# CHARACTERIZING PROPERTIES AND EXPLANATION IN MATHEMATICS

## JOSEPHINE SALVERDA\*

#### Abstract

Mark Steiner proposes one of the earliest contemporary accounts of mathematical explanation, which appeals to characterizing properties of entities referred to in proofs. Unfortunately Steiner's remarks are often quite vague, sometimes described as 'very puzzling indeed', and this lack of clarity has led to a lack of understanding and a tendency to reject Steiner's account in the philosophical literature.

I argue that Steiner's account repays deeper analysis by providing a sympathetic reading that makes sense of his puzzling remarks and draws out some important questions.

I focus on a simple mathematical example involving sums of number sequences and identify three key conditions that the proof must meet to count as explanatory for Steiner. I propose a suitable characterizing property and show that on my suggestion, the proof indeed fits Steiner's account. Subsequently, I present a few potential problems relating to Steiner's focus on the generalizability of proofs, and show how my reading of generalizability helps to avoid these worries.

Finally, I show how (my extension of) Steiner's proposal can account for what I take to be the primary epistemic function of an explanation, namely, to help us see why the fact to be explained is true.

#### 1. Steiner's Account of Mathematical Explanation

All proofs show that their conclusions are true; some also explain why they are true. But what makes a proof (or argument) explanatory, if it is? There has been a recent surge of interest in this topic, reflecting increased philosophical attention to mathematical activity and practice.<sup>1</sup>

The distinction between proofs that show that their conclusion is true and those that explain why it is true is not just something philosophers worry

\* University College London. I would like to thank audiences in Brussels, Cambridge and London for stimulating and helpful discussion. Particular thanks are due to Luke Fenton-Glynn, Alexander Geddes, Marcus Giaquinto, and an anonymous referee for this journal. I was supported by the Arts and Humanities Research Council while writing this paper.

<sup>1</sup> See Mancosu (2008) for an overview and Lange (2014) for a recent proposal. For work on the related but separate topic of mathematical explanation in science, see e.g. Baker (2009).

about. For example, in an interview renowned mathematician and Fields medallist Michael Atiyah recalls:

'I remember one theorem that I proved and yet I really could not see why it was true. It worried me for years ... I kept worrying about it, and five or six years later I understood why it had to be true. Then I got an entirely different proof ... Using quite different techniques, it was quite clear why it had to be true.' (Minio 1984, 17)

Mark Steiner (1978) proposes one of the earliest contemporary accounts of explanatory proof. I provide a sympathetic reading of his account and propose an extension that I hope lays the ground for future research.

#### 1.1. Three Conditions on Explanation in Mathematics

According to Mark Steiner:

'... an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem, such that from the proof it is evident that the result depends on the property. It must be evident, that is, that if we substitute in the proof a different object of the same domain, the theorem collapses; more, we should be able to see as we vary the object how the theorem changes in response. In effect, then, explanation is not simply a relation between a proof and a theorem; rather, a relation between an array of proofs and an array of theorems, where the proofs are obtained from one another by the 'deformation' prescribed above. (But we can say that each of the proofs in the array 'explains' its individual theorem.)' (Steiner 1978, 143)

So, a proof must fulfil three conditions in order to be explanatory:

- 1. The proof makes reference to a characterizing property of an entity or structure mentioned in the theorem.
- 2. It is evident from the proof that the result depends on the property, that is if we substitute in a different object of the same domain, the theorem collapses.
- 3. We should be able to see as we vary the object how the theorem changes in response: the proof is generalizable.

Steiner's suggestion is usually broken down into two conditions where points 1 and 2 are combined (see e.g. Hafner and Mancosu (2005), Resnik and Kushner (1987)), but the first condition here is important to Steiner. As we will see in the next section, he stresses the claim that the characterizing property must be a property of something mentioned in the theorem. Therefore, I include condition 1 as a separate criterion.

Steiner does not explicitly say that these conditions are meant to be necessary and sufficient criteria for a proof to be explanatory. Nevertheless, he puts forward examples which he takes it meet the schema and are therefore explanatory, suggesting that the conditions are sufficient on his view. Additionally, he considers a possible counterexample proposed by Feferman: a proof that is thought to be explanatory but does not seem to meet Steiner's criteria. In response, Steiner finds an appropriate characterizing property in order to fit the proof to his schema (Steiner 1978, 148). So, it also seems that the three conditions are necessary for a proof to be explanatory, on Steiner's view.

At first glance, there is an immediate tension in Steiner's account between conditions 2 and 3. To show that a proof is explanatory, we are required to show both that the theorem 'collapses' for a different object of the same domain and that it generalizes for some other object (presumably also in the domain; Steiner does not specify otherwise). So the second condition can't be understood as saying that the theorem 'collapses' for *any* other object in the domain; rather, the theorem should collapse for *some* object in the domain, and in particular presumably for some object which doesn't have the characterizing property. I will say more about this in section 2.

#### 1.2. The Sum of Positive Integers from 1 to n

Steiner considers three proofs (strictly speaking, proof sketches) of the fact that the sum of positive integers<sup>2</sup> from 1 to *n* is equal to  $\frac{n(n+1)}{2}$ . (Steiner 1978, 136-7)

1. Inductive proof (strictly speaking, this is only the inductive step):

S(n+1) = S(n) + (n+1) = n(n+1)/2 + 2(n+1)/2 = (n+1)(n+2)/2.

2. 'Symmetry' proof:

1	+	2	+	3 + .	+	п	=	S
n	+	(n-1)	+	(n-2) + .	+	1	=	S' = S
(n+1)	+	( <i>n</i> +1)	+	(n+1) + .	+	( <i>n</i> +1)	=	n(n+1)

3. 'Geometrical' proof:



'By dividing a square of dots, n to a side, along its diagonal, we get an isosceles right triangle containing

 $^2$  Except where otherwise specified, I use 'positive integers' and '(natural) numbers' interchangeably.

 $S(n) = 1 + 2 + 3 + \ldots + n$ 

dots. The square of  $n^2$  dots is composed of two such triangles – though if we put the triangles together we count the diagonal (containing n dots) twice. Thus we have

 $S(n) + S(n) = n^2 + n$ , q.e.d.' (Steiner 1978, 137)

According to Steiner, Proof 1 is not very explanatory, Proof 2 is 'more illuminating' and Proof 3 is 'perhaps an even more explanatory proof' (ibid., 145).

