A Social Epistemological Approach to Mathematical Rigor

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree

Doctor of Philosophy

in the Graduate School of The Ohio State University

By

Zoe Ashton

Graduate Program in Philosophy

The Ohio State University

2024

Dissertation Committee:

Professor Stewart Shapiro, Advisor

Professor Neil Tennant

Professor Christopher Pincock

Abstract

There are many features a good mathematical proof may exhibit – it may be simple, surveyable, interesting, explanatory, pure, perspicuous, and even beautiful. The focus of this dissertation is rigor, a necessary feature of modern, mathematical proof. This dissertation proposes an account of mathematical rigor developed around the concept of conviction.

In Chapter 1 and Chapter 2, I focus on the standard view of rigor. The standard view of rigor connects informal rigor to formal proof in a chosen formal deductive system. A proof is rigorous just in case it can be translated into a formal derivation. The standard view faces a number of problems with respect to mathematical practice. Many have argued that it fails to account for changing standards of rigor over time, diagrammatic proofs, and the psychology of mathematical knowledge. In Chapter 2, I work through three general categories of standard view. I present new objections to each of the three categories. One of the main takeaways is that mathematicians are convinced of the steps of the informal proof itself, not that some other formal proof could exist.

In Chapter 3 I give a new account of rigor which is driven by the imagined universal audience. I argue that a proof is completely rigorous when each step is one that the mathematician's universal audience assents to. Each inference is judged to be rigorous when it convinces one's universal audience. For the mathematician, this amounts to the judgment that the inference would convince everyone. The audience view escapes the objection I posed to the standard view, since the mathematician judges that each inference is

convincing, not that some other object could exist. I also argue that my account is superior to the standard view since the audience view accommodates a gradable notion of rigor. A proof is more rigorous than another when it has more inferences to which the universal audience assents.

In Chapter 4, I connect the audience view of rigor to core issues in social epistemology. The audience view claims that rigor judgments depend on social features, including who participates in mathematical practice. Given that participatory injustices seem to occur in mathematical practice, I argue a mathematician's universal audience will be influenced by those injustices. I argue that eliminating participatory injustice will lead mathematicians to have a more robust universal audience. A more robust universal audience leads to rigor judgments that are more stable over time. I argue that participatory injustice is detrimental to the proofs, as well as the participants. This provides a new reason to both eliminate participatory injustice in mathematics and to diversify the profession. In Chapter 5, I conclude and discuss some future issues that a social epistemological approach to rigor must address.

Dedication

To my nieces, Aubrey and Willow

Acknowledgements

Philosophically, my work has greatly benefited from the support of my committee and the department. I am particularly grateful to Stewart with whom I've had some of the most philosophically enriching and incredibly fun conversations. I am also indebted to both Neil and Chris for their time, discussions, and support. Their feedback has greatly improved and sharpened my thinking over the years. I was also very lucky to benefit from the support and guidance of numerous department members, including Lisa Shabel and Robert Kraut, who always made time for me.

Outside the department, I'm grateful to all the conference attendees and referees who have listened to me over the years. But I'm especially grateful to Andrew Aberdein who introduced me to philosophy of math and spent countless hours walking me through it.

I'm especially grateful to my friends for their unwavering support. Most of all, I'm grateful for Haley and Dzintra, without whom I would not have survived the MA or the PhD. There are no words for how lucky I am to know them. I'm grateful for all of the friends I've made along the way but special recognition goes to Dave, Lyla, Connor, Jordan, and Soyeong whose support has been invaluable.

I'd also like to acknowledge my parents for all of the sacrifices they made to support me and my education. Finally, I'd like to thank my brothers who kept me grounded and never let me stray too far into the Platonic realm.

Vita

2016 .... BS Mathematical Sciences, Florida Institute of Technology

2016 .... BA Humanities, Florida Institute of Technology

2018 .... MA Philosophy, Simon Fraser University

2018-2024 .... Graduate Teaching Associate, The Ohio State University

Publications

Aberdein, A., **Ashton, Z.** (2024) Argumentation in Mathematical Practice in: Sriraman, B., *Handbook of the History and Philosophy of Mathematical Practice*, Springer, pgs. 2665-2687.

Ashton, Z. (2021) Audience role in mathematical proof development. *Synthese*, 198, pgs 6251–6275.

Moreno, M., Bain, T., Moreno, M., Carroll, K., Cunningham, E., **Ashton, Z.**, Poteau, R., Subasi, E., Lipkowitz, M., Subasi, M. (2021) Identifying Clinical and Genomic Features Associated with Chronic Kidney Disease (CKD). *Frontiers in Big Data*, 3, 9 pgs.

Ashton, Z. (2018) Mathematical Problem Choice and the Contact of Minds in: Zack, M. and Schlimm, D., *Research in History and Philosophy of Mathematics*, Springer, pgs. 191-203.

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

Ashton, Z., Mizrahi, M. (2017) Intuition Talk is Not Methodologically Cheap: Empirically Testing the "Received Wisdom" about Armchair Philosophy. *Erkenntnis*, 83(3), pgs 595-612.

Fields of Study

Major Field: Philosophy

Table of Contents

List of Figures

Chapter 1

Introduction

Mathematics is rigorous. With the demand for rigor come tangible benefits; rigor is connected to features like reliability, correctness, and certainty. But what exactly does it mean for a proof to be rigorous? Burgess [2015] provides us with a seemingly straightforward answer which divides the demands of rigor into four conditions:

> (1) Mathematical rigor requires that every new proposition must be deduced from previously established propositions (coming either from earlier in the same paper, or from the earlier literature).
> (2) On pain of circularity or infinite regress, if later propositions must be proved from earlier ones, then we must start from some unproved propositions or *postulates*.
> (3) Mathematical rigor requires that every new notion must be defined in terms of previously explained notions (either from earlier in the same paper, or from the earlier literature).
> (4) On pain of circularity or infinite regress, if later notions must be explained in terms of earlier ones, then we must start from some unexplained notions or *primitives*. [Burgess, 2015, 6-7]

And thus we know exactly what rigor demands: start from some unexplained primitives and unproved postulates. The only way to rigorously introduce new propositions is by deduction from previous propositions deduced in turn, eventually, from the postulates. The same holds for new concepts, they are to be defined only in terms of the previously explained notions or the primitive notions.

Of course, (1)-(4) are almost never fully satisfied in practice. So it cannot be that (1)-(4) define rigor as witnessed by proofs in mathematical journals. Burgess himself notes this and settles on a less stringent version. He writes "what rigor requires is that each new result should be obtained from earlier results by presenting enough deductive steps to produce conviction that a full breakdown into obvious deductive steps would in principle be possible" [Burgess, 2015, 97]. We now seem to have two sets of demands from two types of rigor. First there is the clearly defined rigor of (1)-(4), which when combined with a specified formal language is *formal rigor*. Then, there is a second notion of rigor which is "enough" for mathematicians.

What informal proofs possess is an informal rigor. Most proofs found in mathematics textbooks and journals are informally rigorous. In modern mathematical contexts, rigor is a necessary condition for being a proof. I follow that convention and so 'non-rigorous proof' doesn't exist. There are still non-rigorous purported proofs and non-rigorous arguments. According to Burgess above, informal rigor is related to formal rigor. A proof is informally rigorous when it exhibits enough formal rigor so as to convince mathematicians that a formal proof is possible. This relationship between informal and formal rigor is a version of *the standard view* which will comprise the bulk of our discussion in Chapter 1 and Chapter 2.

In the rest of Chapter 1, I'll review some early discussions, and examples of, informal rigor. I'll then discuss the literature on the standard view. While the standard view is appealing at first glance, it faces a number of objections. Finally I'll outline some extant alternatives to the standard view and the issues that arise with them. This chapter sets up the rest of the dissertation, where in Chapter 2, I'll raise new objections to the standard view. These objections motivate my own alternative view, presented in Chapter 3 and explored further in Chapter 4.

1.1 Examples of Informal Rigor

As we noted above, full formal rigor is not usually required of typical mathematical proof. Although there is no clear agreement about what informal rigor is, there is some discussion of its importance through the importance of informal proof. One of the most important features highlighted by Kreisel [1967, 1987], Lakatos [1976, 1978a], and Robinson [1991, 1997] is the role informal rigor plays in definitions and understanding.

Kreisel [1967] is one of the earliest to argue for the importance of informal rigor. His analysis is also unique since it focuses on axioms and notions, not inferences. According to him, rigor aims at two things: '(i) to make [the analysis of intuitive notions] as precise as possible ... and (ii) to extend this analysis, in particular not to leave undecided questions which can be decided by full use of evident properties of these intuitive notions" [Kreisel, 1967, 138-139]. In other words, informal rigor aims to take intuitive notions and produce precise, mathematical definitions of them. This also leads us to the axioms. Kreisel [1967] goes on to provide examples of informal rigor at work. For example, he argues that our intuitive notion of 'set' was a vague concept and it was only by examining the intuitive notions that we arrived at mathematical definitions and axioms. Kreisel [1987] argues that informal rigor should be invoked to understand discussions of Church's Thesis. In each of his examples, the focus is on producing definitions and axioms from informal notions. He justifies the use of informal rigor but does not tell us exactly how it works. He writes that

> We do not pretend to have a theory of a mechanism which explains how we come to form intuitive notions which are so astonishingly successful ... But we regard it as absurd to reject the use of this ability just because we don't have a theoretical explanation; this is what the formalist doctrine of precision does [Kreisel and Krivine, 1967, 169].

Kreisel's view is that we somehow undertake an analysis of intuitive notions and produce definitions and axioms from them. The ideal version of this analysis is informal rigor. He provides us with examples including the notion of set, the notion of validity, and Turing's description of Church's Thesis. But he does not tell us what informal rigor is and he doesn't address the rigor of inferences in proof.

While Kreisel was focused on the importance informal rigor plays in the discovery of notions, Lakatos [1976, 1978b] provide examples of the importance of informal proofs, which we assume to be informally rigorous. Lakatos [1978b] separates formal proofs from 'pre-formal' and 'post-formal' proofs. Both pre-formal and post-formal proofs are informal proofs. Lakatos argues that these informal proofs play a key role in his Popperian, quasi-empiricist account of mathematics. For Lakatos, formal systems should be formalizations of pre-established informal mathematical theory. He outright rejects the characterization of an informal proof as a formal proof with steps and references hidden.

