduction

One of the central questions in the philosophy of mathematics is whether mathematical explanations occur in mathematical practice (Mancosu 2018).[1] According to Hanna et al. (2010, p. 2), "philosophers of mathematics have turned their attention more and more from the *justificatory* to the *explanatory* role of proof" (emphasis in original). This suggests that proof plays an equal, dual role in mathematics: a justificatory role and an explanatory role. As Zelcer (2013, p. 173) puts it, "the distinction between *proofs that merely prove* and *proofs that are in some way enlightening*" [...] has attracted the attention of philosophers" lately (emphasis added).[2]

Now, instances of the word 'explain' and its cognates are not difficult to find in mathematics. Here are a few examples (emphasis added):

Our proof *explains* this bump in graph theoretic terms (Cherlin 2016, p. 342).

The above proof *explains why* we consider weak solutions of FBSDEs associated with the problem (Rozkosz 2013, p. 1079).

---

[1] Cf. Mancuso (2008b, p. 135) on the problem of "giving an account of mathematical explanation of empirical phenomena," which is a different problem from the one about the explanatory and justificatory role of proof in mathematics. In this paper, I am concerned with the latter, not the former. See also Mancuso (2018). For recent work on mathematical explanations in science, see Andersen (2018) and Pincock (2015).

[2] On the question, "What are mathematical explanations?" see Inglis and Mejía-Ramos (2019). Again, this question is beyond the scope of this paper. The focus of this paper is "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018). That is, in mathematical practice (specifically, in the published work of practicing mathematicians), are there "proofs that merely prove" or "proofs that are in some way enlightening" as well?

There are also instances of mathematicians reflecting on their practice and saying that proofs are explanatory. Here are a few examples (emphasis added):

> A good proof *explains why* a result is true — it is how we mathematicians come to grips with something (Weintraub 1997, p. xi).

> a proof *explains why* something is true. The only requirement is that the explanation must be logical, so that other people will understand it (Fenton and Dubinsky 1996, p. 38).

> a proof *explains*, via deductive reasoning, why a certain conjecture should be considered true; that is, why the conjecture is really a theorem (Cullinane 2013, p. 125).

Likewise, instances of the word 'justify' and its cognates are not difficult to find in mathematics, either. Here are a few examples (emphasis added):

> The following formal proof *justifies* such a permutation by the absence of any free occurrence of A in $X \epsilon U$ (Nievergelt 2002, p. 118).

> The same proof *justifies* the right to left direction of the lemma used by Sikorski… (Hinkis 2013, p. 332).

There are also those who claim that proofs both explain and justify. For instance, according to Niss (2006, p. 57):

> although some proofs not only justify but also explain why a proposition is true, many proofs justify without providing any explanation; and sometimes there are convincing explanations that cannot easily be formalised into valid proofs without a given theoretical framework (e.g. Stokes' theorem in vector analysis).

To find out how ubiquitous such usage of 'explain' and 'justify' as applied to mathematical proofs is, however, a more rigorous method than selective quotation is needed. I propose that the sort of text mining and corpus analysis methods commonly used by data scientists and corpus linguists can be useful in shedding light on questions about the explanatory and justificatory roles that proofs play in mathematics. After all, the aforementioned quotations suggest that explanations do occur in mathematical practice, but they cannot tell us how frequently explanations and justifications occur in mathematical practice.[3]

Accordingly, the aim of this paper is to examine the explanatory and justificatory roles of proof in mathematical practice by taking an empirical approach.[4] I propose that applying the

---

[3] Cf. Pease et al. (2018) who report some empirical support for their conjecture that "there is such a thing as explanation in mathematics." Their empirical study, however, was not designed to find out how frequently explanations and justifications occur in mathematics.

[4] The empirical methods employed in this paper are the methods of data science and corpus linguistics, such as text mining and corpus analysis, rather than the empirical methods of social science. For examples of the former methods applied to questions in the philosophy of mathematics and logic, see Pease et al. (2018) and Mizrahi (2019). For an example of the latter methods applied to questions in the philosophy of mathematics, see Inglis and Aberdein (2014).

methods of data science to "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018) might be a fruitful way to shed new light on the problem. By carefully selecting indicator words for explanations and justifications, and then systematically searching for these indicators in databases of scholarly works in mathematics, we can get an idea of how practicing mathematicians use these terms and with what frequency. By mining texts from such databases, and running searches designed to find out whether the mathematical practice term 'proof' occurs in explanatory contexts or justificatory contexts, we can say with some confidence whether, and with what frequency, mathematicians use proof in its explanatory role and justificatory role.

Overall, the results of this empirical study suggest that mathematical explanations do occur in mathematical practice, specifically, in research articles published in mathematics journals, as indicated by the occurrence of explanation indicators. When compared with the use of justification indicators, however, the data suggest that justifications occur much more frequently than explanations in mathematical practice. The results also suggest that justificatory proofs occur much more frequently than explanatory proofs, thus suggesting that proof may be playing a larger justificatory role than an explanatory role in mathematical practice. Before I explain these results in detail (Section 3), however, I will describe the methodology of this empirical study in the next section (Section 2). In Section 4, I will discuss the implications of the results of this empirical study to "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018).

2. Methods

In introductions to logic and argumentation, it is customary to distinguish between arguments and explanations. The former attempt to prove, whereas the latter explain why. For instance, according to Fogelin and Sinnott-Armstrong (2005, p. 425):

> Explanations answer questions about how or why something happened. We explain how a mongoose got out of his cage by pointing to a hole he dug under the fence. We explain why Smith was acquitted by saying that he got off on a technicality. The purpose of explanations is *not to prove* that something happened, but *to make sense of things* (emphasis added).

As many authors of logic and argumentation textbooks do, Fogelin and Sinnott-Armstrong (2005, pp. 42-43) provide a list of words that can be used as indicators or markers of arguments. The list includes words, such as 'therefore', 'because', and 'since'. However, words like 'because' and 'since' can indicate both arguments and explanations. As Copi et al. (2011, p. 18) put it when they distinguish between arguments and explanations in their logic textbook:

> If our aim is to establish the truth of some proposition, *Q*, and we offer some evidence, *P*, in support of *Q*, we may appropriately say "*Q* because *P*." In this case we are giving an argument *for Q*, and *P* is our premise. Alternatively, suppose that *Q* is known to be true. In that case we don't have to give any reasons to support its truth, but we may wish to give an account of *why* it is true. Here also we may say "*Q* because *P*"--but in this case we are giving not an argument *for Q*, but an explanation *of Q* (emphasis in original).