In this paper, I will focus on Proof 2 above and will give a detailed analysis of how the proof might fit Steiner's schema. Neither Steiner nor his respondents do so. Steiner makes only the following brief remark:

'Both explanatory proofs that the sum of the first *n* integers equals n(n+1)/2 proceed from characterizing properties: the one by characterizing the symmetry properties of the sum 1+2 + ... + n; the other its geometrical properties. By varying the symmetry or the geometry we obtain new results, conforming to our scheme.' (ibid.)

Resnik and Kushner (1987) focus on a more straightforward example: the irrationality of  $\sqrt{2}$ , which I will discuss in section 4. And Hafner and Mancosu write:

'Steiner's remarks imply that he apparently takes the symmetry properties as well as the geometrical properties of the sum 1+2+...+n as something – entities or structures? – *mentioned* in [the theorem]. This is very puzzling indeed and just highlights the need for precise definitions here. In the absence of such definitions ... we don't even have a clear enough grasp of Steiner's theory in order to apply and assess it in general'. (Hafner and Mancosu 2005, 233)

Given the lack of analysis in the literature, I think it is useful to look at this example in more depth. I hope my investigation will help us to understand Steiner's three conditions and bring out some important questions for Steiner's account.

# 2. Identifying a Characterizing Property

#### 2.1. Commutativity and Associativity of Addition

Recall Steiner's example, Proof 2 from the last section:

 $\frac{1}{n} + \frac{2}{(n-1)} + \frac{3}{(n-2)} + \dots + \frac{n}{1} = \frac{S}{S' = S}$  $\frac{n}{(n+1)} + \frac{n}{(n+1)} + \frac{n}{(n+1)} + \dots + \frac{n}{(n+1)} = \frac{n}{(n+1)}$ (Steiner 1978, 136) As we have seen, it is important that the characterizing property applies to something mentioned in the theorem, for Steiner. Steiner's formulation of the theorem to be proved is as follows:

(SUM) 
$$S(n) = 1 + 2 + 3 + ... + n = n(n + 1)/2$$
. (Steiner 1978, 136)

The theorem mentions the numbers 1, 2, 3, some arbitrary number n, the sum S(n), the operations of addition, multiplication, subtraction and division, and the relation of equality. Is Steiner's point that these entities (taken broadly) have symmetry and geometrical properties? Hafner and Mancosu suggest that Steiner takes the symmetry properties themselves to be mentioned in the theorem, but I don't think this is required: recall Steiner's stipulation that 'an explanatory proof makes reference to a characterizing property of an entity or structure mentioned in the theorem' (Steiner 1978, 147). Although the statement is ambiguous, I read this as follows: the *proof* must mention the characterizing property; and the characterizing property must be a property of something mentioned in the *theorem*.

Now, Proof 2 presents the sum of integers from 1 to n in two different ways. First, the sum is given in that order: 1 + 2 + 3 + ... + n. Then, the sum is given in reverse: n + (n - 1) + (n - 2) + ... + 1. One central insight of the proof is that, since addition is commutative and associative, these sums are the same. So we can add both sums together to get 2S(n), forming a new sum of n elements. Each element of this new sum is equal to n + 1, using the commutativity of addition again. So 2S(n) is equal to n lots of n + 1, leading to the desired result.

An initial suggestion, then, might be to propose the following characterizing property, P: the commutativity and associativity of addition for positive integers. This can be seen as a 'symmetry propert[y] of the sum' in the sense that the two different representations of the sum S(n) are symmetrical around their midpoint, and the property 'behind' this symmetry is property P.

Proof 2 implicitly appeals to property P (and in this sense perhaps makes reference to it), and the theorem mentions addition, some integers, and the sum of integers. So it seems that P holds of some entity or structure mentioned in the theorem, as required, hence fitting Steiner's first condition.

However, I will now show that P cannot be the property Steiner has in mind, because it doesn't fit Steiner's second condition. Recall that condition 2 stipulates that the result must depend on the property in the sense that the theorem collapses if we substitute in a different object.

It is true that the theorem depends on P in the following sense: if addition for integers were not commutative and associative, we wouldn't be able to get to the desired result using this proof method. But counterfactuals like this are difficult to understand in the mathematical case where results hold necessarily. According to Steiner, the appropriate counterfactual to consider instead runs as follows: 'If we substitute in a different object, the theorem would collapse'. It is important to consider here the domain from which the substituted object must come. The domain of the theorem is clearly (sums of) elements of  $\mathbb{N}$ , and Steiner explicitly specifies that the 'different object' should come from the same domain. So we want to substitute in a sum of some sequence of natural numbers.

Note that no matter which sum we choose, the cause of the theorem's collapse will not be a failure of the new object to instantiate property P. Any sum of elements in  $\mathbb{N}$  is commutative and associative, so the theorem couldn't fail because of a breakdown of associativity or commutativity. But intuitively we want to choose a new object which is 'different' in the sense that it lacks the characterizing property of the original object.<sup>3</sup> If not, it would be hard to understand Steiner's counterfactual condition as a way to flesh out the idea of a result depending on the characterizing property.

This is a problem for our initial choice of characterizing property, and points to a feature any successful characterizing property must have: there must be at least one object in the domain for which the characterizing property does not hold. Otherwise condition 2 would be impossible to fulfil (or is fulfilled only vacuously). I doubt that Steiner has vacuous fulfilment in mind, so on a charitable reading of Steiner's account, I must have identified the wrong characterizing property.

I will suggest a new one in the next section; nevertheless, I hope the discussion here has helped to clarify the content of Steiner's second condition. In particular, I suggest that the best reading of condition 2 runs as follows: 'If we substitute an object from the same domain which lacks the characterizing property, then the theorem collapses'. I will say more about how to understand the 'collapse' of the theorem in the next section.

