Do Mathematical Explanations Have Instrumental Value?

Rebecca Lea Morris

January 25, 2019

Abstract

Scientific explanations are widely recognized to have instrumental value by helping scientists make predictions and control their environment. In this paper I raise, and provide a first analysis of, the question whether explanatory proofs in mathematics have analogous instrumental value. I first identify an important goal in mathematical practice: reusing resources from existing proofs to solve new problems. I then consider the more specific question: do explanatory proofs have instrumental value by promoting reuse of the resources they contain? In general, I argue that the answer to this question is "no" and demonstrate this in detail for the theory of mathematical explanation developed by Marc Lange.

Contents

1 Introduction

Philosophers of science have investigated the nature of scientific explanations. As part of this investigation, they have tackled the issue of their value. In addition to any intrinsic value, scientific explanations are generally recognized as having important instrumental value. More precisely, they are recognized as helping scientists to make predictions and

control the environment (see e.g. Hempel and Oppenheim (1948), Salmon (1989), Douglas (2009), Lombrozo (2011)).

More recently, philosophers have begun to ask whether there are explanations within mathematics. Most philosophers[1] who have explored the topic conclude that there are. Moreover, a variety of contrasting accounts of mathematical explanation have been proposed (see e.g. Steiner (1978), Kitcher (1989) and Lange (2014)). However, the question of the value of mathematical explanation has received little attention.[2]

In this paper, I assume that there are mathematical explanations and focus on the question of their value. I consider only mathematical explanations of mathematical facts[3] in the form of proofs. By "proof" I mean the written form of a mathematical argument that appears e.g. in a journal article or textbook. My aim is to begin an investigation into whether mathematical explanations help mathematicians to achieve important goals. If they do, then mathematical explanations have instrumental value in an analogous way to scientific explanations, which help scientists achieve the important goals of prediction and control. If they do not, then this is an interesting difference between scientific and mathematical explanation.

The first step to determining whether mathematical explanations have instrumental value in an analogous way to scientific ones is to identify important mathematical goals. I examine mathematical practice and identify one such goal: reusing resources from existing proofs to solve new problems. This is an important goal because achieving it means that mathematicians can solve more problems and do so more efficiently than they could otherwise. I then provide a preliminary analysis of proofs that help mathematicians to reuse their resources in order to identify general features common to such proofs. One important general feature that such proofs share is good mathematical design and presentation.

Having identified an important goal for mathematicians, I then address the issue of whether explanatory proofs help them to achieve it. As explanatory proofs are not usually taken to be sensitive to issues of design and presentation, I suggest that explanatory proofs should not in general be expected to help mathematicians achieve the goal of reusing proof resources. I then illustrate this in detail for the specific theory of explanatory proofs developed by Marc Lange. Explanatory proofs thus fail to have instrumental value by helping mathematicians to achieve the goal of reusing resources.

This leaves open the possibility, however, that there are other important goals which explanatory proofs help mathematicians to achieve. If that turns out to be the case, then mathematical explanations would have instrumental value in an analogous way to scientific explanations. Thus further investigation into mathematical goals and their relationship to explanatory proofs is needed to fully determine whether mathematical explanations have instrumental value. Finally the importance of the goal of reusing mathematical resources

[1] Zelcer (2013) is an exception.

[2] Weber and Verhoeven (2002) briefly discuss the value of mathematical explanations, focusing on unification. Mancosu (2008) has addressed the *philosophical* importance of mathematical explanations, highlighting the significance of this topic not only to philosophy of mathematics and science, but also to metaphysics and epistemology more broadly. My focus here, however, will be on the significance of mathematical explanations to *mathematics* and *mathematicians*. In other words, I will be concerned with the extent to which such explanations may be of benefit to mathematics and the mathematical community.

[3] I will not consider mathematical explanations of scientific facts. See Mancosu (2008) for a discussion of this terminology.

calls for a more thorough analysis into proofs that help mathematicians to achieve it.

2 A Goal from Mathematical Practice

2.1 Introduction

Prediction and control are often cited as two central scientific goals (see e.g. Hempel and Oppenheim (1948), Salmon (1989), Strevens (2006) Lombrozo (2011)). In searching for mathematical goals, then, a natural approach would be to consider whether prediction and control, suitably adapted, also apply to mathematics. I do not take this approach, however, because it is not clear that the adapted goals would be as important to the mathematical community as the originals are for the scientific community. Instead, I examine mathematical practice and identify a goal that is explicitly highlighted as important to members of the mathematical community.

As a starting point, consider the following exchange between an interviewer and a research mathematician:

> I. What do you hope to gain out of reading these proofs [published by other mathematicians]?
>
> M5. As a researcher, I want to understand the idea of the proof and to see if that idea could be applied elsewhere.
>
> I: The point that you made about ideas is something that I've been hearing from your colleagues too. Can you elaborate on that?
>
> M5: Sure. Sometimes when a mathematician answers a hard question, he has a new way of looking at the problem or a new way of thinking about it. As a researcher, when you see this, sometimes you can use this idea to solve problems that you are working on. (Weber, 2010, 34)

Similar themes are reflected in the comments of a variety of mathematicians, a selection of which are presented below:

> Proofs are for the mathematician what experimental procedures are for the experimental scientist: in studying them one learns of new ideas, new concepts, new strategies—devices which can be assimilated for one's own research and be further developed. (Rav, 1999, 20)

> The value of a proof of an outstanding conjecture should be judged, not by its cleverness and elegance, and not even by its "explanatory power," but by the extent in which it enlarges our toolbox. (Bressoud, 1999, 190)

> The point of the proof of Fermat's last theorem is to open up new possibilities for mathematics ... [Wiles and his collaborators] develop a host of new techniques that will lead to further connections between number theory and algebraic geometry. Mathematicians of the future will benefit from Wiles's lead. They will find untold new applications of these newly found methods to the solution

> of other problems, even to problems of great practical interest. The value of Wiles's proof lies not in what it proves, but in what it opens up, in what it makes possible. (Rota, 1997, 190–191)

> [W]henever I would learn something new, or discover something new, I would always be asking myself, "Is there any way this can be useful in my own work?" Or, "Is there any way these results can be extended even further?" This is something that almost all mathematicians do, although perhaps some more than others. (Lady, n.d.)

The core theme within these comments is this: mathematicians aim to *reuse* the ideas, techniques, concepts, lemmas etc that they learn from reading proofs written by their colleagues to make progress on other problems. I will refer to the ideas, techniques, concepts, lemmas etc contained within a proof as *proof resources*. Thus one mathematical goal is to reuse mathematical resources from existing proofs to solve new problems.

I will not here develop a full account of proof "resources" or the manner in which they can be reused, as to do so would require an entire paper by itself. Instead I will provide an illustrative example of reuse, make some remarks about the features of proofs that help mathematicians to reuse their resources and then finally discuss the value of reuse. This will be sufficient to address the question that is the focus of section (3): do explanatory proofs have instrumental value by promoting reuse of their resources?

