Mathematical Explanation and Understanding: A Noetic Account

Abstract

We defend a noetic account of intramathematical explanation. On this view, a piece of mathematics is explanatory just in case it produces an appropriate type of understanding. We motivate the view by presenting some appealing features of noeticism. We then discuss and criticize the most prominent extant version of noeticism, due to Matthew Inglis and Juan-Pablo Mejía-Ramos, which identifies explanatory understanding with the possession of detailed cognitive schemas. Finally, we present a novel noetic account. On our view, explanatory understanding arises from meeting specific explanatory objectives, the theory of which we briefly set out.

1 Introduction

We defend a noetic account of intramathematical explanation. On this view, a piece of mathematics is explanatory just in case it produces an appropriate type of understanding. We begin with motivation for noeticism in §2. In §3, we discuss and criticize [Inglis & Mejía-Ramos 2021], the most prominent noeticist theory currently on the market. We present our account in §4.

2 The appeal of noeticism

According to the noetic theory of explanation, to possess an explanation of some fact is to be in an appropriate state of understanding with respect to that fact. Call this state *explanatory understanding*. Noeticists can identify an explanation itself with a body of information that could, typically would, or actually does produce explanatory understanding.

Noeticism has several attractive features. For one, virtually everyone agrees that explanation and understanding are closely related. Any theory of explanation should be able to account for this link, and noeticism does so in the most direct possible way.

Second, noeticism can explain its rivals' successes. Alternative theories of explanation work as well as they do because they've correctly identified various possible sources of understanding. So it's no surprise that these theories often make true predictions. While non-noeticist views go wrong in mistaking a *common route* to explanation for a *necessary and sufficient condition*, they contain genuine insights which the noeticist can embrace with open arms. Rival theories of explanation have a harder time here. It's not obvious what a dependence theorist should say, for example, about the appeal of unificationism. No doubt some story can be told, but it's likely to be relatively complex, and in the end the proponent of one theory may have to dismiss much of her opponents' views as confusion or error.

Finally, noeticism has no trouble handling different types of explanantia. It gives a unified account of theorems, proofs, diagrams and other mathematicalia: any type of information that confers the appropriate sort of understanding is fit to serve as an explanation, and all count as explanatory for the same reason. Alternative theories often struggle on this front. Some, for instance, focus exclusively on explanatory proof, and what they say about proofs is hard to generalize to other cases. (The accounts of [Steiner 1978] and [Lange 2014]are in this boat.) Meanwhile, the counterfactual account ([Baron et al. 2020]) only applies straightforwardly to theorems explained by theorems. This is a bad sign for these views. It's implausible that there are several fundamentally different kinds of mathematical explanation, each working in a special *sui generis* way. A theory that suggests such a disunified picture—or worse, one that has to deny the existence of explanations embraced by mathematical practice—is a theory deserving of skepticism.

3 Inglis and Mejía-Ramos's account

3.1 The theory

We aren't the first to advocate this sort of view. Adopting the terminology of [Wilkenfeld 2014], Matthew Inglis and Juan-Pablo Mejía-Ramos—hereafter IMR—defend what they call a *functional theory* of explanation ([Inglis & Mejía-Ramos 2021]). On this view, what makes something an explanation is that it functions as a source of understanding. IMR's account is therefore a version of noeticism. (We prefer 'noeticism' over 'functionalism' because the latter gives no hint about the nature of the relevant function.)

A basic component of IMR's account is the concept of a *schema*, introduced by Piaget and still widely used in cognitive and educational psychology. A schema is, roughly speaking, a structured mental representation of a subject matter S that facilitates remembering, recognizing, predicting and reasoning about information related to S.