For these reasons, words such as 'because' and 'since' are not reliable indicators or markers of explanations. For, as Copi et al. (2011, p. 18) point out, "those words are used both in explanations and in arguments." In addition to 'explain' and its cognates, then, we need words other than 'because' or 'since' that can serve as reliable indicators or markers for explanations as opposed to arguments. Following Overton (2013, p. 1386), I have used 'account', 'explicate', and 'elucidate' as additional indicator words for explanation (in addition to 'explain').[5] Unlike 'because' and 'since', we can be fairly confident that 'account', 'explicate', and 'elucidate' indicate explanations rather than arguments in texts.

As mentioned above, unlike explanations, "which provide reasons for why or how an event occurred" (Baronett 2016, p. 18), the premises of an argument provide justification for its conclusion (Marcus 2018, p. 112). As Govier (2010, p. 2) puts it, "an argument is a reasoned attempt to *justify* a claim on the basis of other claims" (emphasis added). Now, markers or indicator words for arguments include the following: 'therefore', 'thus', 'so', 'proves that', 'shows that', and 'demonstrates that' (Govier 2010, pp. 5-6). The words 'therefore', 'thus', and the like are problematic indicators or markers of justifications for the same reason that words such as 'because' and 'since' are not reliable indicators or markers of explanations, i.e., they can be used to indicate both explanations and arguments. As Walton (2002, p. 279) puts it, "the indicator-words, 'thus', 'therefore', 'consequently', and so forth, are similar, in many ways, in arguments and explanations." When it comes to mathematical proofs, the "distinction between nonexplanatory and explanatory mathematical proofs is often formulated in terms of the difference between proofs that merely establish *that* the conclusion is the case and proofs that establish *why* the conclusion is the case; the former *demonstrate*, while the latter *explain*" (Dutilh Novaes 2019, p. 71; emphasis in original).[6] For these reasons, I have used 'demonstrate', 'show', and 'prove' as additional markers or indicator words for justification (in addition to 'justify'). Unlike 'therefore' and 'thus', we can be fairly confident that 'demonstrate', 'show', and 'prove' indicate justifications rather than explanations in texts. The explanation and justification indicators I have used in this empirical study are listed in Table 1.[7]

Table 1. Indicator words for explanation and justification

Explanation indicators	Justification indicators
account	demonstrate
explain	justify
explicate	prove

[5] Cf. Pease et al. (2018) who use 'expla*' and 'underst*' as explanation indicators.
[6] Dutilh Novaes (2019, p. 73) argues that, "In an explanatory proof, there should be no surprises: each step in the proof must be clear and evident, eliciting immediate understanding in whoever inspects the proof, thus ruling out unexpected 'turns'."
[7] For more on the relationship between philosophy of mathematics and argumentation theory, see Pease et al. (2009). Ashton and Mizrahi (2018a) use a similar methodology and the tools of data science to investigate appeals to intuition in philosophy. See also Ashton and Mizrahi (2018b) and Mizrahi (2019).

| elucidate | show |

The data driving this empirical study is taken from JSTOR Data for Research (jstor.org/dfr). This database allows researchers to search full texts for exact phrases and access the metadata associated with the search results. I have used this database to search for the explanation and justification indicators listed in Table 1 through research articles written in English. JSTOR allows for truncation or "wildcard" searches (using the asterisk * symbol). Accordingly, a search for 'account*' will yield results that include the word 'account' and its cognates, such as 'accounting', 'accounted', etc. Similarly, a search for 'demonstrat*' will yield results that include the word 'demonstrate' and its cognates, such as 'demonstrating', 'demonstrated', etc.

The methods of data science allow us to overcome the limitations of relying on selective quotation (see Section 1). For selected quotations may or may not be representative of mathematics as a whole. However, empirical methodologies have limitations of their own. As far as the methods of data science and corpus linguistics are concerned, there are two major limitations. First, we can only study and analyze what is explicitly mentioned in the corpus. For the purpose of this study, then, our corpus of mathematical texts must contain explicit mentions of explanations and justifications, e.g., instances of 'explain', 'justify', and the like (see Table 1), for us to be able to analyze means, proportions, and patterns of usage. It is reasonable to assume that there would be such explicit mentions of explanation indicators and justification indicators in mathematical texts if proofs really do play an explanatory role and a justificatory role in mathematics.

Second, as with any empirical methodology, there may be some false positives and/or false negatives. When it comes to the methods of data science and corpus linguistics, false negatives could occur when we search for a specific term t in a corpus, but do not find it, even though the corpus contains a synonym of t. For example, although unlikely, it is possible that our corpus of mathematical texts contains no instances of 'explain', and so a search for 'explain' would return zero results, because mathematicians use 'elucidate' instead of 'explain' in all the research articles that make up our corpus. On the other hand, false positives could occur when we find instances of a term t in our corpus, but those instances contain irrelevant uses of t. For the purpose of this study, then, the corpus of mathematical texts must contain not only explicit mentions of explanations and justifications, e.g., instances of 'explain', 'justify', and the like (see Table 1), but also explicit mentions of explanation and justification indicators in the context of talk about proofs. For example, instances of 'explain' that are not about proofs (as in "this proof explains") would be considered false positives for the purposes of this study.

Now, there are two things we can do to overcome the limitations of our empirical, data-driven approach. First, we can refine our search terms. For the purposes of this study, I have followed Pease et al. (2018), who use 'expla*' as an explanation indicator, but also Overton (2013, p. 1386) and added 'account*', 'explicat*', and 'elucidat*' (see Table 1). This search methodology is designed to minimize the number of false negatives, i.e., occurrences of explanation and justification in research articles published in mathematics journals that are indicated by words other than 'explain' and 'justify', by using synonymous indicator words, such as 'account' and 'elucidate'.