#### 2.2. Arithmetic Sequences

Take the following property, Q: being an arithmetic sequence in  $\mathbb{N}$ . (An arithmetic sequence is one with a constant difference between consecutive terms). Let us see whether this new property helps to fit Proof 2 to Steiner's account. For ease of reference, I repeat the theorem and proof (or proof sketch):

Theorem: (SUM) S(n) = 1 + 2 + 3 + ... + n = n(n + 1)/2.

Proof 2:

1	+	2	+	3 + + n	=	S
п	+	(n-1)	+	$(n-2) + \dots + 1$	=	S' = S
(n+1)	+	( <i>n</i> +1)	+	$(n+1) + \ldots + (n+1)$	) =	n(n+1)

(Steiner 1978, 136)

<sup>3</sup> Unfortunately I do not have the space to address questions about mathematical ontology in this paper. I will speak interchangeably of entities and objects.

Steiner's three conditions are satisfied as follows.

# **1.** The proof makes reference to a characterizing property of an entity or structure mentioned in the theorem.

The proof and theorem both mention an entity S or S(n), the sum of the first n natural numbers. The first n natural numbers form a sequence in  $\mathbb{N}$  and this sequence is also mentioned, in the sense that its initial terms are explicitly listed. The sequence has property Q.

Now, the proof does not explicitly make reference to property Q, but it does implicitly appeal to Q: if the sequence did not have a constant difference between consecutive terms, then the terms of the sequence and its mirror image would not 'match up' to sum to (n+1) in each case. I suggest that property Q is represented in the diagram appealed to in the proof, so Steiner's first condition is fulfilled.<sup>4</sup>

We might worry whether property Q really fits with the fact that Steiner calls the desired property a 'symmetry propert[y] of the sum 1 + 2 + ... + n' (Steiner 1978, 145). Property Q is in fact a property of the sequence 1, 2, ..., n rather than of the sum. But since Steiner chooses to write 1 + 2 + ... + n here rather than S(n), I think we can reasonably take him to be referring to the summand – in this case the terms of the arithmetic sequence – rather than the sum itself. The sum is simply a number in  $\mathbb{N}$ , and it is not clear what symmetry property might hold of the sum.

There is a certain symmetry, on the other hand, in the way terms of the sequence are regularly spaced; it's this fact which means that the terms in the sequence and its mirror image match up in each case. Steiner's remark is vague enough that I think this level of symmetry seems like a reasonable fit.

# 2. It is evident from the proof that the result depends on the property, that is if we substitute in a different object of the same domain, the theorem collapses.

Consider the following sequence: 1, 4, 6, 25, 49, 101. The theorem collapses for this sequence because the terms of the sequence and its mirror image do not 'match up' in each case:

S*	:	1	+	4	+	6	+	25	+	49	+	101
S*	:	101	+	49	+	25	+	6	+	4	+	1
2S*	:	102	+	53	+	31	+	31	+	53	+	102

We can see from the diagram that the method used to calculate S in Steiner's example does not work for  $S^*$ , since  $2S^*$  is not a sequence with constant difference between terms, unlike 2S. It's not that there is nothing to say about the sum  $S^*$ ; we can still calculate its value. However, we can

<sup>4</sup> I do not have space to go into issues of reference in this paper.

only do so by adding each of the terms. We don't get an equation like the one in Steiner's original theorem to calculate the value of the sum.<sup>5</sup>

So, we have: (i) taken another object from the same domain (taken to be sequences of elements of  $\mathbb{N}$ ), where (ii) the object lacks the characterizing property and (iii) the theorem collapses. I hope that points (i) and (ii) are clear, but (iii) needs some further discussion. Steiner talks loosely about the theorem collapsing, but it seems from the discussion of the non-arithmetic sequence above that it is really the proof method that collapses. I think this is right; focusing on the proof method or argument, rather than the theorem, is a more fruitful way to interpret Steiner's second condition.

To back up my claim, consider the example of geometric sequences. Geometric sequences lack the characterizing property Q, and so in the spirit of Steiner's account, the theorem should collapse for geometric sequences. On my reading of 'the theorem collapses', this is true: the proof method above does not work for calculating the sum of a geometric series. Consider a simple example, such as the case where each term of the sequence is double the previous term for n > 1:

S'':	1	+	2	+	4	+	8	+	16	+	32
S'':	32	+	16	+	8	+	4	+	2	+	1
2S'' :	33	+	18	+	12	+	12	+	18	+	33

Just as before, 2S'' is not a sequence with constant difference between terms and the proof method fails to help us find a value for the sum S''. However, unlike in the previous example, it's not the case that no version of the *theorem* holds for the new sequence. There is a formula for finding the value of sums of geometric sequences: in general  $S_n = \frac{a_1(1-r'')}{1-r}$ ,  $r \neq 1$ , where *n* is the number of terms,  $a_1$  is the first term and *r* is the common ratio. So it's not clear that the theorem collapses, on a reading where this does not refer to the proof method used.

Hence I will take the following approach. Wherever Steiner writes 'the theorem collapses', I will take this to be shorthand for the following: Given a certain proof using a characterizing property R of an entity referred to in the theorem, the *theorem collapses* just in case the same argument applied to objects lacking the characterizing property R is not a proof of the modified proposition that is now the conclusion.

With this in mind, let's move on to condition 3 and see how the original proof is generalizable.

<sup>5</sup> Don't be deceived by the kind of symmetry still present in 2S\* as it doesn't get you an equation for calculating the sum that doesn't just involve adding up the terms. We just get

$$S*(n) = \begin{cases} \frac{1}{2} \left( 2(a_1 + a_n) + 2(a_2 + a_{n-2}) + \dots + 2(a_{n/2} + a_{(n+2)/2}) \right) & \text{if } n \text{ is even} \\ \frac{1}{2} \left( 2(a_1 + a_n) + 2(a_2 + a_{n-2}) + \dots + 2(a_{(n-1)/2} + a_{(n+3)/2}) + 2a_{n+1/2} \right) & \text{if } n \text{ is odd} \end{cases}$$

In both cases, this is simply a reformulation of  $S^*(n) = \sum_{k=1}^{n} a_{k}$ 

# 3. We should be able to see as we vary the object how the theorem changes in response: the proof is generalizable.

In keeping with my reading of the theorem collapsing, I will read condition 3 as follows. Given a certain proof using a characterizing property R of an entity referred to in the theorem, the *proof generalizes* just in case the same argument applied to other objects with characterizing property R is a proof of the modified proposition that is now the conclusion. I think we can charitably assume that Steiner's talk of the theorem changing as we vary the object is simply a loose shorthand for the account of generalizability just sketched.