Schemas are thought to play a role in many cognitive tasks. For instance, you have a *house* schema, which encapsulates your understanding of the important general attributes of houses (e.g. that they have parts, are made of materials, and have functions) as well as some default specific values for those attributes (e.g. that the parts of a house are rooms, that the materials are wood or brick, and that the function is human habitation). Accessing this schema can provide reasoners with various benefits. Since its components are strongly cognitively interlinked, the schema allows you to quickly classify something as a house on the basis of a few key characteristics. And once classified, the schema helps you recognize other expected house-features, make inferences about features you didn't directly perceive, and remember what you saw later.

We also have schemas for more abstract subjects, like chess games and mathematical problems, which play cognitive roles similar to schemas for everyday objects. A skilled chess player will have a detailed *Sicilian Defense* schema, for instance, encoding the situations in which this opening is likely to be played, its most important variations and the counterplays for each, typical middle-game positions to which the Sicilian tends to lead, and so on. A player with many such schemas at her disposal will be able to choose moves quickly and accurately. A player with none will have to painstakingly note the positions of each piece on the board, call to mind their possible moves and calculate the consequences of each, leading to slower and more error-prone performance.

On IMR's view, the creation and consolidation of schemas is constitutive of mathematical understanding. As they write, "one can be said to have understood something when a sufficiently well-organised schema... has been encoded into long-term memory" (S6381).

A notable consequence of this view is that, for IMR, understanding why P is nothing more than having a body of well-structured knowledge relevant to P from which P can be inferred. In other words, IMR believe that "the distinction between... objectual understanding and understanding-why/explanatory understanding... is a matter of degree rather than of kind" (S6377).

3.2 The trouble with schemas: objectual understanding is not explanatory understanding

While we share IMR's noeticist sympathies, our first goal is to argue that this schema-based approach is on the wrong track. Possessing well-organized schemas is neither necessary nor sufficient for explanatory understanding. More generally, it's a mistake to think of explanatory understanding as nothing but a sufficient amount of objectual understanding.

First, the necessity claim. Suppose that T is a theorem belonging to subject $\mathcal{S}$. Suppose also that we want to understand why T is true, and hence to explain T's truth. We claim that one can often do this without having an extensive, detailed or well-structured $\mathcal{S}$-schema.

Example: Consider Lange's well-known "calculator number" case. A calculator number is a six-digit number of the form *abccba*, where the digits *abc* are the numbers forming some row, column or main diagonal of a square 1-9 calculator keyboard. There are sixteen calculator numbers in total, and it turns out that each one is divisible by 37. Why does this curious fact hold? Lange offers the following explanatory proof ([Lange 2014], 488):

> 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 $10^5 a + 10^4 (a+d) + 10^3 (a+2d) + 10^2 (a+2d) + 10 (a+d) + a$. Regrouping, we find this equal to

$$\begin{aligned} a\left(10^5+10^4+10^3+10^2+10+1\right) + d\left(10^4 + 2\cdot 10^3 + 2\cdot 10^2 + 10\right) \\ = 1221\,(91a+10d) \\ = (3\cdot 11\cdot 37)\,(91a+10d)\,. \end{aligned}$$

We agree that this proof explains why calculator numbers are divisible by 37. But it's not clear that the proof either provides or exploits any suitable well-organized schema. The list of concepts, facts and techniques in play is quite small—the notion of an arithmetic progression, multiplication and addition, factoring, and so on. And the argument does nothing very profound or ingenious with these tools. So the average reader is unlikely to upgrade her basic arithmetical and algebraic schemas as a result of working through the proof.

One could try to claim that the proof confers understanding by contributing to a more specialized schema, say *calculator number*. But there's a risk of ad hocness here. Given any proof whatsoever, one can typically concoct some highly problem-specific schema to which the proof can be said to contribute. We need a stronger condition to distinguish explanatory proofs from the rest.

Even setting aside this worry, it's doubtful that the above strategy could work. Does Lange's proof really leave us with a well-organized *calculator number* schema? Surely not. The only interesting fact about calculator numbers appearing in the proof is that the digits *abc* must be in arithmetic progression. Perhaps it's hard to say in general what counts as a well-organized schema, but this isn't a borderline case. Someone who knows basic definitions and the facts used in the proof may still be almost completely clueless about calculator numbers. Such a person will have only the most threadbare schemas. And yet the proof is explanatory (for them).