Second, we can make sure that our search methodology picks out instances of explanation and justification indicators in the corpus that occur in the context of talk about proof. Since the aim of this paper is to examine the explanatory and justificatory roles of *proof* in mathematical practice, we need to search for occurrences of the explanation and justification indicators listed in Table 1 in the context of talk about proofs in mathematics journals. In other words, instead of finding out the proportion of research articles that contain occurrences of the explanation and justification indicators listed in Table 1 in all research articles published in a particular mathematics journal, we need to find out how often these indicator words occur in research articles that discuss proofs. Naturally, we would expect to find that many (perhaps most) research articles published in mathematics journals contain discussions of proofs. But perhaps not all of them do. Accordingly, I have searched for explanations and justifications in the context of talk about proofs by pairing the explanation and justification indicators listed in Table 1 with the mathematical practice term 'proof'. This means that I have searched for explanations and justifications in the context of talk about proofs according to the following formula: (indicator* AND proof). For example, (expla* AND proof), (demonstrat* AND proof), and so on. This search methodology is designed to minimize the number of false positives, i.e., instances of explanation and justification indicators that are not about proofs, by ensuring that instances of explanation and justification indicators in text are paired with the mathematical practice term 'proof'.

In that respect, my methodology is different from Pease et al.'s (2018) not only in my choice of additional indicators but also in the data used. Pease et al. (2018, p. 6) use data mined from "the Mini-Polymath projects, online collaborations on a blog to solve problems drawn from International Mathematical Olympiads," which contain work in progress, whereas I use data mined from academic journals that publish scholarly work in mathematics, and so contain published work. According to Colyvan et al. (2018, p. 233), however, "mathematicians are notorious for covering their tracks in their written work and rarely commit to print judgments of the explanatory powers of proof. But as anyone who has spent time with mathematicians knows, such judgments are forthcoming in the tea room, in the pub, and even in the classroom." If one is inclined to agree with Colyvan et al. (2018, p. 233), then perhaps one would like to distinguish between the "public face" of mathematical practice and what goes on "behind closed doors." Likewise, Hersh (1991) distinguishes between the "front" and the "back" of mathematics. Along the lines of this front/back distinction, then, Pease et al.'s (2018) data come from the so-called "behind closed doors" or the "back" of mathematical practice (e.g., chat rooms), and so their conclusions apply to this pre-publication aspect of mathematical practice, whereas my data come from the so-called "public face" or the "front" of mathematical practice (e.g., research articles published in mathematics journals), and so my conclusions apply to this post-publication aspect of mathematical practice.

As far as scholarly mathematical practice is concerned, JSTOR Data for Research also allows for searches by subject, such as mathematics, which contains 61 journals. However, the mathematics category on JSTOR contains logic and math education journals as well as pure and applied mathematics journals. In order to focus on pure mathematics and rule out scholarly work in logic and math education, I removed from my datasets journals that publish work in logic, applied mathematics, and mathematics education. After removing those journals, as well as

journals that publish work in languages other than English, I was left with the following mathematics journals, which publish work in pure mathematics, as opposed to logic, applied math, or math education, from which data was mined for this empirical study:

- *American Journal of Mathematics* (1878-2013)
- *American Mathematical Monthly* (1894-2017)
- *Annals of Mathematics* (1884-2019)
- *Journal of Computational Mathematics* (1983-2013)
- *Journal of the American Mathematical Society* (1988-2013)

The years in parentheses indicate the years from which JSTOR has issues of the journal in the database.

Accordingly, I have created datasets of research articles that contain the mathematical practice term 'proof' from each of the mathematics journals listed above, which then served as the base rates for calculating the proportions of explanation and justification indicators in each dataset. This search methodology is designed to test hypotheses about the role of proof in scholarly mathematical practice as follows:

(1) If proofs play an explanatory role in scholarly mathematical practice, then we would expect explanation indicators to occur in the context of talk about proofs in research articles published in mathematics journals.
(2) If proofs play a justificatory role in scholarly mathematical practice, then we would expect justification indicators to occur in the context of talk about proofs in research articles published in mathematics journals.

Testing these hypotheses about the role of proof is philosophically significant because it might provide some empirical insight into "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018). Moreover, as mentioned above, according to Hanna et al. (2010, p. 2), "philosophers of mathematics have turned their attention more and more from the *justificatory* to the *explanatory* role of proof." This suggests that proofs play an equal, dual role in mathematics: a justificatory role and an explanatory role. If proofs play a justificatory role and an explanatory role in scholarly mathematical practice, more or less equally, then we should find that both explanation indicators and justification indicators occur in the context of talk about proofs in research articles published in mathematics journals with more or less equal frequency.

3. Results

As discussed in Section 2, the aim of this paper is to examine the explanatory and justificatory roles of *proof* in scholarly mathematical practice, which is why we need to search for occurrences of the explanation and justification indicators listed in Table 1 in the context of talk about proofs. The proportions of research articles that contain the mathematical practice term 'proof' in each of the mathematics journals tested in this empirical study will then serve as our base rates for calculating the proportions of explanation and justification indicators in research articles published in those journals. These results are summed up in Table 2. All the data

reported in this section were first collected on July 5, 2019, and then checked for accuracy on April 28, 2020.

Table 2. Numbers and percentages of research articles contain the mathematical practice term 'proof' in five mathematics journals (Source: JSTOR Data for Research)

	Total number of research articles	**Number of research articles that contain 'proof'**	**Percentage of 'proof' research articles**
American Journal of Mathematics	7140	5057	70%
American Mathematical Monthly	39747	12541	31%
Annals of Mathematics	6348	4679	73%
Journal of Computational Mathematics	1601	1189	74%
Journal of the American Mathematical Society	1072	997	93%

As we can see from Table 2, with the exception of the *American Mathematical Monthly*, where 31% of the research articles published in this journal contain the mathematical practice term 'proof', most (between 70% and 93%) of the research articles published in all the other mathematics journals from which data was mined for this empirical study contain some discussion of proofs.