Here are two examples of the theorem from Proof 2 generalized to cover other sums of integers.

**Theorem A:** The sum of the first *n* odd positive integers is equal to  $n^2$ . Or  $S(n) = \sum_{k=1}^{n} 2k - 1 = 1 + 3 + 5 + \dots + (2n - 1) = n^2$ .

**Proof A**:

S(n):	1	+	3	+	5	+	7	+ +	2n - 1
S(n):	2 <i>n</i> – 1	+	2 <i>n</i> – 3	+	2 <i>n</i> – 5	+	2 <i>n</i> – 7	+ +	1
2S(n):	2 <i>n</i>	+ +	2 <i>n</i>						

Since there are *n* terms in the sequence, we have  $2S(n) = n \cdot 2n$  and hence  $S(n) = n^2$ .

**Theorem B:** The sum of the first *n* terms of the following sequence: 1, 4, 7, 10, 13, ... is equal to  $\frac{1}{2}(3n^2 - n)$ . Or  $S(n) = \sum_{k=1}^{n} 3k - 2 = \frac{1}{2}(3n^2 - n)$ .

Proof B:

S(n):	1	+	4	+	7	+	10	+	+3n-2
S(n):	3n - 2	+	3n - 5	+	3n - 8	+	3n - 11	+	+ 1
2S(n):	3 <i>n</i> – 1	+	3 <i>n</i> – 1	+	3n - 1	+	3n - 1	+	+3n-1

Since there are *n* terms in the sequence, we have  $2S(n) = n \cdot (3n - 1) = 3n^2 - n$  and hence  $S(n) = \frac{1}{2}(3n^2 - n)$ .

Strictly speaking, these are proof sketches. We might ask for further clarification of the fact that the  $n^{th}$  term in Theorem B is 3n - 2, for example.<sup>6</sup> But Steiner seems happy with proof sketches, given the way he presents Proof 2 above.

Now, in both Theorems A and B we use exactly the same proof method from Proof 2 to get to a result about the value of the sum of the first *n* terms of some other sequence in  $\mathbb{N}$ . In Steiner's terms, explanation can thus be seen as a relation between the array of Theorems SUM, A and B and Proofs

<sup>6</sup> In general the  $n^{th}$  term of an arithmetic sequence is  $a_1 + (n-1)d$ , where  $a_1$  is the first term in the sequence and d is the common difference between two successive terms in the sequence.

2, A and B. Indeed, the array can be further expanded to cover all arithmetic sequences in  $\mathbb{N}$ , since for any sequence of numbers with a constant difference between consecutive terms, the same proof method can be used to get to a result about the sum of the first *n* terms of that sequence.

**Theorem C:** The sum of the first *n* terms of an arithmetic sequence  $\{a + (k-1)d\}_{k=1}^{n}$  is equal to  $\frac{1}{2}n(2a + (n-1)d)$ .

# Proof C:

S(n):	а	+	a + d	+ +	a+(n-2)d	+	a+(n-1)d
S(n):	a+(n-1)d	+	a+(n-2)d	+ +	a + d	+	а
$\overline{2S(n)}$ :	2a + (n-1)d	+	2a + (n-1)d	+ +	2a + (n-1)d	+	2a + (n-1)d

Since there are *n* terms in the sequence, we have 2S(n) = n(2a+(n-1)d), so  $S(n) = \frac{1}{2}n(2a+(n-1)d)$ .

That is, Proof 2 in fact generalizes to cover all sums of (finite segments of) arithmetic sequences.

So, it seems that Proof 2 meets Steiner's third condition on the basis of characterizing property Q. In the next section, I will look at Steiner's account of characterizing properties in more depth.

#### 3. Characterizing Properties in more Detail

#### 3.1. Unique, Partial and Multiple Characterization

In setting out his account, we have seen that Steiner writes 'I shall speak of 'characterizing properties', by which I mean a property unique to a given entity or structure within a family or domain of such entities or structures.' He goes on to say that '... a given entity can be part of a number of differing domains or families. Even in a single domain, entities may be characterized multiply' (Steiner 1978, 143).

Towards the end of the paper, he allows that '... an arbitrary equation with rational coefficients has not a unique Galois group, in the sense that no other equation has it....The concept of 'characterization' will have to be weakened to allow for partial characterization. The Galois group of E characterizes it in that the properties of the Galois group tell us much about E' (ibid., 149-50).

So, it seems that Steiner's account allows for unique, multiple and partial characterization. How should we understand these notions? Steiner gives an example: 'Thus, one way of epitomizing the number 18 is that it is the successor of 17. But often it is more illuminating to regard 18 in light of its prime power expansion,  $2 \times 3^2$ ' (ibid., 143).

That is, 18 is uniquely characterized as having prime power expansion  $2 \times 3^2$ . It is also uniquely characterized as being the successor of 17. Hence

34

we see that 18 is multiply characterized in  $\mathbb{N}$ : there are (at least) two ways of picking 18 out uniquely from other objects in the domain.

Now, let us consider Steiner's remark about partial characterization. I suggest that 18 is partially characterized by being an *abundant* number, where *n* is abundant if the sum of the divisors of *n* is at least 2n. This partially characterizing property 'tells us much' about 18 in the sense that it picks out 18 as a member of the set of abundant numbers, a proper subset of the set of positive integers. Some interesting mathematical results rely on picking out this set, so it seems that the partial characterization is mathematically relevant. For example, it is easy to show that prime numbers are not abundant number, and mathematicians including Erdős (1934) have proved various results about the density of abundant numbers in  $\mathbb{N}$ .

Note that 18 is also partially characterized by being an even number, and by being equivalent to  $0 \equiv 18$ . So it is clear that multiple partial characterization is also possible.

Let us now examine the characterizing property I suggested in the last section, Q: being an arithmetic sequence in  $\mathbb{N}$ . Property Q is clearly partially rather than uniquely characterizing. It does not pick out just one entity in the set of sequences in  $\mathbb{N}$ , but rather it picks out all objects with property Q, in the same way 'being an abundant number' picks out 18 among many other numbers (12, 20, 24, 30, ...).