We conclude that it's possible to possess understanding and grasp an explanation without having anything resembling a high-quality schema. Now we want to argue that the reverse implication fails too.

Our basic objection is that we often know quite a lot about a subject or problem without grasping why some of the associated facts are true. In such cases, it seems we have a well-organized schema without possessing understanding.

To take a well-known example, consider the four color theorem. Planar graphs and graph colorings have been major areas of study for well over a century, and graph theorists have amassed a great deal of detailed, sophisticated and intricately structured knowledge about these subjects. Appel and Haken's computer-assisted 1976 proof of the four-color theorem makes masterful use of the techniques developed to date. (As they note: "Over the past 100 years, a number of authors including A. B. Kempe, G. D. Birkhoff, and H. Heesch have developed a theory of reducibility to attack the problem. Simultaneously, a theory of unavoidable sets has been developed and the fusion of these has led to the proof" ([Appel & Haken 1976], 711). The body of knowledge embedded in Appel and Haken's work surely amounts to a well-organized schema if anything does.

But the proof of the four color theorem is also paradigmatically unexplanatory. As the graph theorist Paul Seymour writes, "We would very much like to know the 'real' reason the [four color theorem] is true; what exactly is it about planarity that implies that four colours suffice? Its statement is so simple and appealing that the massive case analysis of the computer proof surely cannot be the book proof" ([Seymour 2016], 417).[1] Similarly, according to Cristopher Moore and Stephan Mertens, the Appel-Haken proof leaves us with "an unsatisfying state of affairs. The best proofs do not simply certify that something is true—they illuminate it, and explain *why* it must be true. While the Four Color theorem may be proved, it is not well understood" ([Moore & Mertens 2011], 125; emphasis in original). So we seem to have a proof that turns on various well-organized schema, and which synthesizes those schemas to impressive effect, but which nevertheless fails to explain or produce understanding.

As the above examples show, one can have good explanatory understanding while possessing virtually no objectual understanding (as in the calculator number case). Conversely, one can have abundant objectual understanding while totally lacking explanatory understanding (as in the four color theorem case). Focusing exclusively on objectual understanding—that is, on the quality of an inquirer's schemas—is a false start for a noeticist theory of explanation.

[1] "The book" is God's book, which contains the best possible proof of every theorem. (The idea is Paul Erdős's.)

4 A different style of noeticism

As noeticists, we agree with IMR and others that a mathematical explanation is nothing but a piece of mathematics that provides an appropriate type of understanding. We think existing accounts haven't yet succeeded at identifying this type of understanding. We try to do so below.

4.1 Motivating the view

As our discussion of IMR has shown, possessing information relevant to P—even abundant, well-structured and cognitively tractable information—is in general neither necessary nor sufficient for understanding why P. So the first task for an adequate noeticist theory is to avoid equating explanatory understanding with a mere accumulation of knowledge.

The key insight needed to solve this problem, we take it, is as follows. As seekers of explanatory understanding, we typically aim to resolve specific confusions, clear up specific mysteries, uncover specific sorts of reasons-why, obtain specific insights. In a word, we aim to meet some *explanatory objective* or other. A proof is therefore explanatory when, and only when, it's capable of meeting a relevant objective. Conferring or drawing upon a vast store of knowledge is no use if the knowledge in question doesn't help achieve our explanatory goals.

This picture lets us make sense of the numerous cases discussed above. Consider the calculator number phenomenon. We were initially puzzled by the fact that all calculator numbers are divisible by 37; our objective was to understand why this occurs. A proof would count as explanatory relative to our objective just in case it intelligibly presents a reason that accounts for the divisibility fact. The proof in question did so by observing that calculator numbers have the form $10^5a+10^4\left(a+d\right)+10^3\left(a+2d\right)+10^2\left(a+2d\right)+10\left(a+d\right)+a$, which simplifies to $\left(3\cdot 11\cdot 37\right)\left(91a+10d\right)$. Meeting our explanatory objective in this case didn't require deep or extensive knowledge about calculator numbers, but just a couple elementary facts of the right kind.