Now that we have our prior probabilities of research articles that contain the mathematical practice term 'proof' in the five mathematics journals tested in this empirical study, we can use them to calculate the proportions of explanation indicators and justification indicators in the context of talk about proofs. As discussed in Section 2, this methodology can help us address the question of the role of proof in scholarly mathematical practice. That is, to find out what role proof plays in scholarly mathematical practice, we need to search for the explanation and justification indicators listed in Table 1 in the context of talk about proofs and then compare the results. In practice, this means using the following syntax to run queries in JSTOR Data for Research's dataset construction interface: jcode:(journal's jcode) (proof AND indicator*). For example, the jcode for the *American Journal of Mathematics* is amerjmath. Accordingly, to find out how many instances of the explanation indicator 'account' and its cognates there are in the context of talk about proofs in research articles published in the *American Journal of Mathematics*, we would run the following search query: jcode:(amerjmath) (proof AND account*). This query will yield the number of research articles that contain instances of the explanation indicator 'account' and its cognates in the context of talk about proofs. Likewise, to find out how many instances of the justification indicator 'demonstrate' and

its cognates there are in the context of talk about proofs in research articles published in the *American Journal of Mathematics*, we would run the following search query: jcode:(amerjmath) (proof AND demonstrat*). This query will yield the number of research articles that contain instances of the justification indicator 'demonstrate' and its cognates in the context of talk about proofs.[8]

Let's begin with the data on the explanation indicators listed in Table 1. These results are summarized in Table 3.

Table 3. Numbers of research articles that contain each explanation indicator in the context of talk about proofs by journal (Source: JSTOR Data for Research)

	Number of research articles that contain 'proof'	account* AND proof	expla* AND proof	explicat* AND proof	elucidat* AND proof
American Journal of Mathematics	5057	1005	1171	14	47
American Mathematical Monthly	12541	1586	2277	48	98
Annals of Mathematics	4679	1095	1334	17	31
Journal of Computational Mathematics	1189	138	113	1	1
Journal of the American Mathematical Society	997	252	514	2	7

Since we would like to be able to compare the proportions of research articles in which explanation indicators occur in the context of proof talk with the proportions of research articles in which justification indicators occur in the context of proof talk, we need to calculate the proportions of research articles that contain each of the explanation indicators in the context of proof talk. These results are depicted in Figure 1.

---

[8] The jcodes for the other mathematics journals tested in this empirical study are as follows: *American Mathematical Monthly* (amermathmont), *Annals of Mathematics* (annamath), *Journal of Computational Mathematics* (jcompmath), and *Journal of the American Mathematical Society* (jamermathsoci).

Figure 1. Proportions of research articles that contain each of the explanation indicators in the context of talk about proofs by journal (Source: JSTOR Data for Research)
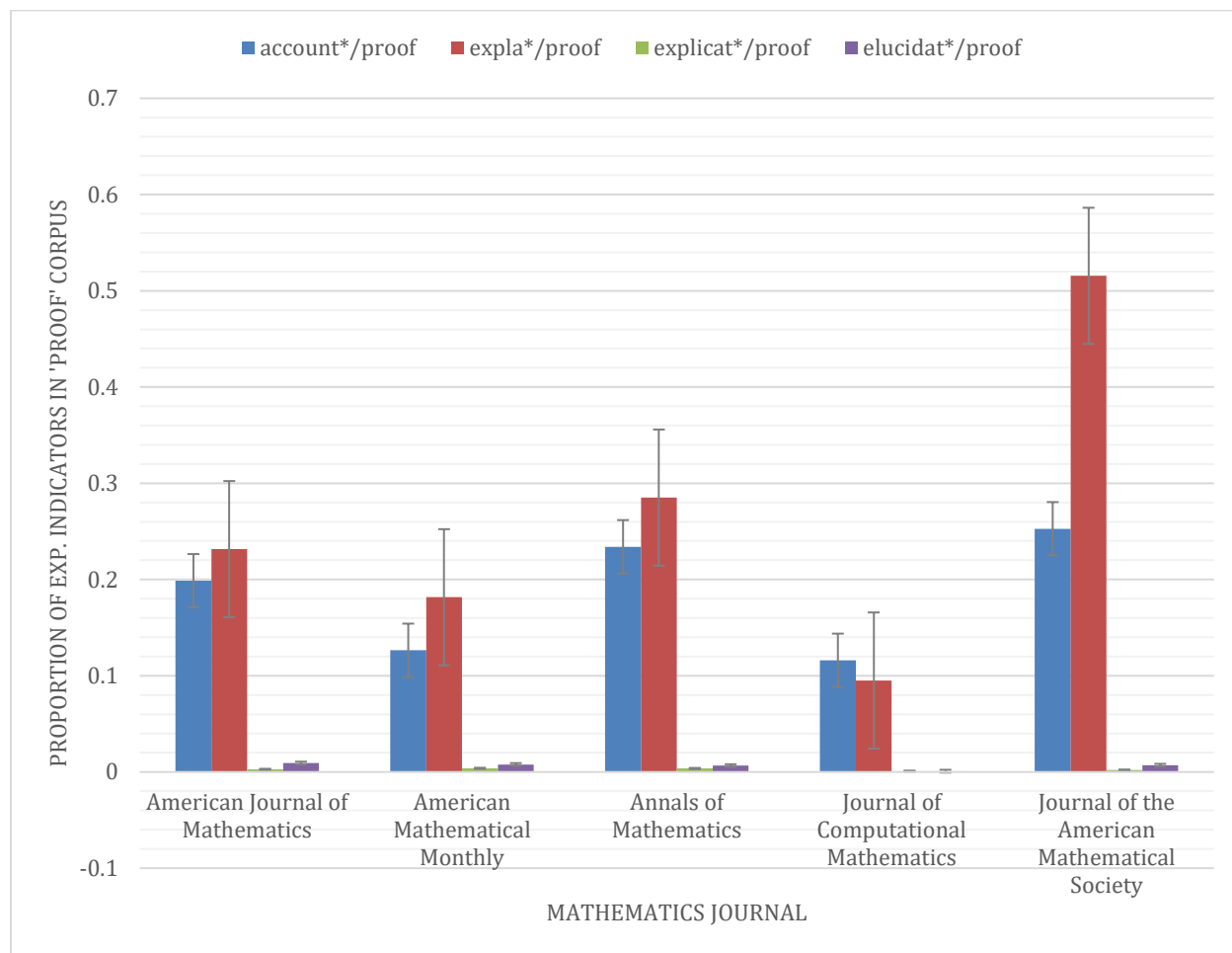
As we can see from Figure 1, the highest proportions are those of the explanation indicator 'explain' and its cognates across the mathematics journals tested for this empirical study, except for the *Journal of Computational Mathematics*, where the highest proportion is that of the explanation indicator 'account' and its cognates.