Although Steiner initially focuses on uniquely characterizing properties, I don't think this is problematic. For one, Q is 'better' at characterizing entities than the commutativity and associativity of addition for integers, as it does not apply to everything in the domain of sequences in  $\mathbb{N}$ . Property Q allows us to distinguish between number sequences, in a way which is relevant to the proof at hand. And we will need to allow for partially rather than uniquely characterizing properties to make sense of the suggestion that the proof should generalize to other entities with the same property, as stipulated in my reading of Steiner's third condition.

In the next section, I explore my reading in more depth.

#### 3.2. Varying the Property

Steiner's third condition is stated in his words as 'We should be able to see as we vary the object how the theorem changes in response: the proof is generalizable' (Steiner 1978, 143). The meaning of 'vary the object' is vague. On my reading, we should take the condition to read as follows: 'Given a certain proof using a characterizing property R of an entity referred to in the theorem, the *proof generalizes* just in case the same argument applied to other objects with characterizing property R is a proof of the modified proposition that is now the conclusion'. The idea is that 'varying the object' means we take a given characterizing property and look for other objects with the same property.

However, other interpretations of Steiner's account in the literature suggest that we should instead consider different (but closely related) properties in order to see whether the theorem generalizes.

For example, Weber and Verhoeven (2002) discuss a case where theorems about right triangles and obtuse angles are proved by holding the proof-idea constant but using a different characterizing property (cosine of right angles versus cosine of obtuse angles).<sup>7</sup> Here the idea of 'varying the object' seems to be to look for closely related objects, where the objects are closely related in virtue of having closely related characterizing properties. This reading seems to fit with one of Steiner's remarks that 'generalizability through *varying* a characterizing property is what makes a proof explanatory' (Steiner 1978, 145, emphasis mine).

In order to stay faithful to this remark, I could simply modify my proposal in section 2 to suggest an alternative characterizing property,  $Q_{1,1}$ : 'being (an initial segment of) the arithmetic sequence with initial term 1 and constant difference 1'. To generalize the proof, then, we would simply consider objects with closely related characterizing properties, such as  $Q_{1,2}$  'being (an initial segment of) the arithmetic sequence with initial term 1 and constant difference 2' (the odd numbers), and so on.

However, I want to resist this move for now because I think my original reading is faithful to the spirit of Steiner's account and helps us to understand some of the remarks he makes about essences. Steiner writes:

'My view exploits the idea that to explain the behavior of an entity, one deduces the behavior from the essence or nature of the entity. Now the controversial concept of an essential property of x (a property x enjoys in all possible worlds) is of no use in mathematics, given the usual assumption that all truths of mathematics are necessary. Instead of 'essence', I shall speak of 'characterizing properties'' (Steiner 1978, 143).

Although I do not have space here to discuss the controversial topic of essential properties, I suggest we can see an explanatory proof as one which explains why all objects with a certain nature have a certain 'behaviour' pattern (fulfilling the relevant theorem), while all objects lacking this nature do not.

To clarify, my aim is not to propose a reading that is faithful to Steiner's account at all costs; rather, to propose a constructive and charitable reading that captures what seems to me the guiding idea behind his account, while making the account as interesting and persuasive as possible. In the next section I defend my reading of generalizability by showing that it helps to overcome a number of potential worries about Steiner's account.

<sup>7</sup> Thanks to an anonymous referee for suggesting this case.

#### 4. Generalizability

It is clear that generalizability is a crucial component of explanation for Steiner; a proof does not count as explanatory if it does not generalize. But suppose a mathematician has proved a result about a certain case. Putting aside cases of simple error, it might take some time for even a successful research mathematician to prove that the result generalizes. In such a case, it would seem odd to forbid the mathematician from classifying the proof as explanatory until she has proved the generalization. Why think a proof that doesn't (yet) generalize can't (yet) be classified as explanatory?

In response, we could argue that as long as what leads us to classify a proof as explanatory is the characterizing feature – in our case, the constant difference between consecutive terms – then it doesn't matter if we don't actually make the generalizing step. This fits with what I take to be the spirit behind Steiner's conception of explanation: it is the characterizing property (the 'essence' of a mathematical entity) which makes something explanatory, not whether we happen to have exploited that characterizing property to its full potential.

But a deeper point still remains. What if there is no further generalization (not just that we haven't discovered one)? It seems very plausible to me that a proof with no further generalization could nevertheless be explanatory. To deny this, I think, needs further argument.

In section 4.2, I show that my reading of generalizability addresses these concerns. First, however, I want to point out another problem with taking generalizability as we usually understand it to be the cornerstone of explanation.

The problem is that generalizability admits of degree: some proofs generalize more widely than others, as I illustrate in section 4.1. And we have seen that explanatoriness and generalizability are closely related for Steiner. Yet Steiner writes that his 'proposal is an attempt at explicating mathematical explanation, not *relative* explanatory value' (Steiner 1978, 143); so it seems that Steiner's conception of explanation does not admit of degree. In the next section I discuss this apparent mismatch by focusing on a particular example.

#### 4.1. Degrees of Generalizability and the Square Root of Two

Theorem: The square root of 2 is not rational.

**Proof:** We proceed by contradiction. Suppose  $\sqrt{2}$  were rational. Then  $\sqrt{2} = \frac{a}{b}$  for a,  $b \in \mathbb{N}$  and  $b \neq 0$ . Then  $2 = (\frac{a}{b})^2$  so  $2b^2 = a^2$ . As Steiner puts it:

'... by using the Fundamental Theorem of Arithmetic – i.e. that each number has a unique prime power expansion (e.g. 756 is uniquely  $22 \times 33 \times 71$ ) – we can argue for the irrationality of the square root of two swiftly and decisively. For

in the prime power expansion of  $a^2$  the prime 2 will necessarily appear with an even exponent (double the exponent it has in the expansion of *a*), while in  $2b^2$  its exponent must needs be odd. So  $a^2$  never equals  $2b^2$ , q.e.d.' (Steiner 1978, 137-8)

How does this prime factorisation proof meet Steiner's three criteria on explanation?