In other situations, however, our objectives may be quite different, and a suitable explanation will have to meet a distinct set of needs. For instance, many areas of contemporary mathematics aim to transfer problem-solving resources between domains, or to unify seemingly disparate problems within a common abstract framework ([Lehet 2021]). A number theorist might seek a specifically geometric understanding of why the Riemann hypothesis is true. Or it might occur to a category theorist to wonder whether dissimilar-looking results in linear algebra, set theory and group theory have a common explanation in terms of adjoints and limits. In cases like these, developing rich and extensive schemas will be crucial to obtaining explanatory understanding. But again, however, not just any problem-relevant knowledge will do. All the analytic number theory in the world won't help the number theorist achieve her goal, even if it leads to a perfectly good proof of the Riemann hypothesis. And mastering linear algebra, set theory and group theory individually won't deliver the category theorist's unifying perspective. Whether or not a given piece of mathematics counts as explanatory, then, depends sensitively on the type of explanation one is after.

This conception of explanation has some affinities with erotetic accounts like Lange's, which identify mathematical explanations with answers to antecedently specified why-questions. We find Lange's work insightful, but our view departs from his in important ways, as we'll discuss below.

4.2 Explanatory objectives

We've suggested that a piece of mathematics confers explanatory understanding when it meets an appropriate explanatory objective. An explanatory objective, in turn, is a demand for insight of a certain sort—insight into why a theorem is true, why a phenomenon occurs, how a result should be interpreted, or how one piece of mathematics relates to others, for instance.

Explanatory objectives may be more or less specific. Upon first learning that P, one might be at a loss to see how P could possibly be true, in which case any compelling reason would do. Alternatively, one might seek an explanation using particular methods or possessing particular epistemic properties. (This might be the case even if one already understands why P in a certain sense or from a certain viewpoint.) It's widely agreed that a given mathematical fact may have several independent and equally good explanations, and our account of objectives makes sense of this.

4.2.1 Explanatory objectives and why-questions

Many explanatory objectives can be expressed as why-questions: "Why is every planar map four-colorable?" "Why is every calculator number divisible by 37?" Often enough, we're in a position to formulate these questions in advance of inquiry, and we know what kind of information would count as a satisfying answer. So it's no surprise that why-question theories like Lange's seem quite appealing.

Recall that, on Lange's view, a proof is explanatory when it "exploits a certain kind of feature in the problem: the same kind of feature that is outstanding in the result being explained" ([Lange 2014], 489). For instance, explaining a theorem which exhibits a striking symmetry calls for a proof that makes use of a corresponding symmetry in the problem setup. Thus it only makes sense to talk about explanatory proof when a why-question has been prompted by some specific noteworthy quality of a theorem. As Lange says, his account "predicts that if [a] result exhibits no noteworthy feature, then to demand an explanation of why it holds, not merely a proof that it holds, makes no sense. There is nothing that its explanation over and above its proof would amount to until some feature of the result becomes salient" ([Lange 2014], 507).

We think Lange has identified a common and important type of explanatory objective: to answer questions of the form "Why does theorem T exhibit surprising feature F?" But this account isn't the whole story about mathematical explanation, or even about explanatory proof. We discuss several ways in which explanatory objectives can diverge from Lange's picture.

First, not all successful explanations start with a sharply posed question about a specific feature of a result or phenomenon. In some cases, our initial explanatory objective is relatively vague: we've discerned something mysterious, obscure, or suggestive about the observed facts, and we seek better understanding without knowing quite what form it will take. In other situations, an objective is revealed only during or after inquiry, as for example when a proof reveals an illuminating but unanticipated aspect of a phenomenon.