2.2 Example: Reuse of Techniques

In order to illustrate how a proof can promote reuse of its resources, I present here an example from elementary real analysis.[4] Recall that to say a sequence (s_n) converges means that for all $\epsilon > 0$ there is some N such that for all $n \geq N$ $|s_n - l| < \epsilon$. Recall also the triangle inequality, which states that for all $x, y \in \mathbb{R}$, $|x + y| \leq |x| + |y|$. Now let's examine the standard proof that every convergent sequence has a unique limit (see e.g. Sutherland (1975, 8)):

Proof. Let (s_n) be a sequence and suppose that it converges to l and l', where $l \neq l'$. Then let $\epsilon = \frac{|l-l'|}{2} > 0$. As (s_n) converges to l there is N_1 s.t. for all $n \geq N_1$ we have $|s_n - l| < \epsilon$. Similarly, there is N_2 such that for all $n \geq N_2$ we have $|s_n - l'| < \epsilon$. Put $N = \max(N_1, N_2)$ and then we have

$$
\begin{aligned}
|l - l'| &= |l - s_N + s_N - l'| && (1)\\
&\leq |l - s_N| + |l' - s_N| && (2)\\
&< 2\epsilon = |l - l'| && (3)
\end{aligned}
$$

Contradiction. □

For a mathematical agent who is not familiar with much analysis, this proof can promote reuse of a surprisingly useful technique: adding and subtracting the same term to an expression (see line (1)). It does this not just by exhibiting that adding and subtracting the

[4] For a summary of the basic concepts, see e.g. Sutherland (1975, Ch 1).

same term leads to a solution, but also by providing information about when it is useful to try implementing this technique and how to use it. More precisely, it shows that it is useful to try implementing the technique when you want to find an upper bound for an absolute value (in this case $|l - l'|$). It also illustrates how to use it: by carefully selecting the value of the term added and subtracted so that we obtain absolute values (via the triangle inequality) whose upper bounds are known and can be easily manipulated (see lines (2) and (3)).

After reading the proof, a mathematical agent could extract the technique by formulating it in a more general, semi-abstract way. For example, she could extract it and represent it semi-abstractly as follows:

Technique of adding and subtracting the same term
Aim: To find an upper bound for the expression $|a - b|$ for real a, b.
Procedure: Choose real q such that $|a - q| < l_1$, $|b - q| < l_2$ where l_1 and l_2 are known and easily manipulated.
Output: $|a - b| = |a - q + q - b| \leq |a - q| + |b - q| < l_1 + l_2$.

Notice that the information about how and when to use the technique of "adding and subtracting the same term" is needed to extract it from the proof and formulate it in this semi-abstract manner. This information is crucial for successful reuse, because otherwise the mathematical agent would not know under what circumstances to use it (i.e. when the aim is to find an upper bound for the expression $|a - b|$) or what she must do to apply it (i.e. the details of the procedure).

Having acquired this new technique, a mathematical agent can apply it to establish many other results. For example, suppose such an agent wants to prove the following theorem: every convergent sequence is a Cauchy sequence, i.e. if (s_n) converges, then for all $\epsilon > 0$ there is an N such that for all $m, n \geq N$, $|s_n - s_m| < \epsilon$.

When trying to prove this, our agent should set $\epsilon > 0$ and note that she's looking for an integer N, dependent on ϵ, such that, for all $m, n \geq N$, $|s_n - s_m| < \epsilon$. Our agent is thus in a situation where she wants to find an upper bound for an absolute value, so the technique for adding and subtracting the same term could be useful. The procedure tells her that she needs to find a real q such that $|s_n - q|$ and $|s_m - q|$ have known, easily manipulable upper bounds. As (s_n) is convergent, the agent knows that there is some integer M such that for all $m, n > M$, $|s_n - l| < \epsilon$ and $|s_m - l| < \epsilon$. Thus she may try to apply the technique and find that, for all $m, n \geq M$:

$$\begin{aligned} |s_n - s_m| &= |s_n - l + l - s_m| \\ &\leq |s_n - l| + |l - s_m| \\ &< \epsilon + \epsilon = 2\epsilon \end{aligned}$$

This doesn't quite give our agent what she wanted, as she ended up with a bound of 2ϵ instead of just ϵ. But this is easily fixed. As $\epsilon > 0$, so is $\epsilon/2$. And as (s_n) is convergent, this means that there is some M' such that for all $n \geq M'$, $|s_n - l| < \frac{\epsilon}{2}$. So if our agent tweaks her above reasoning, she will have found an $N(=M')$ such that for all $m, n \geq N$, $|s_n - s_m| < \epsilon$. Writing this out fully, our agent obtains the following proof (see e.g. Sutherland (1975, 9)):

Proof. Let (s_n) converge to l and let $\epsilon > 0$. Then as (s_n) is convergent, there is an N such that for all $n \geq N$ $|s_n - l| < \frac{\epsilon}{2}$. So for any $m, n \geq N$ we have

$$\begin{aligned} |s_n - s_m| &= |s_n - l + l - s_m| \\ &\leq |s_n - l| + |l - s_m| \\ &< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon \end{aligned}$$

Thus (s_n) is a Cauchy sequence. □

In summary, the first proof can help a mathematical agent who is not familiar with analysis to add a new technique to her mathematical arsenal and reuse it appropriately to prove a new theorem.

2.3 General Features of Proofs that Promote Reuse

From the example in section (2.2), we see that proofs which promote reuse of the resources they contain, or, for short, *proofs that promote reuse*, have a number of features. First, they contain at least one resource that the agent is not familiar with, or use a familiar resource in an unfamiliar way. In the example from section (2.2), the technique of adding and subtracting the same term to an expression was unfamiliar to the agent who was assumed to be inexperienced in real analysis.

Second, proofs that promote reuse must "make it easy" for the agent to extract the resource from the proof by formulating it in a more general, semi-abstract way, as it is this formulation that makes the component amenable to use in other circumstances. To do this successfully, the agent first of all needs to be able to identify the resource as it is instantiated in the proof. This means that the reader needs to be able to identify where the use of this resource begins and ends, and where it interacts with other resources, for example.

However, identifying the resource as it is instantiated in the proof is not sufficient to extract it. The mathematical agent also needs information about *how* and *when* to use it. For example, the semi-abstract formulation of the technique of adding and subtracting the same term involves a "procedure" and an "aim" which need to be filled in with information regarding how and when to use the technique. Thus "making it easy" for a reader to extract a resource involves "making it easy" for a reader (i) to identify the resource as instantiated in the proof; and (ii) to identify how and when to use it.