Lakatos gives an example of pre-formal proofs. Specifically he turns to a proof of Euler's theorem on simple polyhedra: V - E + F = 2, where V is the number of vertices, E the number of edges, and F the number of faces of the polyhedron. Lakatos claims that the pre-formal proof, which involves moves like 'assuming the polyhedron is made of thin rubber', 'flattening it on a surface', and even 'cutting and re-arranging it', is "intuitively showing that the theorem was true" [Lakatos, 1978b, 64-65]. There is no definition of proof at the pre-formal level which allows us to decide whether this counts as a proof or "mere persuasive argumentation, rhetorical appeal" [Lakatos, 1978b, 65]. This example is explored in-depth in Lakatos [1976] where the discovery of various counterexamples to the proof results in refinement of the definition of polyhedron. As Lakatos [1976] shows, it is important for the eventual definition, and subsequent formal proofs, that we give and evaluate pre-formal proofs. Post-formal proofs, like those of the duality principles in geometry and undecidability are also informal proofs. According to Lakatos, these informal proofs are open to falsification:

> [Informal proofs] prove something about that sometimes clear and empirical, sometimes vague and 'quasi-empirical' stuff, which is the real though rather evasive subject of mathematics. This sort of proof is always liable to some uncertainty on account of hitherto unthought-of possibilities. [Formal proof] is absolutely reliable; it is a pity that it is not quite certain ... what it is reliable about [Lakatos, 1978b, 69]

In other words, while formal proof gives us reliability, it does not take center stage in Lakatos's discussion of proof. Rather, informal proof, with its potential for falsification through counterexample, is where most of the mathematical work happens.

Robinson [1991] investigates the relationship between formal proofs and informal proofs. His main worry is that formal proof is incapable of satisfying the goal of informal proof. Informal proofs aim to increase understanding.[1] According to Robinson [1991, 1997], following an informal proof involves understanding the story as it develops and "grasping the meaning of the words and diagrams" [Robinson, 1991, 269]. For

[1]Robinson does not adopt a specific philosophical view regarding either understanding or explanation. He seems to believe we have a strong enough intuitive sense of what understanding and explanation are.

Robinson [1997], "PROOF = GUARANTEE + EXPLANATION" and he argues that formal proofs can satisfy the guarantee but not the explanation component because they hide the meaning. He gives multiple examples of experiments which "go through an intuitive, rigorous proof which has a high degree of explanatory power" Robinson [1997]. Among these examples are the mutilated chessboard and the 27-subcube problem. Each proof is informal and invokes meanings, intuitive perception, and automatic operations. In sum, the rigor judgments in these experiments are meant to be independent of their ability to be formalized since the formalization would obscure their explanatoriness.

The literature reviewed in this section did not give a definition of informal rigor like the one in Burgess [2015]. However, they do comprise some of the earliest explicit discussions of the importance of informal rigor and informal proof. We also see some characterizations of what informal rigor and proof are about – understanding, communication, refining notions, and discovering definitions and axioms. This literature recognizes a role for formal proof in terms of correctness while still demonstrating the importance of informal rigor and proof. But this leaves us with a question: how are we to relate informal and formal rigor, if at all? We turn to this question next.

1.2 Some Varieties of Standard View

We posed the obvious question at the end of the last section: how are informal and formal rigor related? The typical response in the literature is the standard view. Broadly, formulations of the standard view endorse the following: (a) privilege formal proof as being the *ideal* and (b) informal proofs qualify as rigorous because of their relationship to corresponding formal proofs. This is not just the claim that all informal proofs *can* be formalized (aka a formalization thesis). The standard view assumes a formalization thesis and adds that *informal rigor is determined by reference to formal rigor.*[2] In what follows, I review some major accounts of the standard view.

One of the earliest, and most frequently cited, explicit formulations of the standard view is Azzouni [2004]'s derivation-indicator view. Azzouni argues that it is derivations in an 'algorithmic system' which are characteristic of mathematical practice and, in particular, explain widespread mathematical consensus. An algorithmic system is a system which codifies the deductive rules for derivations such that "the recognition procedure for proofs is mechanically implementable" [Azzouni, 2004, 83]. Of course, Azzouni admits that derivations are not the currency of mathematical practice. Informal proofs are, according to Azzouni, indicators of a corresponding derivation. Azzouni asks quite a bit of his algorithmic systems – they are algorithmic but they are not restricted to a particular logic, subject-matter, or explicit language. He does this to ensure mathematicians can transcend these systems and are seen as "sprinting up and down algorithmic systems, many of which he or she invents for the first time" [Azzouni, 2004, 103].

Whatever one may say of the algorithmic system, Azzouni must also tell us how mathematicians manage these systems and indicate the existence of derivations. This is particularly difficult since mathematicians don't have to know the rules codified in an algorithmic system. To do so, Azzouni [2005, 2009] invoke the idea of an 'inference package' which is a "capacity to recognize the implications of several assumptions by means of the representations of objects wherein those several assumptions have been knit together (psychologically)" [Azzouni, 2009, 20]. These packages allow mathematicians to do what they normally do in informal proofs – recognize and tease out the implications of their assumptions without being aware of a formal derivation. The inference packages are rich enough

[2] It's probably for this reason that Tanswell [2017] refers to the standard view as a formalist-reductionist thesis.

to accommodate a wide set of algorithmic systems with which we reason compatibly.

There are a number of questions and objections to Azzouni's view. In part these objections focus on the seemingly mysterious nature of algorithmic systems, inference packages, and the indicator process. Let's turn to a more precise formulation of the standard view given by Hamami [2019]. Hamami separates two different accounts of rigor – descriptive and normative. A descriptive account of rigor characterizes the mechanisms by which proofs are judged to be rigorous in practice. A normative account stipulates conditions that a proof ought to satisfy in order to count as rigorous. The standard view embeds both a descriptive and a normative account which is held together by the conformity thesis. The conformity thesis states that if a proof is rigorous in the descriptive sense, then that proof is rigorous in the normative sense.

According to the normative component of the standard view, a proof is rigorous iff it can be routinely translated into a formal proof. Hamami [2019] interprets 'routinely translated' as 'algorithmic' and attempts to provide algorithmic translations through four levels of granularity. He offers three algorithmic translations which move (1) from vernacular proof to a higher-level proof comprised of higher-level inference rules, (2) from higher-level proof to intermediate-level proof comprised of primitive rules of inference, and (3) from intermediate-level proofs to lower-level proofs comprised of rules in a system. In effect, Hamami [2019] de-mystifies the connection between informal and formal proofs by characterizing it as a sequence of algorithmic translations. Using his algorithmic method, and his distinction between the descriptive standard view and the normative standard view, he can answer a number of worries raised in the literature. According to the descriptive standard view, informal rigor is judged according to whether the mathematician can perform translation (1). The truth of the descriptive standard view, in conjunction with the conformity thesis, ensures the normative standard view. Moreover, the descriptive component can be empirically tested, so, in theory, Hamami [2019] has provided an empirically testable argument for the standard view. We'll discuss Hamami [2019]'s view in more detail in Chapter 2.

As we noted before, Burgess [2015] also seems to endorse the standard view. Tatton-Brown [2019] discusses Burgess [2015] as an example of the standard view. Although Tatton-Brown does not reference Hamami [2019], he gives an argument which could be seen as support for the descriptive standard view. Tatton-Brown roughly, and with different terminology, argues that students in analysis courses learn to prove with very detailed proofs. Then, as they progress through their courses, they learn to perform translations like Hamami's first translation when faced with less detailed proofs. By a sort of induction, Tatton-Brown argues that professional mathematicians are also engaged in giving proofs which they think can be translated into the highest level of detail and explicitness.[3] We'll re-visit Tatton-Brown's arguments in Chapter 2.

Many are implicitly committed to some version of the standard view since it straightforwardly connects formal rigor and explanatory, communicative informal rigor. Azzouni's and Hamami's views comprise two vital interpretations of the standard view – an indicator-based view and a translation-based view. As we will see in the next section, both of these interpretations face a number of objections.

[3]Tatton-Brown has nothing to say about how mathematicians trained without a modern analysis course made rigor judgments. Nor does he explain how his argument works for those of us who learned analysis from Rudin [1976], an elegant industry-standard book which, starkly in contrast to Tatton-Brown's choice of Abbott [2015], includes no explicit discussion of what proof is, which inferences are appropriate, or how much detail to provide.

1.3 Objections to the Standard View

Having now set out some concrete versions of the standard view, we now turn criticisms which comprise a large part of the literature on rigor. I investigate three themes of criticism in this section. These three themes are the meaning-based objections, the diagram-based objections, and the translation-based objections. These are not the only ways to object to the standard. For example, Larvor [2016a] objects on the basis that a naive derivation-indicator view produces a problematic regress. Tanswell [2015, 2017] both argue that the standard view is over-generative: for any informal proof, there are many corresponding formal proofs. Tanswell argues that there is no agent-independent link which selects the 'correct' formal proof. Antonutti Marfori [2010] objects that the standard view creates an epistemic demand on mathematical knowledge that is not met in practice. Further, I'll introduce new objections in Chapter 2.

1.3.1 The Meaning Objections

The name 'the meaning objection' hardly does justice to the complex of ideas which goes into the objections. Broadly speaking, these objections argue that the standard view cannot be the correct account of informal rigor since rigor is related to grasping the concepts being manipulated. As Lakatos noted, the formal proof is certain, but it's not clear what it's talking about. We can see this as an early characterization of the meaning objection. Burgess [2015] calls this the paradox of rigor: any truly rigorous treatment of a subject matter will ipso facto cease to be a treatment of the subject matter alone. This is because formal logic is topic-neutral. Any formal proof cannot encapsulate topic-specific information like the meaning of 'polygon.' So in formalizing a proof, one moves from engaging in topic-specific subject matter to topic-neutral logical inferences on the syntactic forms. This is a problem for the standard view because informal proofs seem to require a grasp of the subject matter, not manipulation of symbols. This is exhibited in both Lakatos [1976]'s and Robinson [1991, 1997]'s examples. One of the most sustained versions of the objection is found in Rav [1999, 2007] who directly engages with Azzouni's derivation indicator view.