- 1. The proof makes reference to a characterizing property of an entity or structure mentioned in the theorem: as Steiner suggests, 'the prime power expansion of a number is a characterizing property' since by the Fundamental Theorem of Arithmetic, each number has a unique prime power expansion (Steiner 1978, 138 and 144). In this case the proof makes reference to the unique prime expansion of the number 2, which is an entity mentioned in the theorem.
- 2. It is evident from the proof that the result depends on the property, that is if we substitute in a different object of the same domain, the theorem collapses. Here, if we substitute in the number 4, we don't get a result about the irrationality of the square root of 4 because 'the prime power expansion of 4, unlike that of 2, contains 2 raised to an even power, allowing  $a^2 = [4]b^2$ '. (Steiner 1978, 144)<sup>8</sup>
- 3. We should be able to see as we vary the object how the theorem changes in response: the theorem is generalizable to cover more numbers. For example, we can substitute in 5 or any other prime, p, to get directly to a result about the irrationality of  $\sqrt{p}$ . Indeed, we can generalize further to the claim that 'the square root of n is either an integer or irrational ... [and] almost the same reasoning gives us the same for the  $p^{th}$  root of n'. (ibid.)

As we see, this example fits Steiner's three criteria. Drawing on earlier discussion, we can easily identify the characterizing property<sup>9</sup>; we can easily find counterexamples and generalizations of the theorem; and both the theorem and proof collapse when we substitute in another entity like 4, since the square root of 4 is rational.

Now, there are many different proofs of the irrationality of the square root of two, and some of these generalize less widely than the one just discussed. Take the following visual proof, presented in (Miller and Montague 2012, 110).<sup>10</sup>

<sup>8</sup> Steiner has 'allowing  $a^2 = 2b^2$ ' here, which must simply be an error.

<sup>9</sup> Here I put aside concerns that the correct characterizing property is a bit more subtle than suggested by Steiner. The proof relies not simply on 2 having a unique prime expansion, but on the fact that the unique prime expansion of a prime number is that prime itself; or more generally that whenever n is not a perfect square, one of the exponents in the unique prime expansion of n is not even.

<sup>10</sup> Printed with permission. Copyright 2012 Mathematical Association of America. All Rights Reserved.

#### **Tennenbaum's proof**

We now describe Tennenbaum's wonderful geometric proof of the irrationality of  $\sqrt{2}$ . Suppose that  $\sqrt{2} = a/b$  for some positive integers *a* and *b*; then  $a^2 = 2b^2$ . We may assume that *a* is the smallest positive integer for which this is possible. We interpret this geometrically by constructing a square of side *a* and, within it, two squares of side *b* (see Figure 1). Since the combined areas of the squares of side *b* equals the area of the square of side *a*, the shaded, doubly-counted square must have the same area as the two white squares. We have therefore found a smaller pair of integers *u* and *v* with  $u^2 = 2v^2$ , which is a contradiction. Thus  $\sqrt{2}$  is irrational.



FIGURE 1. Geometric proof of the irrationality of  $\sqrt{2}$ 

This proof seems a good candidate for an explanatory proof to me (at least as good a candidate as the prime factorisation proof presented by Steiner). But Tennenbaum's proof does not automatically generalize to cover all primes. Indeed, although Miller and Montague generalize the proof to cover the case n = 3, they write that 'For the irrationality of root 5 we have to modify our approach' and they show that further generalizations to cover triangular numbers only work up to n = 10 (Miller and Montague 2012, 111-113). So the proof generalizes less widely than Steiner's example.

I suggest that we can see the new visual proof as posing a dilemma for Steiner. On the one hand, Tennenbaum's proof seems like a good candidate for being an explanatory proof. If this is right, then it will be a point against Steiner's account if his schema for explanatoriness cannot readily accommodate the proof.<sup>11</sup> On the other hand, if Steiner's account can accommodate the proof (given some reasonable characterizing property), then we have two explanatory proofs of the same result which generalize to a different degree. This is a problem if we think, as Steiner seems to, that generalizability tracks explanatoriness and that explanation does not admit of degree.

Now, an easy way out of this dilemma for Steiner would simply be to allow that explanation does after all admit of degree. For example, Steiner's

<sup>&</sup>lt;sup>11</sup> I do not have the space here to discuss potential characterizing properties.

third condition could be modified along the following lines: 'The further a proof generalizes, the greater the degree of explanatory value'.

This would allow Steiner's account to accommodate both proofs of the irrationality of  $\sqrt{2}$ . For example, we could simply hold that both proofs meet a minimum explanatory threshold: that they generalize to cover three further cases, say. The unique factorisation proof is nevertheless more explanatory than Tennenbaum's proof, since it generalizes to cover many more cases.

Although this may seem like a promising approach, I want to suggest one good reason not to take it. Apart from problems of where to draw the somewhat arbitrary threshold, the problem is that modelling explanatory value directly as a function of generalizability could lead to problems of incommensurability. Different proofs of the same result may generalize not only to different degrees, but also to cover different kinds of cases. This means it may be impossible to directly compare the generalizability of two proofs, in order to determine which is more explanatory.

For example, in one paper Stan Wagon presents fourteen different proofs of a result about tiling a rectangle, comparing and classifying the proofs according to their possible generalizations. Some of these proofs generalize to cover the cylinder, while others generalize to the torus. As Wagon points out, 'no one of the proofs is best in terms of its ability to generalize' (Wagon 1987, 601).

Although there is more to be said on this point<sup>12</sup>, in this paper I want to focus on a way of circumventing the dilemma that fits better within Steiner's framework. In the next section, I argue that my reading of Steiner's generalizability condition deals neatly with the problems discussed so far.

# 4.2. Advantages of my Reading

Recall the reading of generalizability under which I interpreted Steiner's third condition: 'Given a certain proof using a characterizing property R of an entity referred to in the theorem, the *proof generalizes* just in case the same argument applied to other objects with characterizing property R is a proof of the modified proposition that is now the conclusion'.