Having identified that proofs that promote reuse both contain an unfamiliar resource and "make it easy" for a reader to extract that resource, I will now focus only on this second feature. As we have seen, this involves "making it easy" both to identify the resource as it appears in the proof and to identify how and when to use it more generally. While I will not offer a full analysis of what it means to "make it easy" for an agent reading a proof to successfully achieve these tasks, I will offer some general considerations.

Importantly, "making it easy" for a mathematical agent to identify a resource and how and when to use it does not mean that the proof must include an explicit statement of this information. Certainly, the real analysis proof did not do this. Rather, that proof made it easy for the agent to find the information needed to successfully extract the resource because the structure of the argument was clear. The structure of the argument was clear

in part because it was a relatively simple example, but also because the proof itself was *well designed* and *presented*. In particular, the proof *managed information* sufficiently well by presenting the agent reading it with relevant information and not distracting her with irrelevant information. This is a very simple idea, which is also powerful and important.

The way proofs manage information becomes particularly pressing when we move beyond elementary examples. This is because mathematical agents have limited cognitive resources, and even relatively small inefficiencies in how information is managed can quickly multiply in a moderately complex proof. The result of such inefficiencies is an unwieldy proof that requires a mathematical agent to expend more cognitive resources than necessary simply to parse it, let alone complete more demanding tasks such as extracting its resources. Agents reading such a proof may thus be unable to extract its resources because the cognitive cost required to do so is just too high. On the other hand, a proof that manages information well will keep the cognitive resources that an agent must spend to extract its resources low. Thus a mathematical agent will be more likely to successfully extract the resources from such a proof, so that they are at her disposal and can potentially be applied elsewhere. Consequently proofs that manage information well tend to promote reuse, while proofs that do not tend not to.

This connection between reuse and information management raises the question: how should a proof be designed so as to manage information well? Jeremy Avigad has pointed out that *modular* mathematics manages information well. He also argues that mathematics designed in this way has a number of benefits, including increased "comprehensibility, reliability, flexibility, and *reuse*" (Avigad, 2018, 12, emphasis added). In rough terms a piece of mathematics is "modular when it can be decomposed into smaller . . . resources, with limited or controlled interactions between them" (Avigad, 2018, 1). This means that information is confined to the parts where it is relevant and thus "hidden" from places where it is not needed.

In a proof, this kind of "information hiding" can occur at both a local and a global level. To see an example of both types, let's look at an example from elementary number theory: Fermat's Little Theorem. This theorem states the following: if p is a prime and m is an integer that is not divisible by p then $m^{(p-1)} - 1$ is evenly divisible by p. If we expand the language of elementary number theory by adding congruence notation, we can express the theorem more concisely. If c is a positive integer then we say that integer a is *congruent* to integer b *modulo* c, i.e. $a \equiv b \pmod{c}$, if and only if $a - b$ is divisible by c. Using this terminology, we can state Fermat's Little Theorem as follows: if p is a prime and m is an integer that is not divisible by p then $m^{(p-1)} \equiv 1 \pmod{p}$.

Two proofs of Fermat's Little Theorem are given below. The first follows closely the presentation by James Ivory (1806) and fails to be modular. The second follows closely the modern, modular presentation given by Thomas Koshy (2007).

Ivory (1806)

Take all the multiples of m by the numbers in the series 1, 2, 3, 4,. . . , $(p-1)$, i.e. m, $2m$, $3m$, $4m$, . . . , $(p-1)m$. None of these multiples will be divisible by p. For suppose $a \cdot m$ is one of the multiples. Then p does not divide m, by assumption, and $a < p$, so p does not divide a either. Hence p does not divide $a \cdot m$. Now consider the remainders of these

multiples when divided by p. No two remainders will be the same. For suppose that $a \cdot m$ and $b \cdot m$ are two multiples which, when divided by p, have the same remainder r. Then $a \cdot m - r$ and $b \cdot m - r$ will both be multiples of p. Consequently $a \cdot m - b \cdot m = (a - b) \cdot m$ will also be a multiple of p. But then either $a - b$ or m must be divisible by p. But this cannot be the case, since $a - b$ is less than p and m is not divisible by p by assumption.

The remainders must thus be the numbers $1, 2, 3, 4 \ldots, (p-1)$ in some order. This means we can represent the multiples $m, 2m, 3m, 4m, \ldots, (p-1)m$ as $a_1p + 1$, $a_2p + 2$, $a_3p + 3$, $a_4p + 4$, $\ldots a_{p-1}p + (p-1)$ in some order, where $a_i, 1 \leq i \leq p-1$ are integers.

We thus have

$$\begin{aligned} & m \cdot 2m \cdot 3m \cdot 4m \cdot \ldots \cdot (p-1)m \\ & = (a_1p + 1)(a_2p + 2)(a_3p + 3)(a_4p + 4) \ldots (a_{p-1}p + (p-1)). \end{aligned}$$

If we expand the expression on the right hand side of this equation, we will find that all of the terms, except the last, contain p or a power of p as a factor. The last term will be $(p-1)!$. Thus we can rewrite the right hand side as $K \cdot p + (p-1)!$, where K is some integer. The left hand side, on the other hand, can be written as $(p-1)!m^{p-1}$. Thus we have $(p-1)!m^{p-1} = K \cdot p + (p-1)!$. Rearranging, we find that $(p-1)!(m^{p-1} - 1) = K \cdot p$. Consequently p must divide either $(p-1)!$ or $m^{p-1} - 1$. However, p does not divide $(p-1)!$. Hence p must divide $m^{p-1} - 1$. $\square$

Koshy (2007)

Definition: Least Residue The least residue of an integer a modulo b is its remainder when it is divided by b.

Lemma: Cancellation Law If $ac \equiv bc \pmod{d}$ and $(c, d) = 1$ then $a \equiv b \pmod{d}$.

Proof of Cancellation Law: Suppose $ac \equiv bc \pmod{d}$ and $(c, d) = 1$. Then d divides $ac - bc$, i.e. d divides $c(a - b)$. But $(c, d) = 1$, so d divides $a - b$, i.e $a \equiv b \pmod{d}$. $\square$

Lemma: Permutation If p is a prime and a any integer such that p does not divide a, then the least residues of the integers $a, 2a, 3a, \ldots, (p-1)a$ modulo p are a permutation of the integers $1, 2, 3, \ldots, (p-1)$.

Proof of Permutation Lemma: We will prove this by first proving that $ia \not\equiv 0 \pmod{p}$ for all $1 \leq i \leq p-1$. We will then prove that the least residues of ia and ja are distinct if i and j are distinct.

First we show that $ia \not\equiv 0 \pmod{p}$ for all $1 \leq i \leq p-1$. By assumption, p does not divide a. As $i < p$, p does not divide i. Thus p cannot divide ia. Hence $ia \not\equiv 0 \pmod{p}$.