To check that the search methodology described in Section 2 returns genuine instances of the phenomenon in question (namely, instances of explanation indicators in the context of talk about proof), I have selected at random three search results from the dataset for explanation indicators (emphasis added):

1. "to *explain* the idea of the general *proof*, we demonstrate the proof of the functional equation in a simple example, which also explains the symbols used in the next sections" (Komori 2013, p. 1020).
2. "We include here a *proof* in order to *explain* the relation between x and s" (Freniche 2010, p. 443).

3. "The rest of the *proof* is exactly the same so this *accounts* for..." (Bestvina et al. 2013, p. 1458).

These instances of explanation indicators in research articles published in mathematics journals also provide context to the statistical results reported above. They illustrate how mathematicians use explanation indicators when they talk about proofs in scholarly mathematical practice.

Let's move on to the data on the justification indicators listed in Table 1. These results are summarized in Table 4.

Table 4. Numbers of research articles that contain each justification indicator in the context of talk about proofs by journal (Source: JSTOR Data for Research)

	Number of research articles that contain 'proof'	demonstrat* AND proof	justif* AND proof	prov* AND proof	show* AND proof
American Journal of Mathematics	5057	579	588	4734	4677
American Mathematical Monthly	12541	1493	990	9541	8466
Annals of Mathematics	4679	617	710	4331	4271
Journal of Computational Mathematics	1189	241	61	1038	1013
Journal of the American Mathematical Society	997	131	176	993	815

As before, since we would like to be able to compare the proportions of research articles in which justification indicators occur in the context of proof talk with the proportions of research articles in which explanation indicators occur in the context of proof talk, we need to calculate the proportions of research articles that contain each of the justification indicators in the context of proof talk. These results are depicted in Figure 2.

Figure 2. Proportions of research articles that contain each of the justification indicators in the context of talk about proofs by journal (Source: JSTOR Data for Research)
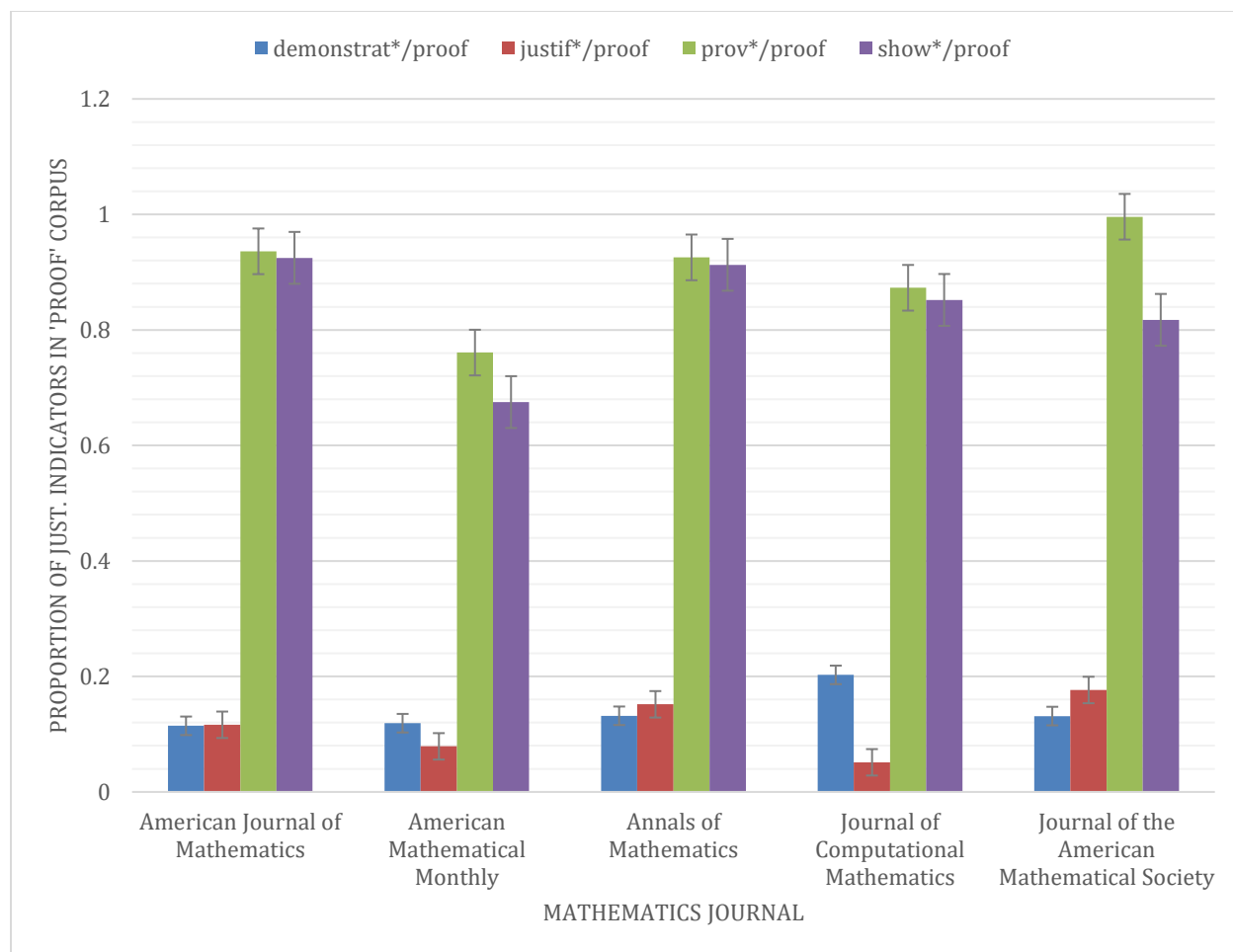
As we can see from Figure 2, the highest proportions are those of the justification indicator 'prove' and its cognates across all the mathematics journals tested for this empirical study.

Again, to check that the search methodology described in Section 2 returns genuine instances of the phenomenon in question (namely, instances of justification indicators in the context of talk about proof), I have selected at random three search results from the dataset for justification indicators (emphasis added):

1. "*Proof of (b). To justify* (b) for…" (Martel and Merle 2011, p. 847).
2. "*Proof.* We first *demonstrate* (i)" (Ren 2012, p. 540).
3. "Resuming the *proof* of the main theorem, it has been *shown* that…" (Sylvester 1886, p. 245).