Note that in fact generalizability does not admit of degree, on this reading. Suppose the characterizing property is R. The theorem generalizes just in case the same argument (resulting in a suitably modified proposition) applies to all other objects with property R. In some cases, there will be many objects with property R. And in some cases, there will be few such objects – perhaps only one. Condition 3 is met if the same argument can be

 $<sup>^{12}\,</sup>$  The obvious suggestion is to see generalizability as only one of multiple explanatory dimensions.

applied to *all* such objects: whether there are many or only one. It is an all-or-nothing condition, and cannot be partially met.

How does this help us to resolve the dilemma posed in section 4.1? Well, it is not the case that the prime factorisation proof that  $\sqrt{2}$  is irrational is automatically more generalizable than Tennenbaum's pictorial proof, simply because the former covers more numbers (namely all the primes). Rather, each proof is generalizable just in case the proof's argument can be applied to all cases of objects with the relevant characterizing property. In the prime factorisation case, it's obvious that the proof is generalizable in this sense (it's easy to see that substituting in another prime will work). In the pictorial case, this is less obvious and needs further work. Perhaps the pictorial proof is not generalizable in this way. But whether it is generalizable in this sense or not, is not a matter of degree.

So, Steiner can get around the dilemma I presented by claiming that neither explanation nor generalizability admit of degree on the best reading of his three conditions.

My reading of generalizability also has a number of further advantages. First, generalizability as presented here is independent of the interests of persons. So explanatoriness is not relative to persons' abilities and interests, on my reading of Steiner's account; to this extent it is an objective property and does not depend on whether we happen to have discovered the generalization.<sup>13</sup> Second, the generalizability requirement is not as restrictive as might first appear: (i) It does not exclude proofs which cover only one object and cannot be extended to apply to more objects, if the characterizing property applies only to a single object; (ii) In the same vein, the most general version of a proof (like Proof C in section 2.2) will still count as generalizable, as long as it covers all objects with property R.

We might worry whether my strong reading of generalizability matches the usual way we use the term in mathematics. But we can forestall this objection by taking generalizability on my reading to be short for 'generalizability with respect to the relevant characterizing property' and allowing for 'trivial generalizability' in case (i) just mentioned, where the characterizing property applies to just one thing.

One serious disadvantage, however, is that the account so far gives no indication of how an explanatory proof fulfils the primary epistemic function of an explanation, namely, to help us see why the fact to be explained is true. I will address this problem in the next section, in which I also return to examine Proof 3 from section 1 in light of the discussion so far.

<sup>&</sup>lt;sup>13</sup> I do not have space to explore debates about different kinds of objectivity in this paper. See for example (Burge 2010, 46-54).

## 5. New Directions

## 5.1. Ontic and Epistemic Aspects of Explanation

Wesley Salmon suggests a distinction between ontic and epistemic conceptions of explanation in the scientific case, where 'the epistemic conception takes scientific explanations to be arguments' and '[t]he ontic conception sees explanations as exhibitions of the ways in which what is to be explained fits into natural patterns or regularities' (Salmon 1984, 293). I will understand the terms 'epistemic' and 'ontic' in broadly this way, as distinguishing between an explanation that provides or increases our understanding, and an explanation that describes the structure or pattern of the (in this case mathematical) world.<sup>14</sup>

Steiner's account seems at first glance to have both ontic and epistemic components. For example, Steiner stipulates that in an explanatory proof it should be *'evident* that the result depends on the property', and that *'we should be able to see* as we vary the object how the theorem changes in response' (Steiner 1978, 143, emphasis mine), which seems to point to an epistemic aspect of his account.

One important question is what it means for the result depending on the property to be 'evident' on Steiner's account. Recall the discussion of Proof 2 in section 2. Why, we might ask, was the importance of property Q not immediately apparent when first analysing Proof 2? It's not clear how we might come to identify property Q except by thinking about how the proof might collapse or generalize to cover other sums of number sequences. This is how I came to identify the property. To a more practised mathematician or mathematics teacher, the required property might become apparent at first glance, perhaps based on familiarity with such results. If this is right, then identifying an appropriate characterizing property seems to build in an epistemic aspect to Steiner's account, where a proof is judged to be explanatory based on the cognitive capacities or background of the reader.

On the other hand, Steiner's account clearly focuses on properties and patterns of dependence, and indeed he rejects another proposed criterion of explanation connected with our ability to visualize on the basis that 'this criterion is too subjective to excite' (Steiner 1978, 143). So it seems that Steiner's primary aim is to capture an ontic or at least objective account of explanatory proof.

We could maintain an emphasis on the ontic aspect by arguing that a less experienced mathematician may *incorrectly* classify proofs as explanatory (or non-explanatory) based on incorrectly identifying the characterizing

<sup>&</sup>lt;sup>14</sup> Note that it is not clear we can maintain the distinction between these terms in the same way in all areas: for example, philosophers investigating mechanistic explanation have argued otherwise. See e.g. Illari (2013). My thanks to an anonymous referee for this point.

property. After all, Steiner does not specify that 'evident' means anything like 'easy to grasp'. Instead, we could simply propose that 'evident' be read as 'evident to a mathematician'; where, of course, further work would be needed to say who counts as a mathematician. In this way, a Steinerian account can maintain that a proof is explanatory in some sense independently of whether the average person actually classifies it as such.

In what follows, I want to go beyond Steiner's account and propose a way to combine both epistemic and ontic aspects, because of the problem raised in the last section: how does a proof's meeting Steiner's three conditions help us to see why the theorem proved is true?

My suggestion is that a proof counts as explanatory in an ontic or objective sense if it in fact meets Steiner's three conditions, involving some suitable characterizing property. We can be justified in calling the proof explanatory if we latch onto the relevant characterizing property (even if, as I suggested earlier, we don't latch on to it in full generality or we don't actually make the generalizing step). The proof also counts as explanatory in an epistemic sense if the property is presented in a way that enables us (or a person with suitably advanced mathematical skills) to latch on to the relevant property.

In my proposed extension of Steiner's account, the primary epistemic function of an explanation is fulfilled to the extent that the proof presents the characterizing property in an accessible way: the more accessible, the more readily we see why the theorem proved is true.

I do not have space to defend this suggestion in depth, but I illustrate the proposal in the next section by going back to examine Proof 3 from section 2.