Both types of case are common. Gauss's quadratic reciprocity theorem is an example of the first (cf. [D'Alessandro 2021]). In the early days of the theorem's conjecture and proof, mathematicians recognized it as a striking statement that hinted at a mysterious relationship between pairs of prime numbers. Explaining the theorem was a primary goal of Gauss and his successors: Gauss called it his "golden theorem" and returned to it repeatedly, producing eight different proofs during his lifetime. But these early explorers could hardly have known just what they were looking for. Number theory was still in its infancy, and mathematicians lacked both the basic facts surrounding reciprocity and the tools needed to conceptualize and investigate them. Truly enlightening proofs would come only after more of this machinery had been developed. Yet Gauss was capable of pursuing an explanatory proof even without a precise idea of what such an explanation would amount to. He recognized the theorem as profound, intuition-defying, and indicative of major gaps in understanding, and this was enough to arm him with an actionable explanatory objective.

For an example of the second type, consider the unexpected change in perspective brought about by abstract algebra with respect to familiar number systems. Pre-modern mathematicians had never thought to ask, for instance, why Euclid's lemma is true—that is, why prime numbers p have the property that p divides a product ab only if p divides either a or b alone. This seems at first glance like a basic property of the primes not standing in need of explanation.

But the modern turn in algebra led to a new understanding of primality. According to the revised picture, it's in fact Euclid's lemma that captures the correct general definition ($p \mid ab \implies p \mid a$ or $p \mid b$), while the familiar condition ($n \mid p \implies n = 1$ or $n = p$) represents an accidental feature called *irreducibility* that prime elements possess in some rings but not others. From this perspective, Euclid's lemma is a non-obvious fact which we might well hope to explain: Why do primality and irreducibility in $\mathbb{N}$ happen to coincide, when they come apart elsewhere? The answer is that the integers are a unique factorization domain, and Euclid's lemma holds in all such rings. And this is a perfectly good explanation, even if nobody thought to pose the why-question until after the answer was available.

Lange's account seems to rule out the possibility of both types of case, since it claims that one can only seek or give an explanatory proof after one has formulated a why-question about some particular salient feature of the theorem to be explained. This prediction is incorrect. By contrast, our account allows for such cases. Some explanatory objectives are broad and open-ended. And we sometimes add a given objective to our list only after seeing the associated explanation.

This last point deserves further comment. Although the term 'objective' may suggest a goal to be set aside once achieved, this isn't our usage here. Rather, an explanatory objective may (and typically does) remain in force even after we've found what we sought, in the sense that the question and its answer continue to interest us. Of course, our objectives do sometimes change, if we decide that some are misconceived, unfulfillable or no longer relevant. But the mere fact of finding a satisfying explanation doesn't force us to abandon the associated objective. For the same reason, it's perfectly possible to take up an objective only after discovering an explanation whose existence we didn't suspect. Doing so just means endorsing the explanatory question as worthwhile.

4.2.2 Objective value and the value of objectives

On our view, mathematical explanations are ways of meeting explanatory objectives. There will often be many actual or possible objectives associated with a given explanandum, and we're free to choose which to adopt according to our interests.

This doesn't entail, however, that all explanatory objectives are created equal. On the contrary, there are many dimensions of value along which one objective can be better or worse than another.

For instance, some objectives are simply unsatisfiable, in the sense that there is no explanation answering to our wants. Any objective that asks for an explanation of a false proposition falls into this category. So does an objective that requests a nonexistent type of explanation for a true proposition (say, a purely algebraic proof of the fundamental theorem of algebra), or any explanation at all for an unexplainable truth (say, the fact that $8 + 11 = 19$).

Among objectives that can actually be fulfilled, some are more natural, fruitful, or consequential than others. A successful geometric explanation of the Riemann hypothesis would throw open the floodgates of mathematical knowledge, leading to far-ranging insights about the relationship between number theory and geometry. Explaining why calculator numbers are divisible by 37 is entertaining but comparatively pointless. Both objectives are legitimate, but their epistemic demands and benefits are completely unalike.