These instances of justification indicators in research articles published in mathematics journals also provide context to the statistical results reported above. They illustrate how mathematicians use justification indicators when they talk about proofs in scholarly mathematical practice.

Now that we have the proportions of explanation indicators and justification indicators in the context of proof talk, we are in a position to compare these proportions. From Figures 1 and 2, it is evident that justification indicators occur much more frequently in the context of proof talk than explanation indicators do. For example, the explanation indicator 'explain' and its cognates occur in 51% of the research articles published in the *Journal of the American Mathematical Society* that contain proof talk. Among the mathematics journals tested in this empirical study, it is the highest proportion of research articles that contain any of the explanation indicators listed in Table 1. By contrast, in the same journal, the justification indicator 'prove' and its cognates occur in 99% of the research articles that contain proof talk. Among the mathematics journals tested in this empirical study, it is the highest proportion of research articles that contain any of the justification indicators listed in Table 1.

Nevertheless, it would be useful to test rigorously if these differences between the proportions of explanation and justification indicators are statistically significant. To do so, I compared the most frequently mentioned explanation indicator with the most frequently mentioned justification indicator within research articles published by the same mathematics journal (see Table 5).

Table 5. Results of z-tests for proportions comparing the most frequent explanation indicators with the most frequent justification indicators by journal (Source: JSTOR Data for Research)

	Most frequent explanation indicator	**Most frequent justification indicator**	**z-value**	**p**
American Journal of Mathematics	expla* 0.23	prov* 0.93	70.59	0.00
American Mathematical Monthly	expla* 0.18	prov* 0.76	91.88	0.00
Annals of Mathematics	expla* 0.28	prov* 0.92	63.38	0.00
Journal of Computational Mathematics	account* 0.11	prov* 0.87	36.91	0.00
Journal of the American Mathematical Society	expla* 0.51	prov* 0.99	24.96	0.00

For example, the aforementioned difference is indeed statistically significant. That is, a z-test for proportions was conducted to find that the difference between the proportion of research articles published in the *Journal of the American Mathematical Society* that contain the explanation indicator 'explain' and its cognates in the context of proof talk (0.51) and the proportion of research articles published in the *Journal of the American Mathematical Society* that contain the justification indicator 'prove' and its cognates in the context of proof talk (0.99) is statistically significant ($z = 24.96$, $p = 0.00$, two-sided). As we can see from Table 5, the same can be said about the difference in proportions between those of the top explanation indicator, which is

'explain' and its cognates in all the mathematics journals tested in this empirical study, with the exception of the *Journal of Computational Mathematics* (where the most frequently mentioned explanation indicator is 'account' and its cognates), and those of the top justification indicator, which is 'prove' and its cognates in all the mathematics journals tested in this empirical study. All these differences in proportions are statistically different at 99%. These results suggest that justification indicators occur significantly more frequently than explanation indicators do in the context of talk about proofs in research articles published in mathematics journals.

In addition, an independent-samples t-test was conducted to compare the proportions of explanatory proofs and the proportions of justificatory proofs across all the mathematics journals tested in this empirical study. There was a significant difference between justificatory proofs (M = 0.49, SD = 0.38, N = 20) and explanatory proofs (M = 0.11, SD = 0.13, N = 20), t(24) = -4.15, $p < 0.00$, two-tailed. These results suggest that justificatory proofs are significantly more frequent than explanatory proofs in scholarly mathematical practice.

4. Discussion

As discussed in Section 2, this empirical study was designed to test hypotheses about the role of proof in scholarly mathematical practice as follows:

(1) If proofs play an explanatory role in scholarly mathematical practice, then we would expect explanation indicators to occur in the context of talk about proofs in research articles published in mathematics journals.
(2) If proofs play a justificatory role in scholarly mathematical practice, then we would expect justification indicators to occur in the context of talk about proofs in research articles published in mathematics journals.

The results of this empirical study show that explanation indicators occur in the context of talk about proofs in research articles published in mathematics journals. In that respect, these results shed new light on "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018), for they suggest that explanations do occur in scholarly mathematical practice because explanation indicators do appear in research articles published in mathematics journals. That is, if the explanation indicators listed in Table 1 are reliable indicators for the presence (or absence) of explanations in scholarly mathematical practice, and if what practicing mathematicians say and do in their published work is representative of scholarly mathematical practice, then the results of this empirical study suggest that explanations do occur within mathematics itself. This result is in line with Pease et al.'s (2018, p. 17) data in support of their conjecture that "there is such a thing as explanation in mathematics."

The results of this empirical study also show that justification indicators occur in the context of talk about proofs in research articles published in mathematics journals as well. That is, if the justification indicators listed in Table 1 are reliable indicators for the presence (or absence) of justifications in scholarly mathematical practice, and if what practicing mathematicians say and do in their published work is representative of scholarly mathematical practice, then the results of this empirical study suggest that justifications occur within mathematics itself as well.

Interestingly, however, when we compare the proportion of explanation indicators in mathematics research articles that contain proof talk with the proportion of justification indicators in mathematics research articles that contain proof talk across all the mathematics journals tested in this empirical study, we find that, in general, the latter is significantly larger than the former. In other words, mathematical explanations do occur in scholarly mathematical practice, as indicated by the occurrence of explanation indicators in mathematics research articles, but not as frequently as justifications generally do. In other words, the results of this empirical study suggest that justifications occur significantly more frequently than explanations do in mathematics itself.

As far as the question concerning the explanatory and justificatory roles of proof in mathematical practice, then, the results of this empirical study suggest that proof does play this dual role in scholarly mathematical practice, given that both explanation indicators and justification indicators occur in the context of talk about proofs in research articles published in mathematics journals. However, once again, when we compare the proportion of explanation indicators in the context of proof talk to the proportion of justification indicators in the context of proof talk in research articles published in all the mathematics journals tested in this empirical study, we find that, in general, the latter is significantly larger than the former. This result suggests that proof may be playing a larger justificatory than explanatory role in scholarly mathematical practice. This result is not what we would have expected to find if proof plays an equal, dual role in mathematics: a justificatory role and an explanatory role. As discussed in Section 2, if proofs played a justificatory role and an explanatory role in scholarly mathematical practice, more or less equally, then we would find that both explanation indicators and justification indicators occur in the context of talk about proofs in research articles published in mathematics journals with more or less equal frequency. What we actually find, however, is that justificatory proofs are significantly more frequent than explanatory proofs in scholarly mathematical practice.