#### 5.2. The Importance of Presentation

I suggest that Proof 3 involves the same characterizing property as in Proof 2, presented in a different way. For ease of reference, I repeat Proof 3:



'By dividing a square of dots, n to a side, along its diagonal, we get an isosceles right triangle containing

 $S(n) = 1 + 2 + 3 + \ldots + n$ 

dots. The square of  $n^2$  dots is composed of two such triangles – though if we put the triangles together we count the diagonal (containing n dots) twice. Thus we have

$$S(n) + S(n) = n^2 + n$$
, q.e.d.' (Steiner 1978, 137)

Like Proof 2, Proof 3 also essentially involves counting each element of the sum twice. Here the second sum is upside down, rather than backwards as in Proof 2. But this is just a different geometrical representation of the same idea. The 'geometric' proof, as Steiner calls it, also only works because there is a constant difference between terms in the sum (which are represented by dots).

We could apply the same argument to other instances of arithmetic sequences as in the first two images below.



Whenever the sequence has a constant difference between terms, two dot copies of the sequence will form a rectangle. It's easy enough to see that no rectangle will be formed if there is not a constant difference between terms, as in the third array of dots.

Note that we can't generalize the proof to cover an arbitrary arithmetic sequence as we did with Proof 2, because we can't represent an arbitrary arithmetic sequence using dots. This might tempt us to say that Proof 3 fails to meet Steiner's third condition; but recall the reading of generalizability defended earlier. It's not part of the condition that the argument has to apply to the abstract general description of the case. Rather, the important thing is that the argument can be applied to any individual case of an object with the same characterizing property. This is true (allowing for a broad understanding of applying the argument: in some cases it may be hard to physically draw all of the dots!).

Let me clarify this. The relevant question is: Does the set S of all objects with property Q contain the 'arbitrary arithmetic sequence' a, a + d, a + 2d,...? If so, the argument in Proof 3 can't be used to cover this arbitrary case as the argument in Proof 2 can (as shown in Proof C, section 2.2). This is because

we can't represent a + d, for example, using a concrete number of dots in the way we can represent a particular number. However, I suggest that the arbitrary sequence is not really a sequence in *S*; rather, it stands for any sequence in this set, and the argument in Proof 3 does cover each one of these sequences. So Proof 3 meets Steiner's third condition, I suggest.

My proposal, then: Proofs 2 and 3 make use of the same characterizing property – being a sequence of numbers with a constant difference between terms. They both meet Steiner's three conditions. What differs between the proofs is the degree to which they make the relevant characterizing property accessible, and thereby the degree to which they achieve the primary epistemic function of an explanation: enabling or helping us to grasp why the proof's conclusion is true.

I find the characterizing property easier to spot and presented more clearly in Proof 2, but this may depend on the reader's cognitive background and skills. In general, identifying a suitable characterizing property could be an epistemically challenging task, if the property is presented unclearly or in a way not accessible to someone with a particular set of cognitive skills. Such a proof can nevertheless be declared objectively explanatory, according to Steiner's account. It's just that the average reader can't immediately access the property – and hence the explanation.

To summarise, I hope my proposal has provided reason to think that successful future developments of Steiner's account could include both an ontic component – whether or not the proof contains a suitable characterizing property which meets conditions 1-3 – and an epistemic component – whether the characterizing property is presented in a way accessible to a reader with a certain cognitive background.

#### 6. Conclusion

I have attempted to give a maximally charitable reading of Steiner's account, one which makes sense of the puzzling comments he makes about his sum-of-integers example. I analysed one of Steiner's proofs (Proof 2) and identified a characterizing property in order to show that the proof does indeed meet Steiner's three conditions and hence can be called explanatory on his account.

Subsequently, I raised a few potential worries about Steiner's account, and showed how my reading of his third generalizability condition helps Steiner to avoid these problems.

Finally, I returned to examine Proof 3 and suggested that a proof may display its characterizing property more or less clearly, allowing for a new epistemic component to Steiner's account that I hope lays the ground for further research.

#### References

- BAKER, Alan (2009). Mathematical explanation in science. British Journal for the Philosophy of Science, 60:611–633.
- [2] BURGE, Tyler (2010). Origins of Objectivity. Oxford: Oxford University Press.
- [3] ERDŐS, Paul (1934). On the density of the abundant numbers. *Journal of the London Mathematical Society*, s1-9(4):278–282, 1934.
- [4] HAFNER, Johannes and MANCOSU, Paolo (2005). Visualization, Explanation and Reasoning Styles in Mathematics, chapter: Varieties of Mathematical Explanation, pages 215–250. Springer.
- [5] ILLARI, Phyllis (2013). Mechanistic explanation: Integrating the ontic and epistemic. *Erkenntnis*, 78:237–255.
- [6] LANGE, Marc (2014). Aspects of mathematical explanation: Symmetry, unity, and salience. *Philosophical Review*, 4:485–531, 2014.
- [7] MANCOSU, Paolo (2008). The Philosophy of Mathematical Practice, chapter: Mathematical Explanation: Why it Matters, pages 134–149. Oxford: Oxford University Press.
- [8] MILLER, Steven and MONTAGUE, David (2012). Picturing irrationality. *Mathematics Magazine*, 85(2):110–114, 2012.
- [9] MINIO, Roberto (1984). Interview with Michael Atiyah. *The Mathematical Intelligencer*, 6(1):9–19.
- [10] RESNIK, Michael D. and KUSHNER David (1987). Explanation, independence and realism in mathematics. *British Journal for the Philosophy of Science*, 38: 141–158.
- [11] SALMON, Wesley (1984). Scientific explanation: Three basic conceptions. *Proceedings of the Biennial Meeting of the Philosophy of Science Association*, 2:293–305.
- [12] STEINER, Mark (1978). Mathematical explanation. *Philosophical Studies*, 34: 135–151.
- [13] WAGON, Stan (1987). Fourteen proofs of a result about tiling a rectangle. *The American Mathematical Monthly*, 94(7):601–617.
- [14] WEBER, Erik and VERHOEVEN, Liza (2002). Explanatory proofs in mathematics. *Logique et Analyse*, 179-180:299–307.

Josephine SALVERDA josephine.salverda.10@ucl.ac.uk