The virtues associated with different types of objective may also trade off against one another. For instance, transparent explanations are simple, vivid, and cognitively tractable, but often "shallow", in that they use elementary methods that don't generalize easily. On the other hand, deep explanations are intellectually rich but often "opaque", in that they require complex and unintuitive theoretical machinery. So the goodness of an objective isn't a simple linear matter. An explanatory objective (or successful explanation) that's outstanding in one respect may be subpar in others.

Still, it may make sense to ask whether an objective A is more worthwhile than another objective B, all things considered—that is, whether A's package of merits and demerits is superior to B's, with the elements of each package appropriately weighted by importance. Mathematicians often make such judgments about overall pursuitworthiness, and reasonably so.

The fact that one can choose which objectives to pursue, then, doesn't mean that all objectives are on equal footing. Mathematical progress depends crucially on asking the right questions: simply accumulating answers isn't enough. The evaluation and careful selection of objectives is therefore an important part of explanatory practice.

4.2.3 What makes an objective explanatory?

Mathematicians have many epistemic aims. They seek to gain knowledge, to make fruitful conjectures, to gather evidence, to develop reliable heuristics, and so on. Many of these desiderata don't involve explanatory understanding. What, then, distinguishes explanatory objectives from epistemic interests of other kinds?

On our view, *explanatory objective* is a cluster concept. Thus there's no single property that all explanatory objectives have in common. Instead, there's a set of features which such objectives tend to share; all explanatory objectives will possess at least a few features from the list, with varying degrees of overlap between one objective and another.

This picture is similar to Collin Rice and Yasha Rohwer's cluster-concept account of scientific explanation ([Rice & Rohwer 2021]). Rice and Rohwer are motivated by two observations. First, there seems to be no single theory that fully captures the diversity of types of explanation while avoiding counterexamples. Second, the pluralist's response to this situation is unsatisfying: if we merely say that there exist many types of explanation $E_1, E_2, \ldots, E_n$ sharing nothing interesting in common, we're at a loss to explain why we apply the same label to all the E_i. A cluster-concept theory can accommodate both insights.

We depart slightly from Rice and Rohwer, in that we defend monism about *mathematical explanation* but propose a cluster theory of *explanatory objectives*. That is, we think all explanations work in essentially the same way—by providing explanatory understanding, and hence by meeting explanatory objectives—but we hold that the variety of explanatory objectives are related to one another only by family resemblance. (Still, it's harmless enough to regard our view as a cluster-concept account of mathematical explanation, as long as one keeps in mind that the relevant cluster is located a couple of analytical steps away.)

The following table, inspired by Rice and Rohwer's similar diagram (1037), displays some of the features shared by various explanatory objectives, focusing on styles of explanation previously discussed in the literature:

	Abstract	Deep	Transparent	Unifying	Impure	Counter-factual	Mechanistic	Salience-based (Lange)	Deformability-based (Steiner)
Rationalizes striking feature of a theorem								X	X
Cognitively compelling and tractable			X				X		
Links diverse subject areas		X			X				
Identifies dependence relations	X					X	X		X
Introduces higher level of abstraction	X	X		X					
Aims at broad context and generalizability		X		X	X				X
Seeks a proof with specific formal features				X				X	X
Seeks a proof with specific contentual features	X	X			X				

Some types of explanatory objectives and their characteristic features.

As the table suggests, there's significant overlap between the features possessed by different objectives, but no single feature common to all of them. We can also discern several families of objectives sharing broadly similar explanatory goals.