Now we show that the least residues of ia and ja are distinct if i and j are distinct. Suppose that the least residues of ia and ja are congruent. Then $ia \equiv ja \pmod{p}$. As p does not divide a, $(p, a) = 1$ and so by cancellation, $i \equiv j \pmod{p}$. But i and j are least residues modulo p and thus must be equal. $\square$

Proof of Fermat's Little Theorem: By the permutation lemma, the least residues of $m, 2m, 3m, 4m, \ldots, (p-1)m$ modulo p are the same as the integers $1, 2, 3, 4, \ldots, p-1$ in some order. Thus their products are congruent modulo p:

$$m \cdot 2m \cdot 3m \cdot 4m \cdot \ldots \cdot (p-1) \cdot m \equiv 1 \cdot 2 \cdot 3 \cdot 4 \cdot \ldots \cdot (p-1) \pmod{p}.$$

Simplifying, we thus have:

$$(p-1)!m^{p-1} \equiv (p-1)! \pmod{p}.$$

As $((p-1)!, p) = 1$, we can apply the cancellation law and so obtain $m^{p-1} \equiv 1 \pmod{p}$. $\square$

Koshy's proof is far more modular than Ivory's and engages in information hiding at the local and global level. First, at the local level, Ivory's proof introduces terms for quotients of certain numbers modulo p. However, we do not need to know the value of these quotients (all we need to know is the remainder upon division by p), so his proof introduces terms for information that is irrelevant. Koshy's proof, on the other hand, uses congruence notation which allows him to avoid introducing terms for the quotients. He thus hides irrelevant information that Ivory displays.

At a more global level, Koshy's proof is broken up into a series of self-contained parts: a definition and two lemmas are separated out from the main proof of Fermat's Little Theorem. The main body of Ivory's proof, on the other hand, includes a version of the permutation lemma and its proof without separating them out. Koshy's proof thus hides the details involved in proving the permutation lemma from the main body of the proof of Fermat's Little Theorem whereas Ivory does not. Additionally, by formulating the permutation lemma as a lemma, Koshy is drawing attention to it as an independent proposition expressed in general terms. This makes it easier for an agent reading the proof to extract it and add it to her toolbox, because Koshy has done some of the work for the reader by explicitly identifying it as a resource that can be reused.

The comparison between Ivory's and Koshy's proofs highlights an important point. When we are trying to design a proof that manages information well, we need to pay careful attention to the resources we choose. This is because some resources manage information better than others, even if they are "equivalent." For example, part of what makes Koshy's proof manage information better than Ivory's is the former's use of congruence notation, even though congruence notation is just a definitional extension of number theory. Given the connection between information management and reuse, this means that it's possible for there to be a proof P utilizing resources $c_1, c_2, \ldots, c_n$ that fails to promote reuse but which could nonetheless be rewritten utilizing equivalent resources $c_1', c_2', \ldots, c_n'$ to yield a proof P' which does promote reuse.

In summary, the relationships between reuse, extracting proof resources, information management, and design and presentation are as follows. 1. Extracting resources is a necessary condition of reuse. After all, to reuse a proof resource, a mathematical agent

needs to add it to her toolbox so that it is at her disposal and ready to be applied to solve other problems. 2. The manner in which a proof manages information can make it easier or harder for an agent to extract proof resources by decreasing or increasing the cognitive cost of doing so. 3. The way a proof is designed and presented can significantly affect how efficiently it manages information. Consequently, good design and presentation, as exemplified by modular mathematics, tend to promote reuse, as they make it more likely that mathematical agents can extract the resources and thus have them at their disposal, ready to apply elsewhere. On the other hand, bad design and presentation tend not to promote reuse, as they make it less likely that mathematical agents reading such a proof will even be able to extract its resources.

These observations also allow us to clarify the relationship between *reuse* and *unification*. For while these are closely related concepts, they are not identical. If we are concerned with unification, then we want proofs whose underlying argument-schemes prove a variety of other theorems. In other words, we want to minimize the number of underlying argument-schemes while maximizing the number of theorems proved using them. Unification therefore has a quantitative nature (see e.g. Hafner and Mancosu (2008, 170)). Reuse, however, is more qualitative and agent-focused. If we are concerned with promoting reuse, we want our proofs to manage information well so that it is easy for finite agents to extract their resources. So while both someone concerned with unification and someone concerned with reuse would find Koshy's proof appealing, they would do so for different reasons. Koshy's approach of breaking out generally applicable lemmas from the overall proof is attractive to someone concerned with unification because doing so helps with the project of finding a small number of argument-schemes that prove a large number of theorems. On the other hand, the approach of breaking out lemmas will be attractive to someone concerned with promoting reuse because it hides information and makes it easier for finite mathematical agents to extract the proof resources.

For a more extreme example that highlights the difference between unification and reuse, consider a case study analyzed by Hafner and Mancosu (2008): the theory of Real Closed Fields. Hafner and Mancosu point out that, due to the existence of a decision procedure for this theory, all theorems can be obtained from a single underlying argument-scheme. Thus someone who values unification would value proofs that instantiate this scheme. However, someone who is concerned with promoting reuse would not care for such proofs because they manage information poorly. For example, Hafner and Mancosu quote the mathematician Brumfiel who notes that, in practice, it can be "very tedious, if not physically impossible, to work out" (Hafner and Mancosu, 2008, Brumfiel quoted on pg 159) the details of proofs instantiating this argument-scheme. Moreover, Hafner and Mancosu themselves observe that the argument-scheme is "hardly ever used at all by working mathematicians because of the limited feasibility of the decision algorithm" (Hafner and Mancosu, 2008, 166).

The brief analysis presented in this section is admittedly far from complete. However, it is sufficient to highlight the important connection between mathematical design and presentation, specifically in the form of information management, and whether a proof promotes reuse. This will be important when we come to consider whether explanatory proofs have instrumental value by promoting reuse of their resources.

2.4 Value of Proofs that Promote Reuse

Proofs that promote reuse allow mathematical agents who read them to add new resources to their "mathematical toolbox," i.e. the collection of ideas, techniques, concepts, lemmas etc that they are familiar with and can use appropriately. The result is that, after reading a proof that promotes reuse, a mathematical agent's toolbox increases. She thus has more tools at her disposal when working on her own problems and this means that it is more likely that she will be able to solve them successfully and efficiently.

More precisely, suppose that our mathematical agent is working on a problem that can be solved by existing resource t. Further, let the agent's initial toolbox be T_1 and suppose that, after reading a variety of proofs that promote reuse, her toolbox is expanded to $T_2 \supsetneq T_1$. As T_2 is strictly larger than T_1, it is more likely that t is contained within T_2 than T_1. Assuming that the agent can solve her problem if and only if t is available in her toolbox, she is then more likely to solve her problem after reading the proofs that promote reuse than before.