Rav [1999] distinguishes between derivations which are syntactic objects of formal systems and proofs which are *conceptual* and have *irreducible semantic content*. Rav [1999] criticizes 'Hilbert's Thesis', which is the idea that every informal proof can be converted into a derivation in a suitable formal system. 'Hilbert's Thesis' is not yet the standard view, rather it is a formalization thesis. But, as we noted above, a formalization thesis is presupposed by the standard view. Rav [1999] argues that mathematical proofs, not theorems or derivations, are the sites of mathematical knowledge and, since conceptual meanings are lost in formalization, so is epistemic content. Note, again, that these are arguments against replacing proof with formal proofs and they are used to motivate a primary view of Rav's own which I'll discuss in sub-section 1.4.1.

Rav [2007] is directly against the derivation-indicator view espoused by Azzouni [2004]. He writes that "when it comes to the nature of the logical justification of mathematical arguments in proofs ... Azzouni put[s] his faith in formal derivations, even if just indicated ... I hold that mathematical proofs are cemented via arguments based on the meaning of the mathematical terms" [Rav, 2007, 294]. Rav examines the historical and methodological wealth of proof practices and concludes that the derivation-indication criterion cannot sustain rigor judgments in all these cases. Moreover, Rav argues, once we have an informally rigorous proof, the formal proof does not increase reliability. Formalizing the proof can only occur if the informal proof, including its mathematical concepts and informal logic, are already verified to be reliable.[4] The formal proof adds nothing to reliability. Azzouni

[4]The informal logic of mathematics that Rav cites is Aberdein [2006]'s analysis.

[2009]'s response to the meaning objection is that the inference packages make it seem like mathematicians engage primarily with semantic content. But since inference packages are consistent with algorithmic systems, this concept-based reasoning is in line with the SV. Of course, this response is only satisfactory to the extent that (a) Azzouni's algorithmic systems are sustainable and (b) inference packages provide a satisfactory account of mathematical cognition. We will re-visit (a) in sub-section 1.3.3.

Thus, Rav's meaning-based objection to the standard view is this: topic-specific semantic information is central to justification but cannot be in the formal proof. Judgments about goodness and rigor of informal proof cannot be grounded in the formal proof, since formal proofs typically lack semantic information. This is not the last we'll hear about the meaning-based objection since we will re-visit the positive component in Section 1.4.

1.3.2 The Diagram Objections

The second class of objections are related to the role of diagrams and visual reasoning in informal proof. Take, for example, Euclid's *Elements* which invokes visual reasoning and diagram manipulation. The diagram-based inferences in Euclid have proved to be successful, fruitful, and largely reliable. Proofs in knot theory and low-dimensional topology also employ diagrammatic proof and modern mathematicians classify these proofs as rigorous. But under the standard view, rigor judgments are justified in relation to a formal proof. Thus, an immediate worry arises: diagrammatic reasoning seems to be important and informally rigorous but difficult to relate to formal proof.

Manders [2008] sees mathematical practice as a cooperative effort for control. Under his view, traditional geometrical demonstrations include both a discursive component and a diagram. The discursive text attributes assertions to the diagram and steps in the demonstration can be either attributions or constructions in the diagram. Manders [2008] introduces two types of attributions: exact and co-exact. He claims that the fallacies of diagram use involve the presupposition that one can read exact information off a diagram. The fallacies are the supposed problems with diagram use; specifically they encompass the worry that diagram use is unreliable and leads to false conclusions. This is not the case. Exact attributions cannot, and are not, read off of the diagram. The reason for this is that co-exact features, e.g. that one region includes another, are unaffected by continuous variation of a diagram. On the other hand, exact features, like the straightness of a line or the size of an angle, are not stable under continuous variation. By invoking this distinction, Manders [2008] argues that exact attribution is licensed only by the discursive text, while co-exact attributions can be licensed by either the diagram or the discursive text. By invoking these controls, the Euclidean geometer avoids disarray and incorrect inference, i.e. he avoids the fallacies of diagram use. According to Manders [2008], the diagram-based inferences are rigorous and succeed at the cooperative effort to control.

Manders [2008] is not directly raising an argument against the standard view. He's interested in describing the reliability and practice of Euclidean geometry. But the reliability that Manders [2008] explains does lead to an objection. In diagrammatic proof, reliability and justificatory work is done by methods of control in the diagram. The objection requires another claim: these methods of control cannot be translated to formal proof. There are two ways to interpret "cannot be translated." The first interpretation is an anti-formalization thesis: there is no way to formalize the diagrammatic moves. The second interpretation is that the reasoning in diagrammatic proof cannot be *faithfully* translated. The second objection provides a stronger objection but is less fleshed out.

The first response available to the standard view is in Azzouni [2013] which focuses on diagram use generally, not just Euclidean diagrams. Azzouni admits that proofs invoking diagram use can be rigorous. One must be careful about the scope of conclusions that are read off diagrams but, given such care, there's no reason to reject diagram usage as unreliable. But he rejects the claim that this informal rigor judgment is not related to formal rigor. To fit diagrams into the derivation-indicator view, Azzouni claims that what the diagram licenses is actually *mechanically recognizable*. There are properties of the diagram that can be mechanically recognized as an admissible rule in a suitable algorithmic system. Again, we see that the force of Azzouni's reply lies in a coherent concept of algorithmic systems.

A much stronger reply to the anti-formalizability diagram objection is to show that many diagrammatic proofs can be formalized. Although this is not the intended aim of Avigad et al. [2009], their formal system for Euclid's *Elements* is an impressive move in that direction. Avigad et al. [2009] are clear that they intend only to create a formal system which models the proofs of the *Elements*. They identify and precisify the individual inferences which govern Euclid's proofs. They aim only to provide the description and formalization of the norms, without explaining why the norms arose or why they should be followed. Showing that we can produce a formal system for the *Elements* undermines the claim that diagrammatic reasoning cannot be formalized.

The second interpretation of the diagram objection is harder to characterize. But the core of the objection is this: even though one can formalize diagrammatic proofs, the new formal proof is importantly different. As Larvor [2019] puts it: 'traduttore traditore' which means that every translation traduces. According to Larvor, we should be looking at the Euclidean proofs only in their original idioms. These idioms are both context-dependent and content-dependent. The meaning and allowable actions in that context matters to how we judge the proof. We should not translate and then judge the proofs. He specifically relates this problem to Avigad et al. [2009]; according to Larvor, the proofs of the Avigad et al. [2009] system may have the same results, but they do not have the same inferences. As in Manders [2008], Larvor [2019] goes on to characterize requirements on rigorous diagrammatic proof which do not relate to formal proofs.

Recent work on knot theory, which Larvor [2019] discusses, raises the issue more explicitly. For example, by examining determinate case studies in early knot theory, de Toffoli & Giardino challenge the "model of formal logic as adequate to account for proof" [De Toffoli and Giardino, 2016, 27]. Clearly they are talking about the standard view, as Tatton-Brown [2019] points out. De Toffoli and Giardino [2016] provide an account of how the proof of Alexander's Lemma works, ie how it is judged to be rigorous, without reference to an associated formal proof. Tatton-Brown [2019], in adopting a Burgess [2015] inspired version of the standard view, argues that the inferences used in proving Alexander's Lemma are not inconsistent with the standard view. This leads to a refinement of the original argument, found in De Toffoli [2021]. De Toffoli [2021] accepts that the proof is not inconsistent with the standard view but she rejects Tatton-Brown [2019]'s interpretation which classifies De Toffoli and Giardino [2016] as an anti-formalist argument. However, De Toffoli still endorses an adapted anti-formalist thesis: "there is a reasonable way to individuate proofs such that if topological proofs involving visualization are converted into formal proofs, they are thereby transformed into a different proof" [De Toffoli, 2021, 18]. We do not yet have a description of the individuation criterion that De Toffoli mentions.

To recap, then, we see a similar field of play in both the discussions on traditional Euclidean geometry and modern knot theory. The original objection is raised as strictly anti-formalization: diagrammatic proofs cannot be translated into formal proof. This is too strong and the standard view proponent has promising examples to the contrary. The better diagram objection lies in the claim that the new formal proof is different in some important way. For both Larvor [2019] and De Toffoli [2021] there is some individuation criterion which makes the inferences or the proof different. This matters for the standard view because the informal proof must either indicate or translate into the formal proof. If the formal reasoning is so relevantly different, then justification is not a matter of recognizing routine translation. Even worse, the idea that the informal proof could indicate a vastly different formal proof is seemingly untenable.

1.3.3 The Translation Objections

One way of reading the standard view, in particular Hamami [2019]'s reading is that the connection between the informal proof and the formal proof is one of algorithmic translation. This worry does not arise exclusively for Hamami [2019], as Azzouni [2004] also relies on the notion of algorithmic systems. In this section, we look at translation objections which argue that the translation from informal to formal proof cannot be algorithmic.

Rav [2007]'s second aim is at Azzouni's notion of an algorithmic system. As we hinted at above, and as Rav argues, an algorithmic system must be formally specified so that the derivations are formal proofs. But Azzouni also wants these algorithmic systems to be easily manipulable and transcend-able to match his view of mathematical practice. Again Rav distinguishes between proofs, which is a non-technical term, and derivation which is a technical term. For a derivation, there must be a formal object language T with explicit syntactical and inferential rules. A derivation is a finite sequence of formulas in T where each formula is either a logical axiom, an axiom of T, or a result of applying one of the explicit rules of T. Given Azzouni's claims about derivations in an algorithmic system, Rav points out that the above definition of derivation is a fair characterization. But then an algorithmic system must be algorithmic, ie it must operate on a fixed formal language. The problem arises, though because Azzouni's algorithmic systems are meant to mechanically recognize the validity of informal proofs. Informal proofs do not have a fixed formal language. If the minimal requirement of operating on a fixed formal language cannot be met, then algorithmic systems are really a misnomer for a completely unspecified non-algorithmic system.

Hamami [2019] also claims that each of the translations (1), (2), and (3) described in Section 1.2 are algorithmic. The worry occurs at translation (1) which takes us from the vernacular proof to the higher-level proof. The vernacular proof much like the derivation-indicator proofs is written in English. English is not a fixed formal language with explicit syntactical and inferential rules. So, again, this translation cannot be an algorithm in the technical sense. Moreover, Hamami [2019] needs it to be an algorthmic process in the technical sense because he equates 'routine translation' with 'algorithmic translation.' Routine translation serves as the bridge between informal and formal proof for Hamami [2019]. If he cannot rely on the algorithmic component, he must introduce a new definition of 'routine,' most of which will not have the formal backing that appeals to the standard view. I'll spell out a new objection to translation-based standard views in Chapter 2.