Impure, deep and unifying explanations are similar, for instance, in seeking links between subjects and a bird's-eye perspective. We might call this family of objectives Seeing the Big Picture. (Steinerian explanation belongs here too, to a lesser degree, as it also requires a limited kind of generalizability at the level of proof.) Transparent and mechanistic explanations look for vivid, compelling reasons presented in a cognitively accessible way. Call this family Getting a Handle. Abstract, counterfactual, mechanistic and Steinerian explanations are all interested in dependence relations of various kinds, and thus aim to identify grounds or difference-making features. This family might be called Putting Things in Order. Finally, both Steiner's and Lange's styles of explanation focus on rationalizing striking properties of a theorem with the help of an appropriately matched proof. Call this family Accounting for Surprises. (Incidentally, this taxonomy helps explain why Steiner's account has provoked such widespread interest as well as widespread criticism—the view is a peculiar multi-family hybrid incorporating several kinds of explanatory objectives.)

Each of these families represents what's plausibly a natural and important sort of understanding. Seeing the Big Picture reflects our interest in comprehending patterns and placing facts in context. Getting a Handle captures the drive to grasp phenomena with intuitive clarity. And so on.

4.2.4 Summary

Putting the above all together, our view is that we possess explanatory understanding when we've successfully met an explanatory objective. Such objectives include explicit why-questions but aren't limited to them. The value of an explanatory objective is multifaceted, but mostly independent of particular agents' beliefs and preferences. Finally, *explanatory objective* is a cluster concept: such objectives fall into a few broad and overlapping categories without sharing any single essential feature in common.

5 Conclusion

Our account has all the general virtues of noeticism discussed in §1—it makes sense of the relationship between explanation and understanding, it can account for the insights of rival theories, and it can handle diverse types of explanantia. We've argued that our account is preferable to non-noetic theories of mathematical explanation like Lange's, as well as extant versions of noeticism like IMR's. Finally, our view has the resources to answer important questions facing any noeticist theory: questions about objectivity, value, and the nature of explanatory understanding, for instance. There are many further issues to address, of course, but we hope this sketch makes for a convincing promissory note.

References

[Appel & Haken 1976] Appel, K. and W. Haken. 1976. "Every planar map is four colorable." *Bulletin of the American Mathematical Society* 82, 711-712.

[Baron et al. 2020] Baron, Sam, Mark Colyvan and David Ripley. 2020. "A counterfactual approach to explanation in mathematics." *Philosophia Mathematica* 28, 1-34.

[D'Alessandro 2021] D'Alessandro, William. 2021. "Proving quadratic reciprocity: Explanation, disagreement, transparency and depth." *Synthese* 198, 8621-8664.

[Dellsén 2021] Dellsén, Finnur. 2021. "Understanding scientific progress: The noetic account." *Synthese* 199, 11249-11278.

[Inglis & Mejía-Ramos 2021] Inglis, Matthew and Juan-Pablo Mejía-Ramos. 2021. "Functional explanation in mathematics." *Synthese* 198 (Suppl. 26), S6369-S6392.

[Lange 2014] Lange, Marc. 2014. "Aspects of mathematical explanation: Symmetry, unity, and salience." *Philosophical Review* 123, 485-531.

[Lehet 2021] Lehet, Ellen. 2021. "Impurity in contemporary mathematics." *Notre Dame Journal of Formal Logic* 62, 67-82.

[Moore & Mertens 2011] Moore, Cristopher and Stephan Mertens. 2011. *The Nature of Computation*. New York: Oxford University Press.

[Rice & Rohwer 2021] Rice, Collin and Yasha Rohwer. 2021. "How to reconcile a unified account of explanation with explanatory diversity." *Foundations of Science* 26, 1025-1047.

[Seymour 2016] Seymour, Paul. 2016. "Hadwiger's conjecture." In John Forbes Nash and Michael Rassias (eds.), *Open Problems in Mathematics*, Berlin: Springer, 417-437.

[Steiner 1978] Steiner, Mark. 1978. "Mathematical explanation." *Philosophical Studies* 34, 135-151.

[Wilkenfeld 2014] Wilkenfeld, Daniel. 2014. "Functional explaining: A new approach to the philosophy of explanation." *Synthese* 191, 3367-3391.