Now suppose that our agent has toolbox T_1 and that $t \notin T_1$. Suppose we relax the assumption that the agent can solve her problem if and only if t is available in her toolbox. More precisely, assume that the agent can solve her problem by reinventing t for herself. Inventing new mathematical resources takes considerable time and effort, so if our agent reinvents t in order to solve her problem, it will plausibly take her longer to find the solution than if t were already available to her. So there are now two possibilities for our mathematical agent: (i) she is unable to reinvent t for herself and thus fails to solve her problem; (ii) she reinvents t for herself and solves her problem, but it takes her much longer to do so than if t were already available in T_1. Either option is worse than if resource t were available in her toolbox.

Consequently, reuse of proof resources enables mathematicians to successfully and more efficiently solve new problems. In this way, it contributes to the advancement of mathematical research. The mathematical community is therefore right to value reuse and proofs that promote it.

2.5 Mathematical vs. Scientific Goals

From our brief analysis of reuse, we can already see that it differs significantly from the scientific goals of prediction and control. The latter are externally directed, as they concern phenomena in the world, not just the scientific theory. Reuse, on the other hand, is internally directed, as it is focused on mathematical resources.[5] Moreover, the discussion of reuse, how to achieve it, and its importance has focused heavily on mathematical agents and, in particular, on making it easier for them to solve problems and do so efficiently. This is in contrast to the goals of scientific explanation, which are not taken to make it easier to do science.[6]

If reuse is so different from the scientific goals of prediction and control, is it an appropriate goal to focus on when investigating the instrumental value of explanatory proofs? I

[5] I am grateful to an anonymous referee for pointing out this difference between reuse and prediction and control.

[6] I am grateful to an anonymous referee for pointing out this difference between reuse and explanatory goals in science.

suggest that it is for two reasons. First, reuse is, as I have argued, an important goal for mathematicians, just like prediction and control are important goals for scientists. If mathematical explanations promote attainment of important mathematical goals, then they are analogous to scientific explanations which are already known to promote attainment of important scientific goals. If they don't, then this is an interesting difference between scientific and mathematical explanation. Second, the differences between reuse and prediction and control can plausibly be accounted for by differences between mathematics and science. This means that, while different from prediction and control, reuse is not a strange *mathematical* goal. To see this, first note that pure mathematics is abstract, with no direct connection to the world. As such, it is not surprising that mathematical goals, unlike prediction and control, are internally focused and concerned with improving mathematics itself. And one legitimate dimension along which mathematics can be improved is by making it easier for agents to do mathematics.

Thus while reuse differs from scientific goals like prediction and control, it is nonetheless a legitimate goal to consider when assessing whether explanatory proofs have instrumental value.

3 Explanatory Proofs and Reuse

3.1 Introductory Remarks

Having sketched an account of the goal of reusing proof resources to solve new problems in section (2), we can now address the question whether explanatory proofs help mathematicians to attain this goal. Philosophers and mathematicians have strong and conflicting intuitions about which proofs are explanatory (see e.g. Lange (2009)), so while I will begin by making some tentative general remarks, I will then address the question relative to a specific account of explanatory proofs. Here I choose to focus on the account of explanation developed by Marc Lange (2014), though theories of mathematical explanation have also been proposed by Mark Steiner (1978)[7] and Philip Kitcher (1989).[8] The question whether explanatory proofs, according to these accounts, help mathematicians to reuse their resources should be fully explored.

[7]On Steiner's account, "an explanatory proof depends on a characterizing property of something mentioned in the theorem: if we 'deform' the proof, substituting the characterizing property of a related entity, we get a related theorem" (Steiner, 1978, 147). This seems to suggest that explanatory proofs, on Steiner's account, will promote reuse, since the proof resources are reused when the proof is deformed to obtain new theorems. However, the situation is not clear cut (two anonymous referees had conflicting views, for example). It seems plausible, for instance, for a proof to be explanatory on Steiner's account and yet badly presented or designed, making it difficult for mathematical agents to obtain the deformations, i.e. to reuse the resources in practice. In other words, it seems plausible for there to be explanatory proofs, on Steiner's account, which are deformable in principle but not in practice. If that is correct, then explanatory proofs, on Steiner's account, can't be said to promote reuse after all.

[8]Kitcher's account of explanation is a unificationist one. As we have seen in section (2.3), unification has a quantitative nature, while reuse is more qualitative and agent-focused. Thus while Kitcher's account of explanation may at first seem to have a strong connection to reuse, under closer inspection any such connection appears much weaker. In fact, Hafner and Mancosu (2008) argue that, according to his account, proofs that instantiate the decision procedure argument-scheme discussed in section (2.3) are explanatory. However, we have already seen that such proofs manage information inefficiently and thus fail to promote reuse.

First, the tentative, general remarks. Explanations are not usually taken to be sensitive to issues of design and presentation. For example, Zelcer notes "Mathematical explanations are said to be ... not mere stylistic features that communicate mathematics more clearly or in a psychologically more satisfying or pedagogically more useful way" (Zelcer, 2013, 176). Yet mathematical style and communication fall under the umbrella of mathematical design and presentation and thus *do* affect whether a proof successfully promotes reuse of its resources. For example, we saw in section (2.3) that an abbreviation introduced to number theory, congruence notation, allowed Koshy to suppress irrelevant information from his proof of Fermat's Little Theorem and made it more likely to promote reuse than Ivory's proof which did not make use of the abbreviation.

If mathematical explanations are not sensitive to issues of design and presentation but these issues affect whether a proof promotes reuse, then there is little reason to expect explanatory proofs to promote reuse of their resources. Below we will see a concrete example of a proof that is explanatory on Lange's account but which fails to promote reuse.

3.2 Lange's Account

Lange offers the following explication of an explanatory proof "What it means to ask for a proof that explains is to ask for a proof that exploits a certain kind of feature in the setup—the same kind of feature that is salient in the result. The distinction between proofs that explain why some theorem holds and proofs that merely establish that it holds exists only when some feature of the result being proved is salient"(Lange, 2014, 507). He notes that, while there are a wide variety of features that may strike us as salient, unity, symmetry and simplicity are common such features in mathematics.

To see how Lange's account works in practice, let's examine one of his examples: The Calculator Number Theorem (Lange, 2014, 488, 508–509). A *calculator number* is formed by taking the three digits from a row, column or main diagonal of a calculator and then reversing them, for example, 147741. It is a theorem that *all* calculator numbers are divisible by 37. This result has a surprising unity, being true for all such numbers. As its unity is salient, an explanatory proof of this theorem is one which "... proceeds from a property common to each of these numbers ... and that is common to them precisely because they are calculator numbers" (Lange, 2014, 508). Consider the following proof of the above theorem (Lange, 2014, 488):

Proof. The three digits from which a calculator number is formed are three integers a, $a+d$, $a+2d$ in arithmetic progression. Take any number formed from three such integers in the manner of a calculator number—that is, any number of the form $10^5a+10^4(a+d)+10^3(a+2d)+10^2(a+2d)+10(a+d)+a$. Regrouping, we find this equal to $a(10^5+10^4+10^3+10^2+10+1)+d(10^4+2.10^3+2.10^2+10) = 111111a+12210d = 1221(91a+10d) = (3\times 11\times 37)(91a+10d)$. □

This proof is explanatory, on Lange's account, because "... it traces the fact that every calculator number is divisible by 37 to a property that they have in common by virtue of being calculator numbers [i.e. that they can be represented in the form $10^5a+10^4(a+d)+10^3(a+2d)+10^2(a+2d)+10(a+d)+a$]. In short, an explanation of this result consists of a proof that treats every calculator number in the same way" (Lange, 2014, 509).