Both Azzouni [2004] and Hamami [2019] rely on the technical nature of algorithms to argue that the translations required by the standard view are routine or identifiable. But the technical definition of algorithm requires a fixed formal language. Both authors attempt to apply algorithmic translations to vernacular, informal proofs which lead to contradictory claims about the nature of the translations.

1.4 Alternatives to the Standard View

The standard view and responses to it comprise most of the literature on informal rigor. We've seen three types of objections to the standard view which are compelling. Here we turn to a few alternative accounts of informal rigor. These accounts are typically born out of a criticism of the standard view. Moreover, these accounts are not anti-formalization accounts. Many of them admit that formalizing proofs can be a worthwhile study for both logic and proof theory. The issue arises with the further claim of the standard view which evaluates informal rigor in terms of its relationship with formal rigor.

1.4.1 The Meaning Account

Let's return for the final time, to the Rav-Azzouni debate. In addition to Rav's criticisms of the standard view, he provides his own account of informal proof and informal rigor. Unsurprisingly, the account he provides is one based on what he calls the irreducible semantic content of the proof.

The objections that Rav [1999, 2007] raises against the standard view produces a natural positive account. Rav claims that vital semantic content is lost in translating an informal proof to a formal one. Rav also claims that it is this semantic content that cements the informal proof. Given this characterization of informal proof, we also get a characterization of informal rigor. Judgments of informal rigor are related to the semantic content of the proof. Rav is not the only proponent of the view, Robinson also wrote that "unformalized proofs ... are judged to be rigorous (or not) directly, on the basis of criteria which are intuitive and semantic" [Robinson, 1997, 54]. The meaning-based account, much like the meaning-based objection, is tied to the idea that informal proofs provide a level of conceptual understanding which is lost in the formal proof.

One of the benefits of this view is that it connects a thread lost by the standard view. As we saw in Section 1.1, informal rigor was originally tied to definitions, notions, and understanding. This connection was lost by the standard view where informal rigor merely pointed to formal proof which was topic-neutral and difficult to parse. The standard view often left the understanding of concepts out of rigor's domain. The meaning account brings this back into focus. But, as Hamami [2019] points out, the meaning account is under-specified.[5] We have no clear account of what 'intuitive' or 'semantic' means in these accounts. The meaning account must answer those questions. To answer these questions, one might return to a view like Kreisel [1967] where informal notions from natural language are being examined.

The meaning account aims to incorporate some of our earliest observations about informal rigor: it's related to grasping the meaning of certain terms, it's content and context dependent, it's related to the notions, and it treats Burgess [2015]'s paradox of rigor as a significant problem. That being said, it is not entirely clear what semantic content is or how we reliably make rigor judgments on that basis.

1.4.2 The Action Account

The action account is intimately tied to the diagram objection. The core idea is that certain actions, specifically inferential actions, are rigorous. The view holds that informal rigor is determined on the basis of certain actions, not formal proofs.

Larvor [2012] is the main proponent of this view. Larvor starts with something very similar to the meaning-based claim: "the validity or invalidity of essentially informal arguments does not depend on their logical form alone, but also on their content" [Larvor, 2012,

[5]Hamami [2019] actually raises this against Robinson's version of the meaning objection. But it clearly seems to affect the positive view as well.

720]. The essentially informal arguments Larvor discusses are exactly those arguments which, when translated to a formal proof, *lose something*. According to Larvor, the content-specificity is what defines them. But, contrary to the meaning account, this is not the end of the characterization. Larvor reminds us that each move in a proof is not just a linear shift between propositions, but a genuine inferential action. This gives us the following account of rigor:

> [Informal arguments] are rigorous if they conform to the controls on permissible actions in that domain. An action demonstrating (by performance) the possibility of a new gymnastic feat had better conform to the rules of gymnastics; ice-core samples must be kept free of contamination; and so on. [Larvor, 2012, 724]

So what we have is a domain-specific set of actions used to judge the rigor of a proof. If a proof conforms to those actions, then it is informally rigorous. This account applies well to diagrammatic proof where, as we saw in Manders [2008] and Larvor [2019], rigorous diagrammatic proofs were defined in terms of actions which exhibited certain control.

One of the issues with this view is similar to the worries raised by Azzouni [2004] and Tatton-Brown [2019]. What counts as a rigorous action according to this view is defined in terms of rules. The issue is where we derive the rules. In the case of gymnastics, rules and regulations are set by governing committees. There is no such governing committee for mathematics. Mathematicians are fairly reliable at judging what is rigorous, but as Tatton-Brown [2019] points out, there's no reason to think that a community's way of reasoning is necessarily accurate. Moreover, Tatton-Brown [2019] objects to the community-based view of appropriate actions since new branches of mathematics must be introduced before there is a defined community for them.

Like the meaning-based account, the inferential-action account places judgments of informal rigor in the informal proof. Here the focus is on admissable actions defined by a community. The standard view opponent needs to clarify who is setting the standards, how they earn that power, and why those standards are reliable.

1.4.3 The Agential Virtue Account

The final alternative account is the agential virtue account proposed by Tanswell [2017]. While Tanswell [2015, 2017] both raise objections to the standard view, the agential virtue account is not born directly from these objections.

Tanswell [2017] argues that we should think of rigor as an agential virtue; in other words rigor is a characteristic of the mathematician. This is a result of his argument that we should take a virtue approach to mathematical epistemology. According to this view, it is through virtuous acts that we gain knowledge. Applying this to math, we get the uniquely virtuous acts involved in proving. Rather than permissible acts, as discussed in the last section, we now focus on virtuous acts. Rigor is a "an acquired character trait and excellence specific to mathematical practices" [Tanswell, 2017, 180]. Thus he separates rigor into three types: formal rigor of derivations in formal systems, informal rigor of communal standards and norms, and rigor as an agential virtue.

The agential virtue account raises a number of further questions. For example, what does it add to the communal or action account to say that certain mathematicians are rigorous? And how exactly is a rigorous mathematician defined, if not in terms of community standards? While Tanswell has used the overall framework to examine problems in the philosophy of mathematics, as in Tanswell and Kidd [2020] and Tanswell and Rittberg [2020], the specific aspect of rigor as a virtue needs investigation.

The goal of this chapter has been to review the literature on informal rigor. In Section 1.1, we examined some of the earlier accounts of informal rigor. We then turned to the standard view, described in Section 1.2 and objections to it in Section 1.3. We looked at some alternatives in Section 1.4. Overall, the standard view held some intuitive appeal but it has not yet fully answered the objections against it. On the other side, the alternative proposals raise a number of questions of their own.

1.5 Outline of the Dissertation

This dissertation aims to characterize a new view of rigor which adopts some of the appealing components of each of the alternative views and locates rigor in relation to conviction.

Next, in Chapter 2, I work through three general categories of standard view. I present new objections to each of the three categories. One of the main takeaways is that mathematicians are convinced of the steps of the informal proof itself, not that some other formal proof could exist. The notion of conviction is a key component of the audience view developed in Chapter 3 and Chapter 4.

In Chapter 3 I give a new account of rigor which is driven by the imagined universal audience. I argue that a proof is completely rigorous when each step is one that the mathematician's universal audience assents to. Each inference is judged to be rigorous when it convinces one's universal audience. For the mathematician, this amounts to the judgment that the inference would convince everyone. The audience view escapes the objection I posed to the standard view, since the mathematician judges that each inference is convincing, not that some other object could exist. I also argue that my account is superior to the standard view since the audience view accommodates a gradable notion of rigor. A proof is more rigorous than another when it has more inferences to which the universal audience assents.

In Chapter 4, I connect the audience view of rigor to core issues in social epistemology. The audience view claims that rigor judgments depend on social features, including who participates in mathematical practice. Given that participatory injustices seem to occur in mathematical practice, I argue a mathematician's universal audience will be influenced by those injustices. I argue that eliminating participatory injustice will lead mathematicians to have a more robust universal audience. A more robust universal audience leads to rigor judgments that are more stable over time. I argue that participatory injustice is detrimental to the proofs, as well as the participants. This provides a new reason to both eliminate participatory injustice in mathematics and to diversify the profession. In Chapter 5, I conclude and discuss some future issues that a social epistemological approach to rigor must address.

Chapter 2

New Objections to the Standard View of Rigor

In Chapter 1, we saw some introductory discussion of both the standard view of rigor and three types of objections to the standard view. In this chapter, I'll argue, as others have before me, that the standard view does not give a satisfactory account of rigor. Unlike previous discussions of the standard view, I'll use Burgess and De Toffoli [2022]'s choice points to mount objections to each possible type of standard view. The three choice points are how the informal proof is related to its formal counterpart, whether the formalization must be actual or merely potential, and whether the conversion is routine. The objections of this chapter are more general than those seen in Chapter 1 since they don't focus on any single version of the standard view.

First, in Section 2.1, I set out the target phenomenon of a theory of rigor. I'll outline some goals for an account of rigor and an example of rigorous proof. In Section 2.2, I'll turn to Burgess and De Toffoli [2022]'s summary of the standard view. They give a few choice points resulting in different versions of the standard view. In Section 2.3 through Section 2.5, I discuss each of those versions. In each section, I'll argue that each version of the standard view fails to account for how mathematicians make rigor judgments. I draw some overall conclusions about why the standard view fails and suggest we turn to alternative accounts. Such an account will be given in Chapter 3.

2.1 Target Phenomenon

There are two goals for this section. The first goal is to give an example of a mathematical proof which is rigorous. The second is to refine the goal of our study. More specifically, I'll argue that we're trying to give an account of rigor which answers two questions. First, what features of a proof makes it rigorous? Second, how do mathematicians judge that a proof has those features?

But let us begin with a proof. The proof is from [Casella and Berger, 2002, 58] and shows that the expected value of a random variable X (denoted EX) minimizes the distance $E(X-b)^2$. This fact plays an important role in the development of well-behaved estimators used in linear regression. The proof is reproduced below.