For philosophers of mathematics, I submit, the philosophical significance of these results consists in getting us a bit closer to having a more accurate picture of mathematical practices. After all, to study any practice, we need to have an accurate picture of what that practice is like. The results of this empirical study, then, contribute to this ongoing project in philosophy of mathematics. For if there are significantly more justificatory proofs than explanatory proofs in scholarly mathematical practice, as the results of this empirical study suggest, but philosophers of mathematics focus on one more than the other, then it looks like philosophers of mathematics might be getting a rather distorted picture of scholarly mathematical practice. If philosophers of mathematics want to have an accurate picture of what scholarly mathematical practice is like, then it is important to know which role of proof is the rule and which role of proof is the exception. The results of this empirical study suggest that, as far as scholarly mathematical practice is concerned, justificatory proofs, not explanatory proofs, are the rule. In that case, "taking mathematical practice seriously" (Carter 2019, p. 2) means taking the aforementioned results suggesting that justificatory proofs are the rule, whereas explanatory proofs are the exception, seriously and adjusting our attention as philosophers of mathematics accordingly.

Beyond the aforementioned results, the methodology used to obtain these results has philosophical significance as well. According to Mancosu (2008a, p. 2), "attention to mathematical practice is a necessary condition for a renewal of the philosophy of mathematics." Along these lines, the methodology employed here provides an empirical way to study mathematical practices on a broader scale than the traditional methods of philosophy of mathematics, such as the method of case studies.[9] Relying on a few case studies might provide an inaccurate picture of what mathematical practices are really like, for the selected case studies may simply be outliers.[10] On the other hand, the methods of text mining and corpus analysis used in this paper can provide a more accurate picture of what mathematical practices are like than the case study method can precisely because of the use of more data systematically mined from databases of large corpora of scholarly work done by practicing mathematicians.[11] In other words, if philosophers of mathematics are serious about "taking mathematical practice seriously" (Carter 2019, p. 2), then that means an "[e]xtension of methodologies brought in to deal with [questions related to mathematics]" (Carter 2019, p. 27). In that respect, the methods of data science, such as text mining, corpus analysis, data visualization, and the like, could provide a set of useful tools for studying mathematical practices.

Of course, investigating particular cases of explanatory proofs in mathematical practices may still be a useful and worthwhile endeavor in philosophy of mathematics. In that respect, it is important to recall the distinction between the so-called "behind closed doors" or the "back" of mathematical practice (e.g., chat rooms) and the so-called "public face" or the "front" of mathematical practice (e.g., research articles published in mathematics journals), and that the conclusions of this empirical study apply to the latter, not the former. It is possible, then, that mathematical explanations could occur implicitly in the "back rooms" or "chat rooms" of mathematics, which can be accessed by philosophers of mathematics through case studies and interviews with practicing mathematicians. Accordingly, empirical and quantitative methods, such as those employed in this paper, can serve to complement rather than replace traditional and qualitative methods of philosophical inquiry, such as the method of case studies.

Some philosophers of mathematics might insist that they *should* direct their attention to explanatory proofs, the empirical evidence suggesting that explanatory proofs are significantly less frequent than justificatory proofs in scholarly mathematical practice notwithstanding. This would be a *normative* claim, of course, on which the empirical findings of this empirical study have no bearing directly, unless those philosophers of mathematics share an interest in portraying mathematical practices in our philosophical accounts of mathematics as accurately as possible. If they do not share this research interest, however, then those philosophers of mathematics might argue that philosophers of mathematics *should* direct their attention to explanatory proofs for reasons other than accuracy. For instance, they might think that explanatory proofs are particularly interesting, more so than justificatory proofs, and thus deserving of the attention of philosophers of mathematics. Nevertheless, as Maddy (1997, p. 161) puts it, "If our philosophical account of mathematics comes into conflict with successful mathematical practice,

---

[9] See, e.g., McLarty (2008) for a use of a case study in the philosophy of mathematics.

[10] On methodological issues in philosophy of mathematics, see Cellucci (2013). On the use of case studies in philosophy of science, see Mizrahi (2020).

[11] For more on the application of text mining and corpus analysis methods to philosophy of logic and mathematics, see Pease et al. (2018) and Mizrahi (2019).

it is the philosophy that must give." Accordingly, if we have reasons to believe on empirical grounds that justificatory proofs, not explanatory proofs, are the rule in scholarly mathematical practice, then our philosophical accounts of mathematics need to account for these empirical findings, or so I would suggest.

5. Conclusion

In this paper, I have taken an empirical approach to "the problem of whether mathematical explanations occur within mathematics itself" (Mancosu 2018). In particular, I have applied the sort of text mining and corpus analysis methods commonly used by data scientists and corpus linguists to questions about the explanatory and justificatory roles that proofs play in mathematics. The results of this empirical study suggest that mathematical explanations do occur in scholarly mathematical practice, as indicated by the occurrence of explanation indicators in research articles published in mathematics journals. When compared with the use of justification indicators, however, the data suggest that justifications occur much more frequently than explanations in scholarly mathematical practice. The results also suggest that justificatory proofs occur much more frequently than explanatory proofs, thus suggesting that proof may be playing a larger justificatory role than an explanatory role in scholarly mathematical practice. I propose that our philosophical accounts of mathematics need to explain (or, at the very least, explain away) these empirical findings.

References

Andersen, H. (2018). Complements, Not Competitors: Causal and Mathematical Explanations. *British Journal for the Philosophy of Science* 69 (2): 485-508.

Ashton, Z. and Mizrahi, M. (2018a). Intuition Talk is Not Methodologically Cheap: Empirically Testing the "Received Wisdom" about Armchair Philosophy. *Erkenntnis* 83 (3): 595-612.

Ashton, Z. and Mizrahi, M. (2018b). Show Me the Argument: Empirically Testing the Armchair Philosophy Picture. *Metaphilosophy* 49 (1-2): 58-70.

Baronett, S. (2016). *Logic*. Third Edition. New York: Oxford University Press.