3.3 Reuse on Lange's Account

Having sketched Lange's account of explanatory proofs, we can now ask whether proofs that are explanatory on his account tend to promote reuse. The answer to this question is "no."

As we have seen, mathematical design and presentation affect whether a proof promotes reuse. However, there is nothing in Lange's characterization of explanatory proofs that takes these issues into account. Thus there is nothing in the nature of such proofs that prevent them from managing information inefficiently and failing to promote reuse. I will now illustrate this with a concrete example: another proof of Fermat's Little Theorem.

Recall from earlier that Fermat's Little Theorem states that if p is a prime then all integers m which are not divisible by p are such that $m^{p-1} - 1$ is evenly divisible by p. Alternatively, using congruence notation, if p is a prime then all integers m that are not divisible by p are such that $m^{p-1} \equiv 1 \pmod{p}$. This result is similar to the Calculator Number Theorem in that its unity is surprising, i.e. it is striking that it holds for *all* integers that are not divisible by p. As its unity is salient, we can expect an explanatory proof to be one which establishes the result by treating each number that is not divisible by p in the same way. More precisely, such a proof will work by identifying a common property of these numbers and use it to derive the result.

I present one such proof below, which is in the style of Joseph-Louis Lagrange (1773).[9] However, I give you fair warning that the proof manages information poorly, which makes it painfully difficult to read! After presenting all of the gory details, I will point out in more precise terms how it manages information poorly and provide a concrete example of how it fails to promote reuse. Finally, I will spell out why it is nonetheless explanatory on Lange's account.

Proof. Let m be an integer that is not divisible by p. Then $m = \mu p + \rho$ where ρ is an integer such that $1 \leq \rho \leq p - 1$. Thus one of $(m + 1), (m + 2), \ldots, (m + p - 1)$ must be divisible by p. Consequently the product $(m+1)(m+2)\ldots(m+p-1)$ will be divisible by p.

Now consider the polynomial $(x + 1)(x + 2)\ldots(x + p - 1)$ and expand it:

$$\begin{aligned}&(x + 1)(x + 2)\ldots(x + p - 1)\\ &= x^{p-1} + A'x^{p-2} + A''x^{p-3} + \ldots + A^{(p-1)}.\end{aligned}$$

Substituting $x + 1$ for x yields

$$\begin{aligned}&(x + 2)(x + 3)\ldots(x + p)\\ &= (x + 1)^{p-1} + A'(x + 1)^{p-2} + A''(x + 1)^{p-3} + \ldots + A^{(p-1)}.\end{aligned}$$

Multiply the entire first equation by $(x + p)$ and the entire second by $(x + 1)$ to obtain:

[9] While this proof is in the style of Lagrange, I have modified it to make it manage information even less efficiently than the original.

$$
\begin{aligned}
&(x+p)(x^{p-1}+A'x^{p-2}+A''x^{p-3}+\ldots+A^{(p-1)})\\
&=(x+1)^p+A'(x+1)^{p-1}+A''(x+1)^{p-2}+\ldots+A^{(p-1)}(x+1).
\end{aligned}
$$

Expand and collect like terms:

$$
\begin{aligned}
&x^p+(p+A')x^{p-1}+(pA'+A'')x^{p-2}+(pA''+A''')x^{p-3}+\ldots\\
&=x^p+(p+A')x^{p-1}+\left[\frac{p(p-1)}{2}+(p-1)A'+A''\right]x^{p-2}\\
&+\left[\frac{p(p-1)(p-2)}{2\cdot 3}+\frac{(p-1)(p-2)}{2}A'+(p-2)A''+A'''\right]x^{p-3}+\ldots
\end{aligned}
$$

Comparing term by term, we obtain a series of equations

$$
\begin{aligned}
&p+A'=p+A'\\
&pA'+A''=\frac{p(p-1)}{2}+(p-1)A'+A''\\
&pA''+A'''=\frac{p(p-1)(p-2)}{2\cdot 3}+\frac{(p-1)(p-2)}{2}A'+(p-2)A''+A'''\\
&\ldots
\end{aligned}
$$

From these equations, we can derive the following:

$$
\begin{aligned}
&A'=\frac{p(p-1)}{2}\\
&2A''=\frac{p(p-1)(p-2)}{2\cdot 3}+\frac{(p-1)(p-2)}{2}A'\\
&3A'''=\frac{p(p-1)(p-2)(p-3)}{2\cdot 3\cdot 4}+\frac{(p-1)(p-2)(p-3)}{2\cdot 3}A'+\frac{(p-2)(p-3)}{2}A''\\
&\ldots
\end{aligned}
$$

From these equations and the fact that $\frac{p(p-1)}{2}, \frac{p(p-1)(p-2)}{2\cdot 3}, \frac{p(p-1)(p-2)(p-3)}{2\cdot 3\cdot 4}, \ldots$ are all evenly divisible by p, we can see that $A', 2A'', \ldots, (p-2)A^{(p-2)}$ are also all divisible by p. Consequently $A', A'', \ldots, A^{(p-2)}$ are all divisible by p.

Now consider $(p-1)A^{(p-1)}$. From the previous set of equations, we have:

$$
\begin{aligned}
(p-1)A^{(p-1)} &= \frac{p(p-1)(p-2)\ldots 1}{1\cdot 2\cdot 3\cdot\ldots\cdot p}\\
&\quad+\frac{(p-1)(p-2)\ldots 1}{1\cdot 2\cdot 3\cdot\ldots\cdot(p-1)}A'+\frac{(p-2)(p-3)\ldots 1}{1\cdot 2\cdot\ldots\cdot(p-2)}A''\\
&\quad+\ldots\\
&=1+A'+A''+A'''+\ldots+A^{(p-2)}.
\end{aligned}
$$

Rearranging we obtain

$$A^{(p-1)} + 1 = pA^{(p-1)} - A' - A'' - A''' - \ldots - A^{(p-2)}$$

As $A', A'', A''', \ldots, A^{(p-2)}$ are all divisible by p, it follows that $A^{(p-1)}+1$ is also divisible by p.

Thus, for any integer x,

$$(x+1)(x+2)(x+3)\ldots(x+p-1) - x^{p-1} + 1$$

is always divisible by p.