> Suppose we measure the distance between a random variable X and a constant b by $(X-b)^2$. The closer b is to X, the smaller this quantity is. We can now determine the value of b that minimizes $E(X-b)^2$ and, hence, will provide us with a good predictor of X. (Note that it does no good to look for a value of b that minimizes $(X-b)^2$, since the answer would depend on X, making it a useless predictor of X.)
>
> We could proceed with the minimization of $E(X-b)^2$ by using calculus, but there is a simpler method. ... Using the belief that there is something special about EX, we write

$$\begin{aligned} E(X-b)^2 &= E(X-EX+EX-b)^2 \\ &= E((X-EX)+(EX-b))^2 \\ &= E(X-EX)^2+(EX-b)^2+2E((X-EX)(EX-b)) \end{aligned} \tag{2.1}$$

where we have expanded the square. Now, note that

$$E((X - EX)(EX - b)) = (EX - b)E(X - EX) = 0 \tag{2.2}$$

since $(EX - b)$ is constant and comes out of the expectation, and $E(X - EX) = EX - EX = 0$. This means that

$$E(X - b)^2 = E(X - EX)^2 + (EX - b)^2. \tag{2.3}$$

We have no control over the first term on the right-hand side of [2.3], and the second term, which is always greater than or equal to 0, can be made equal to 0 by choosing $b = EX$. Hence,

$$\min_b E(X - b)^2 = E(X - EX)^2. \tag{2.4}$$

This proof is considered rigorous by mathematicians. The textbook from which it's copied is commonly used as a theory-heavy introduction to mathematical statistics. The proof requires some background knowledge including for example the linearity of expectation, that the expected value of a constant is the constant itself, some notational facts, and how to expand a square. But it is overall a simple and straightforward proof, especially when compared to a calculus-heavy counterpart. What we'll see later is that this proof turns on the clever (but not wildly inventive or insightful) decision to add and subtract EX since we believe "there is something special about EX."

With a rigorous proof directly in our minds, let's now look at what a theory of rigor aims to provide. The mathematician, at the end of the proof, knows that $\min_b E(X - b)^2 = E(X - EX)^2$. Philosophers of mathematics, I take it, are interested in how following the proof can justify the mathematician's knowledge of the theorem. We're not just interested in the mere fact that the truth of the premises guarantees the truth of the conclusion, as that's validity. We're interested in how mathematicians come to know that the conclusion is true. Informal rigor is the element required to have, and share, mathematical knowledge in practice. It governs claims to justification in modern practice. The epistemological aim of giving or reading a proof is to establish the conclusion with as much certainty as the axioms.

Informal rigor, then, seems deeply tied to mathematical knowledge. If that's true, then what we're seeking in a theory of rigor is a theory which makes sense of rigor in practice. We want a theory of rigor that does not just describe which proofs are rigorous or prescribes which proofs should be rigorous. We want a theory of rigor that explicates mathematician's judgments of rigor. In other words, a theory of rigor which only answers "what feature(s) of the proof make it rigorous" is not complete since it fails to explain how rigor is judged, and therefore, fails to explain how knowledge is generated in practice. A theory of rigor must also include a plausible story about how mathematicians judge a proof to be rigorous.

In this section I hope to have motivated two core assumptions. The first core assumption is that a theory of rigor ought to identify the feature(s) that makes proofs like the Casella and Berger [2002] one rigorous. Second, a theory of rigor ought to account for how mathematicians make that judgment. In the following sections, I'll argue that the standard view's answer to the first question precludes a satisfactory answer to the second question.

2.2 A Picture of the Standard View

The goal of a study into rigor is to determine two things: what makes a proof rigorous and how mathematicians judge the rigor of proofs. This chapter focuses on a subset of answers to the first question often called *standard view of rigor*. This section focuses on examining the overall structure of the standard view and the theses it requires. To do so, I follow the survey provided in Burgess and De Toffoli [2022]. In the sections that follow, I'll look at more detailed accounts falling under the standard view.

Burgess and De Toffoli [2022] formulate the standard view as follows:

> STANDARD VIEW: A mathematical argument is a rigorous proof if and only if it can in principle be converted into a formal derivation.

where a formal derivation is a derivation in some suitable formal deductive system. This description of the standard view is sometimes called a formalization thesis – that every rigorous proof has a formal counterpart. Of course, 'formal counterpart' can be spelled out in different ways. Likewise, as Burgess and De Toffoli [2022] note, this definition is incomplete. The phrase "it can in principle be converted" requires definition and refinement. Burgess and De Toffoli [2022] raise three choice points. The answer to each results in different formulations of the standard view. The first question, (A), is "how are the steps of the informal proof and the ones of its formal counterpart related?" The second question, (B), which we'll largely set aside for this paper, is whether the formalization is actual or merely potential. The final question Burgess and De Toffoli [2022] raise is whether the conversion of a proof consists in a routine translation. The third question, (C), is often related to the first; for example, in Hamami [2019] routine translation answers how the steps of an informal proof are related to a formal one.

The second question is a red herring for this debate. It seems obvious that there are many rigorous, informal proofs which do not yet have formal counterparts. Likewise, mathematicians do not seem to require a formal proof to judge the rigor of an informal proof. I am unsure of whether any fully formal proof exists of Casella and Berger [2002]'s proof that $\min_b E(X - b)^2 = E(X - EX)^2$. But I don't need to find one in order to determine whether the textbook proof is rigorous. Focusing on which proofs have formal derivations and which merely have formal derivations *in principle* sidetracks us from the more philosophically rich debates in Burgess and De Toffoli [2022]'s questions (A) and (C).

Question (A) focuses on how the steps of the informal proof relate to those of the formal proof. Burgess and De Toffoli [2022] offer the two following possible answers:

1. (A.1) The formalized versions of the steps of the informal proof are components of the formal proof. Therefore, the process of formalization is (apart from transcription into special symbols) one of filling in.

2. (A.2) Parts of the informal proof, such as passages of reasoning with diagrams, though converitble into a sequence of formal steps, are not themselves expressible as informal counterparts of propositions of any of the usual formal systems.

The choice between (A.1) and (A.2) generally lies in one's interpretation of diagrammatic proof. De Toffoli [2021], for example, argues that some parts of diagrammatic proof are not expressible as informal counterparts of propositions in the formal derivations. If true, these arguments put serious pressure on (A.1), though there are many philosophers who believe that the informal diagrammatic inferences are, in a suitable formal system, the components of the formal proof. These philosophers are bolstered by work such as Avigad et al. [2009]'s formal system for Euclid's *Elements*. As Burgess and De Toffoli [2022] note, an account of rigor shouldn't rule out entire swathes of mathematical proofs. So whichever choice one makes with respect to Question A, it ought to allow for rigorous diagrammatic proof.

The second question I'll focus on is Question (C). (C) asks whether routine translation is the process by which informal proofs are converted to formal proofs. There are two options which cover all possible responses.

1. (C.1) Converting a proof into a formal proof is done by a *routine translation*.

2. (C.2) Converting a proof into a formal proof requires more than a *routine translation*.

As we'll see in Section 2.3, there are multiple ways of interpreting routine translation. But the general idea is that a routine translation is straightforward, simple, or involves very little invention. We shouldn't interpret routine as being synonymous with fast or short. As with many algorithms, there may be many steps which take a long time to perform. But the nature of the steps is routine and non-creative. We can think of routine translation as being similar to the work of Steiner [1975]'s logician midwife who never adds to the creative process but helps the mathematician birth his proof.

The decision between (C.1) and (C.2) is related to the decision made in (A). As Burgess and De Toffoli [2022] themselves note, the choices (A.2) and (C.1) seem incompatible. If we view routine translation as a simple, algorithmic, or non-inventive process, then a routine translation is incompatible with the view that parts of the informal proof are inexpressible in the formal system. The reasoning of the formal proof would have to supplant the original reasoning through a creative, or non-routine method. In other words, inventive, creative, or heuristic work is necessary for converting the inexpressible.

Setting aside responses to (B) and the incompatible combination of (A.2, C.1), we have three general routes a standard view can take. The three routes are characterized in the following way:

1. (A.1, C.1) The formalized version of the steps of the informal proof are components of the formal proof. Formalization is a process of *filling in*. Moreover, *filling in* is performed by a routine translation.

2. (A.1, C.2) The formalized version of the steps of the infomal proof are components of the formal proof. Formalization is a process of *filling in*. But the process of *filling in* requires more than a routine translation.

3. (A.2, C.2) Parts of the informal proof are convertible into a sequence of formal steps. But these parts aren't expressible as informal counterparts of propositions in any usual formal system. Moreover, the formal proof is reached by a non-routine translation.

This leaves us with three versions of the standard view to discuss: (A.1, C.1), (A.1, C.2), and (A.2, C.2). Burgess and De Toffoli [2022] point out that there might be more nuanced answers to either of (A) or (C). But they don't point to authors with more nuanced answers and, to my knowledge, the routes characterized are representative of the positive literature on the standard view. If each faces significant objections, then it will be worthwhile to explore alternative views of rigor in more detail. This is what I'll focus on in the rest of this chapter.

The purpose of this section has been to give a picture of the standard view of rigor using Burgess and De Toffoli [2022]'s choice points. In the following sections, I look at the three possible views and argue that they fail to account for mathematical practice. To do so, I'll take a published version of the view as an example but draw more general conclusions based on the generic features of the view.

2.3 Filling in the Steps by Routine Translation

Classification (A.1, C.1) is one of the most appealing versions of the standard view for proponents of formal rigor since it provides the tightest connection between informal rigor and formal derivations. It's also the most stringent version since it places requirements on both the informal proof and the process of translation. The (A.1, C.1) version of the standard view has the following consequence: an informal proof is rigorous iff each of the steps in the informal proof can be *routinely translated* into those of a formal derivation. For the rest of this section, I'll sketch Hamami [2019]'s view which sees routine translation as an algorithmic translation. Ultimately I'll argue that views of the (A.1, C.1) variety fail to account for the first step in translation – that of filling in the gaps via informal sub-proof.