Bestvina, M. B., Bromberg, K., Fujiwara, K., and Souto, J. (2013). Shearing Coordinates and Convexity of Length Functions on Teichmuller Space. *American Journal of Mathematics* 135 (6): 1449-1476.

Carter, J. (2019). Philosophy of Mathematical Practice--Motivations, Themes and Prospects. *Philosophia Mathematica* 27 (1): 1-32.

Cellucci, C. (2013). Philosophy of Mathematics: Making a Fresh Start. *Studies in History and Philosophy of Science Part A* 44 (1): 32-42.

Cherlin, G. (2016). On the Relational Complexity of a Final Permutation Group. *Journal of Algebraic Combinatorics* 43 (2): 339-374.

Colyvan, M., Cusbert, J., and McQueen, K. 2018. Two Flavours of Mathematical Explanation. In A. Reutlinger and J. Saatsi (eds.), *Explanation Beyond Causation: Philosophical Perspectives on Non-causal Explanations* (pp. 231-249). Oxford: Oxford University Press.

Copi, I. M., Cohen, C., and McMahon, K. (2011). *Introduction to Logic*. Fourteenth Edition. Upper Saddle River, NJ: Prentice Hall.

Cullinane, M. J. (2013). *A Transition to Mathematics with Proofs*. Burlington, MA: Jones & Bartlett Learning.

Dutilh Novaes, C. (2019). The Beauty (?) of Mathematical Proofs. In A. Aberdein, and M. Inglis (eds.), *Advances in Experimental Philosophy of Logic and Mathematics* (pp. 63-93). London: Bloomsbury Press.

Fenton, W. and Dubinsky, E. (1996). *Introduction to Discrete Mathematics with ISETL*. New York: Springer.

Fogelin, R. J. and Sinnott-Armstrong, W. (2005). *Understanding Arguments: An Introduction to Informal Logic*. Seventh Edition. Belmont, CA: Thomson Wadsworth.

Freniche, F. J. (2010). On Riemann's Rearrangement Theorem for the Alternating Harmonic Series. *The American Mathematical Monthly* 117 (5): 442-448.

Govier, T. (2010). *A Practical Study of Argument*. Seventh Edition. Belmont, CA: Cengage Learning.

Hanna, G., Jahnke, H. N., and Pulte, H. (2010). Introduction. In G. Hanna, H. N. Jahnke, and H. Pulte (eds.), *Explanation and Proof in Mathematics: Philosophical and Educational Perspectives* (pp. 1-13). Dordrecht: Springer.

Hersh, R. (1991). Mathematics has a Front and a Back. *Synthese* 88 (2): 127-133.

Hinkis, A. (2013). *Proofs of the Cantor-Bernstein Theorem: A Mathematical Excursion*. Basel: Springer.

Inglis, M. and Aberdein, A. (2014). Beauty Is Not Simplicity: An Analysis of Mathematicians' Proof Appraisals. *Philosophia Mathematica* 23 (1): 87-109.

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

Komori, Y. (2013). Functional Equations of Weng's Zeta Functions for (G, P)/$\mathbb{Q}$. *American Journal of Mathematics* 135 (4): 1019-1038.

Maddy, P. (1997). *Naturalism in Mathematics*. Oxford: Oxford University Press.

Mancosu, P. (2008a). Introduction. In P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (pp. 1-21). New York: Oxford University Press.

Mancosu, P. (2008b). Mathematical Explanation: Why It Matters. In P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (pp. 134-149). New York: Oxford University Press.

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

Marcus, R. (2018). *Introduction to Formal Logic with Philosophical Applications*. New York: Oxford University Press.

Martel, Y. and Merle, F. (2011). Description of Two Soliton Collision for the Quartic gKdV equation. *Annals of Mathematics* 174 (2): 757-857.

McLarty, C. (2008). 'There Is No Ontology Here': Visual and Structural Geometry in Arithmetic. In P. Mancosu (ed.), *The Philosophy of Mathematical Practice* (pp. 370-406). New York: Oxford University Press.

Mizrahi, M. (2019). What Isn't Obvious about 'Obvious': A Data-Driven Approach to Philosophy of Logic. In A. Aberdein, and M. Inglis (eds.), *Advances in Experimental Philosophy of Logic and Mathematics* (pp. 201-224). London: Bloomsbury Press.

Mizrahi, M. (2020). The Case Study Method in Philosophy of Science: An Empirical Study. *Perspectives on Science* 28 (1): 63-88.

Niss, M. (2006). The Structure of Mathematics and Its Influence on the Learning Process. In J. Maasz and W. Schlöglmann (eds.), *New Mathematics Education Research and Practice* (pp. 51-62). Rotterdam: Sense Publishers.

Nievergelt, Y. (2002). *Foundations of Logic and Mathematics: Applications to Computer Science and Cryptography*. Boston: Birkhauser.

Overton, J. A. (2013). "Explain" in Scientific Discourse. *Synthese* 190 (8): 1383-1405.

Pease, A., Smaill, A., Colton, S., and Lee, J. (2009). Bridging the Gap Between Argumentation Theory and the Philosophy of Mathematics. *Foundations of Science* 14 (1-2): 111-135.

Pease, A., Aberdein, A., and Martin, U. (2018). Explanation in Mathematical Conversations: An Empirical Investigation. *Philosophical Transactions of the Royal Society A* 377: https://doi.org/10.1098/rsta.2018.0159.

Pincock, C. (2015). Abstract Explanations in Science. *British Journal for the Philosophy of Science* 66 (4): 857-882.

Ren, Z. (2012). Banded Toeplitz Preconditioners for Toeplitz Matrices from Sinc Methods. *Journal of Computational Mathematics* 30 (5): 533-543.

Sylvester, J. (1886). Lectures on the Theory of Reciprocants. *American Journal of Mathematics* 8 (3): 196-260.

Rozkosz, A. (2013). Stochastic Representation of Weak Solutions of Viscous Conservation Laws: A BSDE Approach. *Journal of Theoretical Probability* 26 (4): 1061-1083.

Walton, D. N. (2002). *Legal Argumentation and Evidence*. University Park, PA: The Pennsylvania State University.

Weintraub, S. H. (1997). *Differential Forms: A Complement to Vector Calculus*. New York: Academic Press.

Zelcer, M. (2013). Against Mathematical Explanation. *Journal for General Philosophy of Science* 44 (1): 173-192.