Substituting m for x we see that $(m+1)(m+2)(m+3)\ldots(m+p-1) - m^{p-1} + 1$ is evenly divisible by p. As $(m+1)(m+2)(m+3)\ldots(m+p-1)$ is evenly divisible by p, so is $-m^{p-1}+1$. Thus $m^{p-1}-1$ must also be divisible by p and Fermat's Little Theorem is proven. □

The Lagrange style proof of Fermat's Little Theorem does a poor job at managing information. This makes both simply parsing the proof, as well as completing more demanding cognitive tasks, like extracting its resources, more difficult than it has to be.

Let's start by considering how the proof manages information at the global level. Note that the proof is made up of a variety of different parts, as summarized below:

1. As $m = \mu p + \rho$ where $1 \leq \rho \leq p-1$, p evenly divides $(m+1)(m+2)\ldots(m+p-1)$.
2. Lemma: The polynomial coefficients A', A'', …, $A^{(p-1)}$ satisfy a particular recurrence relation.[10]
   (a) Corollary 1: $A', A'', \ldots, A^{(p-2)}$ are all evenly divisible by p. $A^{(p-1)}+1$ is also evenly divisible by p.
   (b) Corollary 2: For any x, $(x+1)(x+2)\ldots(x+p-1) - x^{p-1} + 1$ is evenly divisible by p.
3. As $(m+1)(m+2)\ldots(m+p-1)$ and $(m+1)(m+2)\ldots(m+p-1) - m^{p-1} + 1$ are both evenly divisible by p, $m^{p-1}-1$ must also be evenly divisible by p.

The proof itself, however, does not break out any lemmas or corollaries, nor does it clearly delimit these different parts in any other way.[11] This means that the reader has to undertake a significant amount of cognitive work to extract them herself, which she must do if she wants to reuse them. The proof's information management could be much improved in this regard if, for example, the results relating to the recursive relationship between the polynomial coefficients were broken out as lemmas and/or corollaries, in a similar style to Koshy's proof from section (2.3).

[10] The recurrence relation is given by the equations: $A' = \frac{p(p-1)}{2}, 2A'' = \frac{p(p-1)(p-2)}{2 \cdot 3} + \frac{(p-1)(p-2)}{2} A', \ldots$.

[11] Lagrange (1773) does break out lemmas in his original proof. This is one way in which the information management in the Lagrange style proof I gave above is worse than the information management in the original.

Next, let's consider information management at a local level. The proof fails to hide irrelevant information at this level in at least two ways. First, the proof uses resources expressed in terms of divisibility rather than congruence notation, which makes it more clumsy than necessary.[12] Second, the polynomial coefficients $A', A'', \ldots, A^{(p-1)}$ are kept "on display" throughout the proof, even when they are not strictly needed, and this can make the proof much harder to parse. Consider, for example, the following (paraphrased) part of the proof:

> Multiply
>
> $$\begin{aligned}&(x+1)(x+2)\ldots(x+p-1)\\&= x^{p-1} + A'x^{p-2} + A''x^{p-3} + \ldots + A^{(p-1)}\end{aligned}$$
>
> by $(x+p)$ and
>
> $$\begin{aligned}&(x+2)(x+3)\ldots(x+p)\\&= (x+1)^{p-1} + A'(x+1)^{p-2} + A''(x+1)^{p-3} + \ldots + A^{(p-1)}\end{aligned}$$
>
> by $(x+1)$ to obtain:
>
> $$\begin{aligned}&(x+p)(x^{p-1} + A'x^{p-2} + A''x^{p-3} + \ldots + A^{(p-1)})\\&= (x+1)^p + A'(x+1)^{p-1} + A''(x+1)^{p-2} + \ldots + A^{(p-1)}(x+1)\end{aligned}$$

It takes some effort to verify that what we get by multiplying the first equation by $(x+p)$ and the second by $(x+1)$ are the same, at least when presented like this. It is much clearer, however, when we suppress the coefficients, as in the following presentation:

$$\text{Let } L(x) = (x+1)(x+2)\ldots(x+p-1)$$

Then notice that

$$L(x+1) = (x+2)(x+3)\ldots(x+p) = \frac{(x+p)L(x)}{(x+1)}.$$

Rearranging, we see

$$(x+p)L(x) = (x+1)L(x+1).$$

[12] This is not a criticism of Lagrange. He could not have chosen to use the resources of congruence notation because they were not invented when he wrote his proof.

Thus, when we suppress the coefficients, we see that the equation holds almost immediately. In other words, the presence of the coefficients in the extract from the first presentation served only to distract us from the information that we needed to easily check that the relevant equation holds. This distracting information will also make it harder to complete more cognitively demanding tasks, including extracting proof resources so that they can be reused.

The proof's poor information management ultimately makes it harder than necessary for a reader to extract resources so that they can be reused. In other words, the proof fails to promote reuse. Nonetheless, the proof resources can be reused to prove another number theoretic result: Wilson's Theorem. Wilson's theorem states that if p is prime $(p-1)! \equiv -1 \pmod{p}$, i.e. $(p-1)!+1$ is divisible by p. To really appreciate how the proof fails to promote reuse, it is worth looking at it again and trying to see if you can figure out how to use it to prove Wilson's Theorem.

In fact, to prove Wilson's Theorem you don't need to modify the proof of Fermat's Little Theorem at all—it is proved along the way! When the original proof shows that $A^{(p-1)} + 1$ is evenly divisible by p, it has established Wilson's Theorem, since $A^{(p-1)}$ is $(p-1)!$. This is a very minimal case of reuse, then, since it just amounts to recognizing that the very same proof establishes another result along the way. Yet the proof's poor information management means that it fails to promote even such a minimal form of reuse. More precisely, the poor information management does nothing to help a mathematical agent extract the results related to the recursive relationship between the polynomial coefficients (i.e. the Lemma and Corollaries 1 and 2 from the outline given above) which contain Wilson's Theorem. This theorem thus remains hidden beneath a sea of other information in the original proof and will be hard to spot. If, however, the proof was designed to help an agent extract the relevant lemma and corollaries, it would be much easier to recognize that Wilson's Theorem is proven along the way, since it is contained in Corollary 1.

While its poor information management means that the Lagrange style proof of Fermat's Little Theorem fails to promote reuse, it is nonetheless explanatory on Lange's account. To see this, recall that Fermat's Little Theorem is striking in the same way that the Calculator Number Theorem is striking: it holds for all numbers of the relevant sort. And, just like the proof of the Calculator Number Theorem, the Lagrange style proof traces this unity to a property that the relevant numbers have in common in virtue of the fact that they are all numbers of this sort. In the Calculator Number Theorem, the common property is that they can all be represented in the form $10^5a+10^4(a+d)+10^3(a+2d)+10^2(a+2d)+10(a+d)+a$. In Fermat's Little Theorem, the common property is that they can all be represented in the form $\mu p + \rho$ where $1 \leq \rho \leq p-1$ and so, for each m not divisible by p, $(m+1)(m+2)\ldots(m+p-1)$ is divisible by p. Finally, by judging this proof to be explanatory, Lange's account is in agreement with other mathematicians. For example, Henry Smith explicitly describes Lagrange's proof as explanatory in his *Report on the Theory of Numbers* (Smith, 1894, 47).