Hamami [2019] separates an account of rigor into three parts: a descriptive account, a normative account, and a linking thesis. Here is how Hamami describes the difference:

> A *descriptive account* of mathematical rigor provides a characterization of the mechanisms by which mathematical proofs are judged to be rigorous in mathematical practice; a *normative account* of mathematical rigor stipulates one or more conditions that a mathematical proof ought to satisfy in order to qualify as rigorous. [Hamami, 2019, 3]

In other words, a descriptive account is one that explains how proofs are judged to be rigorous in practice while a normative account tells us what a proof ought to satisfy in order to qualify as rigorous. A proof which is rigorous according to a descriptive account is $\textit{rigorous}_D$ while a proof which is rigorous according to a normative account is $\textit{rigorous}_N$. It is important to note that the descriptive accounts still involve evaluation. Rigorous_D judgments are evaluative since they might exhibit implicit rules governing rigor judgments in practice. These norms might be built into the mechanisms by which proofs are judged to be rigorous.

Hamami's descriptive account of rigor aims to give an account of the mechanisms by which mathematicians judge proofs to be rigorous. Each proof P is composed of a set of inferences I. Each mathematician is engaged in a mathematical practice M. A proof is rigorous when it has been completely verified, i.e., all of the inferences it contains are verified. According to Hamami, giving a descriptive account of rigor amounts to giving an account of how proofs are verified to be valid within the mathematical practice.[1]

According to Hamami, judging a proof to be rigorous involves two kinds of processes. The first set of processes involved is decomposing a proof into a set of immediate mathematical inferences. A mathematical inference is immediate for a given agent if she can evaluate it as valid without introducing intermediate steps of deduction. For example, instances of modus ponens are immediate inferences. Once decomposed, the proof is verified according to verification processes. Thus, letting D_M be the set of decomposition processes and V_M be the set of verification processes, we have the following re-formulation of a descriptive account. The set V_M consists of processes V that map an inference I to either valid or invalid.

A mathematical proof P is rigorous$_{D,M}$

$$\Longleftrightarrow$$

For every mathematical inference I in P there exist $D \in D_M$ and $V_1, ..., V_n \in V_M$ such that (1) $D(I) = \langle I_1, ..., I_n \rangle$ and (2) $V_i(I_i) =$ valid for all $i \in [1, n]$.

Providing a full descriptive account of rigor amounts to providing an account of D_M and V_M. As Hamami notes, the decomposition processes D_M "required to turn the mathematical inference $P_1, ...P_n \rightarrow C$ into a sequence of immediate mathematical inferences is identical to the proof search process required to prove the mathematical proposition 'if $P_1, ..., P_n$ then C.' Decomposition processes are therefore proof search processes" (15). Drawing on Fallis [2003]'s discussion of enthymematic gaps, Hamami places two further constraints

[1]Note that this equates rigor to successive judgments of validity. This is a substantive thesis. One prima facie reason to doubt such a thesis is that rigor judgments seems gradeable – some proofs are more rigorous than others – but validity is not.

on D_M. First, the proof search processes in D_M must be part of common background knowledge of the practitioners of M. Second, the proof search process must be completable in a 'reasonable amount of time.' So according to Hamami, decomposition processes are just a subset of proof search processes.

According to Hamami, the standard view's V_M lies in higher-level rules of inference or, as he calls them, hl-rules. An hl-rule is determined entirely by its premise schema and its conclusion schema. These schemas are relative to the language of a mathematical practice M. An immediate mathematical inference is valid whenever it corresponds to an instance of an hl-rule. So each hl-rule R has a corresponding verification process V_R. Specifying V_M is done by characterizing the set of hl-rules that an agent in M acquires during their training.

An agent begins her training in M with only the set of propositions accepted without proof by M. These are identifiable by looking at the elementary textbooks in a practice M and identifying what is accepted without proof. In the case of Casella and Berger [2002], the axiom of countable additivity (or finite additivity, depending on M) would be such an example. The agent is also equipped with starting rules of inference. These rules are "essentially basic rules of elementary logical reasoning necessary to reason with mathematical propositions" [Hamami, 2019, 18]. Hamami leaves open what the basic rules of elementary logical reasoning are. Whenever the trainee derives a proposition from the set of propositions she already accepts, she may add the new proposition to the set of mathematical propositions she knows. Likewise, she can update her hl-rules by turning deductions into new hl-rules. Whenever the mathematician "has derived a mathematical proposition C from a set of mathematical propositions $P_1, ..., P_n$ through a sequence of applications of hl-rules [she already possesses] ... she is entitled to add to her set of hl-rules ... the new rule: $P_1, ..., P_n \to C$" [Hamami, 2019, 19]. Through this process, theorems and definitions also quickly become hl-rules in the set V_M.

For example, $E[aX + b] = aEX + b$ when a, b are constants, could become an hl-rule through this process. The statement is not an immediate inference. But it can be proven by a student using the definition of expected value, the definition of probability distribution and mass functions, and some facts about integration.[2]

By specifiying the sets D_M and V_M, Hamami has given a descriptive account of rigor that explains the methods mathematicians use to evaluate proofs as rigorous. The descriptive view we've discussed so far focuses on (A.1) – how mathematicians fill in the gaps of an informal proof through decomposition and verification. The next step is to specify how (C.1), the routine translation, works. Accordingly, Hamami gives a normative account of rigor which focuses on routine translations that practitioners can undertake. The standard view normative account is:

A mathematical proof P is rigorous$_N$

$$\iff$$

P can be *routinely translated* into a formal proof.

To give a full formulation of the normative account, Hamami must explain what a routine translation is. This is given in terms of a successive series of algorithmic translations through levels of proof granularity.

There are four levels of granularity. The first is the vernacular-level proof. A vernacular-level proof is the general mathematical proof seen in textbooks and journals. It is a series of inferences as they are commonly presented in mathematical texts of the practice M. The Casella & Berger proof in Section 2.1 was a vernacular-level proof. The inferences of

[2]It's fun to note that the proof will most likely invoke the proposition that $E[g(X)] = \int_R g(X) f_X(x)$ where $f_X(x)$ is the pdf of X. This proposition (and its discrete countepart) is often called the "law of the unconscious statistician" which yields some humorous evidence of learned hl-rules.

a vernacular-level proof are not all immediately verified for a mathematician in M. They often must be decomposed which results in the second level of granularity. A higher-level proof is a sequence of inferences which are all hl-rules. There are two more levels of granularity – the intermediate level and the lower level. The intermediate-level proof is comprised of primitive rules of inference and primitive axioms. Thus, the intermediate-level proof only contains inferences and axioms that mathematicians possess at the start of their mathematical training. The lower-level proof is a sequence of inferences comprised of rules of inference and axioms from a formal deductive system adequate for the foundations of mathematics.

For Hamami, routine translation amounts to an algorithmic translation among the four levels of granularity. The first translation, from vernacular-level to the higher-level proof, "corresponds exactly to the first phase of the process that a typical mathematical agent engages in when judging the rigor of a mathematical proof, namely the decomposition of each inference in the proof that cannot be verified directly into a sequence of immediate mathematical inferences" [Hamami, 2019, 22]. This has a corresponding *algorithm* $\tau_{vl\to hl}$ for the standard view's D^*_M and V^*_M which has three steps:

1. It identifies a decomposition process $D \in D^*$ such that (1) $D(I) = \langle I_1, ..., I_n \rangle$ and (2) there exist $V_1, ..., V_n \in V^*$ such that $V_i(I_i) =$ valid for all $i \in [1, n]$.

2. It decomposes I into the sequence of inferences $\langle I_1, ..., I_n \rangle$ using the decomposition process D.

3. It replaces I in P by the sequence of inferences $\langle I_1, ..., I_n \rangle$.

There are two similar algorithmic translations. One takes us from the higher-level proof to the intermediate-level proof and the other takes us from the intermediate-level proof to the lower-level proof.

For the routine translation view to succeed, the mathematician must be able to succeed at performing the routine translations. I want to focus in particular on the translation $\tau_{vl\to hl}$. I will argue that this algorithmic translation is not compatible with mathematical practice and thus, fails as an account of rigor. When discussing both $\tau_{vl\to hl}$ and its role in proving the conformity thesis, Hamami claims that the translation corresponds to the decomposition and verification processes which typical mathematical agents are capable of performing. Additionally, this translation must be *algorithmic*. Hamami, in replying to critics of the standard view, argues that routine translation is *algorithmic*, not simple, easy, etc. Without the claim that these translations are algorithmic, then, the standard view does not have claim to "routine" and remains plagued by the objections of Robinson [1997], Detlefsen [2009], and Larvor [2012].

The problem with the algorithmic translation $\tau_{vl\to hl}$ is apparent when we return to Hamami's claim that decomposition is proof search. This one translation must live up to three demands: it must be algorithmic, it must be proof search, and it must be done by actual mathematicians. An algorithm is a finite sequence of instructions. It is a precise set of steps applied on an input. The claim that translation $\tau_{vl\to hl}$ is algorithmic amounts to the claim that searching for a sub-proof is algorithmic. But proof search, in general, is not algorithmic. Proof search is most often described as a heuristic process. Pólya [1945]'s famous discussion of problem solving highlights the importance of heuristic. According to this view, the process of finding a proof invokes heuristic problem-solving techniques. Heuristic reasoning involves pragmatic or imaginative strategies which are not captured by a specified set of rules. For example, reasoning by analogy or through an example might be heuristc proof techniques. Additionally, consider reflective discussions of mathematical practice by Villani [2015], van der Waerden [1971], Hadamard [1945], etc. All of these discussions describe proof search as a messy and sometimes dialogical process.

As a concrete example, consider the proof in 2.1. The decision to add and subtract EX was based on the hypothesis that "there is something special about EX." Likewise, similar minimization problems involving sums of squares in probability and statistics are solved by using the same method of adding and subtracting some function of the mean. The budding statistician searches for those proofs using heuristic methods like analogy and comparison. The 'trick', once learned, can occasionally be used in some minimization problems. But the 'trick' hardly constitutes a precise set of steps that applies to every minimization problem. So the processes by which mathematicians search for proofs like this do not seem to be algorithmic. But the descriptive account given by Hamami states that that decomposition is proof search and proof search is algorithmic. This is incompatible with how mathematicians reason through proofs. Thus, Hamami's account of the standard view fails at its first level of translation.

Without an argument that proof search is an algorithmic process, and mathematicians are simply wrong about their decomposition processes, there is no reason to believe the first step of translation is an algorithmic process. In other words, the first step in "filling in the gaps" is non-routine. It's the creation of a new sub-proof. This translation may be routine in the cases where the proof of interest is not at the vernacular level. Then the proof search process will occur in a highly restricted and regimented context. But in general, the search for sub-proofs is not highly restricted. For example, in the Casella & Berger proof, the search for a sub-proof that $E(EX - b)^2 = (EX - b)^2$ will not be any easier (or harder) than when that proposition is presented as its own theorem.