Consequently, the proof of Fermat's Little Theorem serves as a concrete example of a proof that is explanatory on Lange's account but which manages information poorly and fails to promote reuse. Moreover, this is not an isolated example, since there is nothing in Lange's account that requires explanatory proofs to be sensitive to issues of mathematical design and presentation, yet these features affect whether a proof promotes reuse. In fact,

given any explanatory proof on Lange's account that happens to promote reuse, we can redesign it to manage information so poorly that it will very likely fail to promote reuse while still preserving its explanatory power. In conclusion, then, explanatory proofs on Lange's account cannot, in general, be said to promote reuse.

4 Concluding Remarks

In this paper, I raised the general question whether mathematical explanations have instrumental value. I then narrowed this to the more specific question whether explanatory proofs promote the attainment of an important goal from mathematical practice: reusing proof resources to solve new problems. I analyzed proofs that promote reuse and showed that they are sensitive to issues of mathematical design and presentation. Explanatory proofs, however, are not taken to be sensitive to the same issues. I thus argued that we should not expect explanatory proofs to promote reuse in general and illustrated this in detail for the case of Lange's account of explanatory proofs.

However, there may be other important mathematical goals. If so, then explanatory proofs may have instrumental value by helping mathematicians to achieve these goals, even though they do not generally tend to promote reuse. Thus an avenue for future work is to examine mathematical practice further and identify other goals so that we can assess whether explanatory proofs help mathematicians to attain them.

Furthermore, in this paper I have focused exclusively on mathematical explanations in the form of proofs. However, mathematical explanations can come in other forms (see e.g. Lange (2014)), so another avenue for future work is to ask whether explanations in these forms have instrumental value.

A final important avenue for future work is a full investigation into proofs that promote reuse. As we have seen, such proofs help advance mathematical research by allowing mathematicians to solve more problems and do so more efficiently than they would otherwise. A full analysis into proofs that promote reuse may thus inform us about the structure of mathematical knowledge and its dissemination. This could, in turn, provide practical advice about the design of future mathematics. That an investigation into such proofs could have important practical consequences is not an unrealistic hope. For example, reusability is a topic that has received attention in software engineering, with design principles being formulated to guide the development of reusable software (see e.g. Anguswamy (2013)).

Ultimately, then, this paper is just the beginning of two related investigations: (i) into the instrumental value of mathematical explanations; (ii) into proofs that promote reuse.

Acknowledgments I am very grateful to Jeremy Avigad, Michael Friedman, Erich Kummerfeld and Wilfried Sieg for helpful feedback on drafts of this paper. I am also grateful to participants at the 2018 Stanford Workshop on Mathematical Reasoning for their helpful questions and discussions on reuse in mathematics. Finally I am grateful to the anonymous reviewers who provided helpful feedback and suggestions.

References

Anguswamy, Reghu. 2013. "Factors Affecting the Design and Use of Reusable resources." PhD diss, Virginia Polytechnic Institute and State University.

Avigad, Jeremy. 2018. "Modularity in mathematics." *The Review of Symbolic Logic*, 1–33.

Bressoud, David. 1999. *Proofs and Confirmations: The Story of the Alternating-Sign Matrix Conjecture*, Cambridge University Press.

Douglas, Heather. 2009. "Reintroducing prediction to explanation." *Philosophy of Science*, 76(4):444–463.

Hafner, Johannes and Mancosu, Paolo. 2008. "Beyond unification." In *The Philosophy of Mathematical Practice*, ed. Mancosu, Paolo, 151–178. Oxford University Press.

Hempel, Carl and Oppenheim, Paul. 1948. "Studies in the logic of explanation." *Philosophy of Science*, 15(2):135–175.

Ivory, James. 1806. Demonstration of a theorem respecting prime numbers In *The Mathematical Repository*, ed. Leybourn, Thomas, 6–8. Gledinning.

Kitcher, Philip. 1998. "Explanatory unification and the causal structure of the world." *Scientific Explanation*, volume 8, ed. Kitcher, Philip and Salmon, Wesley, 410–505. Minneapolis: University of Minnesota Press, 1989.

Koshy, Thomas. 2007. *Elementary Number Theory with Applications*, Elsevier.

Lady, Lee. n.d. How to Do Mathematical Research. URL `http://www.math.hawaii.edu/~lee/how-to.html`. Archived at `https://perma.cc/LZE6-ETTD`

Lagrange, Joseph-Louis. 1773. "Demonstration d'un théorème nouveau concernant les nombres premiers." *Nouveaux Mémoires de l'Académie Royale des Sciences et Belles-Lettres*, 2:125–137.

Lange, Marc. 2014. "Aspects of mathematical explanation: Symmetry, unity, and salience." *Philosophical Review*, 123(4):485–531.

Lange, Marc. 2009. "Why proofs by mathematical induction are generally not explanatory." *Analysis*, 69(2):203–211.

Lombrozo, Tania. 2011. "The instrumental value of explanations." *Philosophy Compass*, 6 (8):539-551.

Mancosu, Paolo. 2008. Mathematical explanation: Why it matters. In *The Philosophy of Mathematical Practice*, ed. Mancosu, Paolo, 134–149. Oxford University Press.

Rav, Yehuda. 1999. "Why do we prove theorems?" *Philosophia Mathematica*, 7(1):5–41.

Rota, Gian-Carlo. 1997. "The phenomenology of mathematical proof." *Synthese*, 111(2): 183–196.

Salmon, Wesley. 1989. "4 decades of scientific explanation." *Minnesota Studies in the Philosophy of Science*, 13:3–219.

Smith, Henry J. S. 1894. *The collected mathematical papers of Henry John Stephen Smith.* Oxford: The Clarendon Press.

Steiner, Mark. 1978. "Mathematical explanation." *Philosophical Studies*, 34(2):135–151.

Strevens, Michael. 2006. Scientific Explanation. In *Encyclopedia of Philosophy* (second edition), ed. Borchert, D. Macmillan Reference.

Sutherland, Wilson. 1975. *Introduction to Metric and Topological Spaces.* Oxford Science Publications.

Weber, Keith. 2010. "Proofs that develop insight." *For the Learning of Mathematics*, 30 (1):32–36.

Weber, Erik, and Liza Verhoeven. 2002. "Explanatory proofs in mathematics." *Logique & Analyse*, 45(179/180):299–307.

Zelcer, Mark. 2013. "Against mathematical explanation." *Journal for General Philosophy of Science*, 44(1):173–192.