While I've focused on Hamami [2019]'s account, the objection applies to all accounts satisfying (A.1, C.1). Almost every informal proof requires some amount of "filling in the gaps" to check if the proof is rigorous. The first level of "filling the gaps" is producing any necessary sub-proof. But to produce a sub-proof is to just find a proof. Finding a proof in practice is often a creative, heuristic process, not an algorithmic one. Proof search may one day be an algorithmic practice done by computers or augmented humans. But we began our investigation by stating our interest in exploring what it means for modern, mathematical proofs to be rigorous in practice.

2.4 Filling in the Steps by Non-Routine Translation

In the last section, I focused on Hamami [2019]'s version of the standard view which is representative of position (A.1, C.1). According to those views, a proof is rigorous iff it can be converted into a formal derivation by filling in the steps of the informal proof via routine translation. I argued that the problem with such views was that a routine translation is incompatible with the first translation step which is finding a sub-proof. In this section, I turn to two representatives for position (A.1, C.2). These views hold that a formal proof is rigorous iff it can be converted into a formal derivation by filling in the steps of the informal proof. The procedure for filling in the steps may be non-routine or creative. The limitation is that the steps of the informal proof are expressible components of the formal derivation. I think there are two representatives of this view – Azzouni [2004, 2009]'s derivation-indication view and Tatton-Brown [2019]'s view following Burgess [2015] about conviction.[3]

2.4.1 Derivation-Indication

One of the earliest, and most frequently cited, formulations of the standard view is Azzouni [2004]'s derivation-indicator view. Azzouni argues that it is derivations in an 'algorith-

[3]It's worth noting that in Azzouni's most recent work, like Azzouni [2020], Azzouni has retained the view that algorithmic processes are essential to rigor but attempted to separate it from the standard view of rigor.

mic system' which are characteristic of mathematical practice and, in particular, explain widespread mathematical consensus. An algorithmic system is a system which codifies the deductive rules for derivations such that "the recognition procedure for proofs is mechanically implementable" [Azzouni, 2004, 83]. Azzouni admits that these derivations are not the currency of mathematical practice. The informal proofs are successful, according to Azzouni, because they are *indicators* of a corresponding derivation. In other words, informally rigorous proofs are proofs which indicate the existence of a formally rigorous derivation. Azzouni asks quite a bit of his algorithmic systems – they are algorithmic but they are not restricted to a particular logic, subject-matter, or explicit language. He does this to ensure mathematicians can be seen as "sprinting up and down algorithmic systems, many of which he or she invents for the first time" [Azzouni, 2004, 103]. Notice that an algorithmic system is not an actual algorithm. An algorithm is a procedure. An algorithmic system is something of Azzouni's own invention.

Whatever one may say of the algorithmic system, Azzouni must also tell us how mathematicians manage these systems and indicate the existence of derivations. This is particularly difficult since, according to Azzouni, mathematicians need not know the rules codified in an algorithmic system. To do so, Azzouni [2005, 2009] invokes the idea of an 'inference package' which is "capacity to recognize the implications of several assumptions by means of the representations of objects wherein those several assumptions have been knit together (psychologically)" [Azzouni, 2009, 20]. These packages allow mathematicians to do what they normally do in informal proofs – recognize and tease out the implications of their assumptions without being aware of a formal derivation. Luckily, our inference packages are rich enough to accommodate a wide set of algorithmic systems with which they allow us to reason compatibly. And by filling out the inference packages, we can, in principle, produce the formal derivation.

Let's return to the Casella & Berger proof to ground the derivation-indicator view. We said before that finding this proof would be non-routine since there's a 'trick' where we add and subtract EX within the squared term. Inference packages, it seems, rely heavily on past experiences. So one can imagine that the professor teaching this proof knows very well that the trick should be applied here. The professor, when asked, might say "this is because EX is special so we want to get it inside the squared term." This would be an example of a psychologically knit-together inference package. Knowing EX is special seems to automatically get the professor to adding and subtracting EX.

Under the derivation-indicator view, proofs are rigorous because they can be filled out into a formal derivation in an algorithmic system by means of inference packages. Inference packages, as implications psychologically knit-together, are most likely not algorithms. If they are algorithms, then we will just return to the objections of Section 2.3. This is what makes the derivation-indicator view a position in the space of (A.1, C.2). The problem with inference packages is that many psychologically knit-together inferences are simply not rigorous. The view over-generates. Think, for example, of affirming the consequent. Anyone who has taught first-year logic knows that it's difficult to train students out of using it. One reason for this might be that affirming the consequent is reminiscent of a nearby inductively strong IBE argument. IBEs are successful psychologically-knit together inference packages. But no argument via IBE is a rigorous proof. So the derivation-indicator view needs a strong, sufficiently robust notion of inference package to ensure there's no over-generation.

The over-generation concern above focuses on the inference packages of people broadly. But focusing on only mathematicians does not yield a better result. Consider the mathematician Srinivasa Ramanujan who made significant contributions to mathematics without formal training. In order to do so, it seems that Ramanujan must have had many well-knit psychological inferences about mathematical concepts. Nevertheless, he failed to produce rigorous arguments for his theorems. Again, just because the inference packages were available to him, that didn't make his arguments rigorous.

Setting aside the specificities of the derivation-indicator account, it seems to point to something important about training. Mathematicians seem to get better and better at producing rigorous proofs with lots of hidden details. Mathematicians seem to be convinced that they can fill in any gaps. Perhaps inference packages are too broad to fill in the details. So in the next account, we look at a more compelling account of knowing that one can "fill in the gaps non-routinely."

2.4.2 Conviction

Azzouni's inference packages are one non-routine route to "filling in the gaps" in an informal proof. If there is a routine filling-in procedure, then it either must be algorithmic which we've discussed above or the routine procedure must be non-algorithmic and also non-creative. Someone may spell out such a view but it's not yet on offer for us to discuss. But view (A.1, C.2) encompasses a number of options for non-routine translations since there are many non-routine ways to reach a formal derivation. Rather than spelling out non-routine methods for filling in gaps, we can instead focus on the question of how mathematicians know that the gaps *can* be filled in. So then how does a mathematician judge this? She cannot know that the gaps can be filled in algorithmically. She must judge just that the gaps can be filled in *somehow*. In Burgess [2015]'s words, "what rigor requires is that each new result should be obtained from earlier results by presenting enough deductive steps to produce conviction that a full breakdown into obvious deductive steps would in principle be possible" [Burgess, 2015, 97]. If this is to be understood as a breed of the standard view in Burgess and De Toffoli [2022], then what a mathematician needs to judge a proof as rigorous is *to be convinced that a formal derivation could be reached by filling out the steps in the informal proof*.

The obvious question arises: how do mathematicians come to know what can be filled out? Tatton-Brown [2019] sketches an account of how mathematicians learn that informal proofs can be converted into formal derivations. He argues that mathematicians are introduced to proofs at the least granular level, i.e. where all of the immediate inferences are explicitly shown. Proofs at the earliest stage rely on only the most obvious deductions as in a formal derivation. Then, as mathematicians proceed in their training, proofs become more coarse. Students prove things over time and add to the list of acceptable higher level rules. Nevertheless, students know that if they enquire into a step, they can also break it down into a greater level of granularity. They proceed in this manner with larger gaps between steps and with the faith that, if they must, they could produce the most detailed proof.

Let's return to the Casella & Berger example from Section 2.1. According to Tatton-Brown [2019], the proof is not one that mathematicians would experience at the very start of their training since it includes higher-level inferences. In particular, the final equality of equation (2.1) involves dropping the expectation of $(EX - b)^2$. In other words, one must know that $E(EX - b)^2 = (EX - b)^2$ since $EX - b$ is a constant. Rather, it's an example of a proof where the mathematician judges it to be rigorous because she could "fill in the gaps" (in a non-algorithmic way). A student can provide that proof in whatever way she finds feasible whether that's algorithmic or heuristic. Moreover, to check if the proof is rigorous, she doesn't have to find that sub-proof. She just needs to be convinced that she could find the sub-proof and *that the sub-proof could be a maximally detailed formal derivation*. Without the second clause in italics, the view is not a standard view because it doesn't relate a rigor judgment to a formal derivation.

The problem with the conviction views discussed here is that the mathematician must be convinced of something very specific. She must be convinced that a full formal derivation could be given by filling out the steps of the informal proof. I agree with Tatton-Brown [2019] that mathematicians, through training, come to believe that the gaps in any rigorous proof can be filled in. But there is an important point that Tatton-Brown [2019] seems to miss: mathematicians fill gaps with more informal proofs. They don't fill gaps with formal derivations. In other words, mathematicians are not convinced that they can fill gaps with formal derivations, they're convinced they can fill gaps with more informally rigorous proofs. Thus the training model described by Tatton-Brown [2019] doesn't support the standard view unless students start their training with formal derivations. But that is an empirical claim and it is false.

Intuitively, for a person to be convinced that X could become Y, they must know what Y is. To be convinced that a bunch of metal can become an engine, I have to know something about the makeup of engines. I don't have to make the engine to be convinced of its possibility. But I must know at least what the outcome should be and what it's made of. Likewise, if mathematicians judge rigor in a similar fashion, then they ought to know what a formal derivation is and what is required of one. But mathematicians are rarely trained in a derivation-forward manner. They don't start their educations with logic classes or discussions of formal derivations. As Tatton-Brown [2019] notes, most mathematicians begin training in proofs during their real analysis courses. The proofs mentioned foremost in Abbott [2015], which is what Tatton-Brown [2019] focuses on, are not formal derivations. The informal, mathematical training is the primary source of training regarding rigor. The issue of training in formal derivations is even clearer when we consider historical proofs. Under this standard view, for ancient Greek mathematicians to claim that a proof was rigorous, they had to be convinced that a formal derivation was in principle possible. This simply could not have been the case since the concept of formal derivation being invoked by the standard view is a definition of the 20th century.

There's an interesting rejoinder to the objection I've just raised. I'll briefly sketch it. The argument is that historical mathematicians judged things to be rigorous while lacking a fully explicated understanding of rigor. The argument compares rigor to things like gold. Ancient scientists had tests for determining whether a substance was gold. But they didn't know that the real test for gold was whether the substance had atomic number 79. The tests were fairly reliable, though, and so when ancient scientists referred to gold, they were referring to the substance with the atomic number 79 without knowing about atomic numbers. The analogous case would be that mathematicians didn't have the modern concept of formal derivation. But they were reliably referring to it by tests of rigor through other means. I think there's a problem with this objection. Gold is not like rigor because there are patterns of deference associated with scientific natural kinds that are not associated with rigor. Ancient scientists had tests for whether something was gold, but they left open for future scientists the nature of gold. Likewise, jewelers employing tests for gold will defer to scientists for the ultimate decision of whether a substance is gold in virtue of its atomic number. But this is not the case in mathematics. Mathematicians regarded Euclid's *Elements* as a paragon of informal rigor, not as an example of superficial qualities. For the analogy to go through, we'd expect mathematicians to defer to logicians regarding rigor in the same way jewelers defer to scientists regarding gold. Again, this prediction is empirically false. Mathematicians converge on whether a proof is rigorous without deference to logicians. When Mochizuki 'proved' the *abc* conjecture, it was evaluated by mathematicians. Mathematicians Scholze and Stix, both arithmetic geometers, not mathematical logicians, were the ones to argue that Mochizuki's proof was not rigorous. If a massive, important chunk of metal was purported to be gold by some jewelers, the issue would be settled by sending it to scientists who can check its atomic properties. But a similar process is not observed in cases of mathematical uncertainty. Without the above sketched deference pattern, I think we can reject the analogy between tests for mathematical rigor and tests for gold.

In other words, the conviction account gets one thing right – mathematicians often judge that a gappy proof is rigorous if they can back it up with a sub-proof. And mathematicians are trained to believe they can fill the gaps in. But mathematicians are convinced about being able to provide informal sub-proofs. Just because they're convinced they can fill in some detail for specific audiences (like filling in extra details for a reviewer or a student), that does not mean they're convinced a formal derivation is possible. This is in part because they're simply not trained in providing formal derivations of informal proofs. That cannot be the content of their conviction. Standard views of the (A.1, C.2) variety fail because the mathematician isn't judging that a formal derivation is possible, they're judging that futher informal proofs can be supplied.

2.5 Bare Conversion

The final variety of standard view we'll consider falls into (A.2, C.2). An informal proof is rigorous iff it can be converted into a formal derivation but the inferences of the informal proof may not be expressible as transitions among the propositions in the formal derivation. In other words, the crucial inferences of the informal proof may not be in the formal derivation and vice versa. The motivation for such a view is that the reasoning involved in diagrammatic proofs is often difficult to faithfully express in a sentential way. De Toffoli and Giardino [2014], De Toffoli and Giardino [2016] gives examples of such proofs in knot theory and topology. I call the (A.2, C.2) view bare conversion since it places no limitation on the process of conversion (it may be non-routine) and it also places no limitation on fidelity. Some of the steps of the informal proof may be abandoned in favor of a formal derivation proving the same conclusion. Under bare conversion, an informal proof is rigorous just in case it is convertible in any way into a formal derivation.

The problem with bare conversion is that it wildly over-generates. The bare conversion view doesn't require a special procedure that takes you from the informal proof to the formal derivation. Consider the following 'proof' that $\sqrt{2}$ is irrational.

Proof. The proof runs as follows. If $\sqrt{2}$ is rational, then the equation

$$a^2 = 2b^2 \tag{2.5}$$

is soluble in integers a, b with $(a, b) = 1$. This produces a contradiction. So $\sqrt{2}$ is irrational. □

There are formal derivations that show $\sqrt{2}$ is irrational. By a very non-routine translation, this 'proof' can be converted into that formal derivation. By the bare conversion view, this proof is rigorous. But no mathematician would consider such a proof to be rigorous. The conviction account is immune because it still requires enough of the steps of the formal derivation be revealed in the informal proof. But choice point (A.2) is built to allow massive components of diagrammatic proofs to be replaced by different reasoning in the formal derivation. So there's no requirement that enough of a formal derivation be shown in the informal proof. Without such a requirement, the above 'proof' should be rigorous. There is a formal derivation of the proof that $\sqrt{2}$ is irrational. So a conversion is possible. The conversion process is simple – replace the contents of the proof with the contents of the derivation.

The bare conversion view fails to be extensionally adequate regarding rigor judgments. Any informal proof would be judged to be rigorous once a known formal derivation exists. It's open to proponents of the standard view to sketch an account satisfying (A.2, C.2) which doesn't amount to bare conversion. But such a view is currently not in the literature. For now, I plan to set it aside and discuss the overall results of this study into the standard view.

2.6 Concluding Remarks

I took Burgess and De Toffoli [2022]'s survey of the standard view as a starting point for discussing the types of standard view. They outlined three choice points. I then argued against the prominent views in each of the resulting choice points. If the list is exhaustive, it should be clear that the standard view has failed to account for rigor judgments in practice and that a new view is called for. There may be room for further nuance within the views. The arguments I've given have varying degrees of generality with respect to nuance.

The arguments I've sketched in Section 2.3 are quite general. The process of conversion will always first involve generating sub-proofs. Generating sub-proofs is a heuristic and creative process, not an algorithmic or routine one. I did not address the possibility that one day we might have an empirically adequate theory under which proof search is routine. This seems to be the same future where automated theorem proving is the norm. That discussion is outside the scope of this paper and far in the future, if it's even possible.

In Section 2.4, my arguments were general with respect to conviction. Any view where the proof must be judged as rigorous by reference to the in-principle existence of a formal derivation will fail to account for historical and pedagogical considerations. It fails to explain mathematical behavior and deference patterns. There is plenty of room for new, nuanced views within (A.2, C.2), though. My arguments may not apply to more nuanced versions which avoid bare conversion.

Overall, I have aimed to survey some new reasons why the standard view provides an inadequate account of mathematicians' rigor judgments. A proponent of the standard view may respond by sketching a new method of indication or translation. But there will likely be historical, diagrammatic, and pedagogical objections when we try to explicate rigor judgments in terms of formal derivations. Given this, in Chapter 3 I'll provide an account of rigor which focuses on the conviction that mathematicians have in the informal proofs without reference to any formal proof.

Chapter 3

An Audience View of Rigor

The purpose of this chapter is to provide a new alternative to the standard view. The alternative view I give here will be called the *audience view*. The core motivation for the audience view is that a rigorous proof is one that convinces the right kind of audiences. The tagline ‘a proof is that which convinces’ is sometimes mentioned, and quickly dismissed, by proponents of the standard view. It is quickly dismissed since there are many things which are convincing – including testimony – but not rigorous proof. Likewise, genuine proofs don’t always convince real people on first exposure. Nevertheless, there seems to be something right about the tagline. Conviction even lurks in the background of some standard views like the one presented by Burgess [2015]. But the standard view seems to get the object of conviction wrong. Mathematicians aren’t convinced some other proof could be written. They’re convinced by the proof before them. So a modification of the standard view wouldn’t make more sense of the intuition that conviction and rigor are related. The standard view’s main focus is formal derivations. Similarly, the main focuses of the alternative views mentioned above are, respectively, the meaning of mathematical concepts, acceptable actions, and the virtues of a mathematician. Prima facie, none of those three help make sense of the intuitive connection between conviction and rigorous proof. So a modification of them will not help make sense of the tagline.

Now the goal is to give an account of rigor that makes sense of the idea that conviction plays a role in rigor. First, in Section 3.1 I review two examples of proofs which mathematicians and philosophers of mathematics judge to be rigorous. Then in Section 3.2, I'll make some over-arching remarks about what a theory of rigor might attempt to do. This will be vital to frame Section 3.3, in which I characterize the different claims of the audience view. The goal of this chapter is not to give necessary and sufficient conditions for a proof to be rigorous. Rigor, as it's used in practice, is open-textured and sometimes vague. The aim of this paper is to provide a useful explication of the concept. The audience view will be broken into three related components. Having provided the audience view in Section 3.3, I'll turn to arguments in favor of the audience view in Section 3.4. I'll argue that the audience view outperforms the standard view with respect to some intuitive conditions on a theory of rigor. The audience view, then, should be a serious contender against the standard view of rigor.

3.1 Target Phenomenon

The goal of this section is to provide two examples to reference. These two examples are chosen because they seem intuitively correct and rigorous, they are often cited in the literature, and both have been the subject of *formalization*. The first example is from a number theory textbook and the second example is from a book of mathematical games.

The first example is one of [Hardy and Wright, 1975]'s proof that $\sqrt{2}$ is irrational.

Theorem 1. *(Pythagoras' Theorem). $\sqrt{2}$ is irrational.*

Proof. The traditional proof ascribed to Pythagoras runs as follows. If $\sqrt{2}$ is rational, then the equation

$$a^2 = 2b^2 \tag{3.1}$$

is soluble in integers a, b with $(a, b) = 1$. Hence a^2 is even, and therefore a is even. If $a = 2c$, then $4c^2 = 2b^2$, $2c^2 = b^2$, and b is also even, contrary to the hypothesis that $(a, b) = 1$. □

Here, the reader must possess a bit of background knowledge. For example, the reader must know that $(a, b) = 1$ means that the only positive integer divisor both a and b share is 1. The mathematician will still likely need to fill out some gaps in the proof. One such gap is between recognizing that a^2 is even and recognizing that a is even. Verifying that the proof is rigorous involves filling these gaps in. Additionally, Wiedijk [2004] has provided a formal version of this proof in Mizar. The Mizar proof is much longer than the original. Simply translating the proof into the Mizar language results in a number of justificatory errors. So a formal proof requires filling in some of the aforementioned gaps. Overall, though, the Hardy and Wright proof was recognized as rigorous without referencing the Mizar article.

The second example is the mutilated chessboard problem which has received attention from Tanswell [2015]. The problem was originally presented as a puzzle for a critical thinking textbook. But Martin Gardner took it up as a mathematical puzzle and then gave a proof. Gardner's formulation of the problem, including the original diagram, and his solution are given below.

> The props in this problem are a chessboard and 32 dominoes. Each domino is of such size that it exactly covers two adjacent squares on the board. The 32 dominoes therefore can cover all 64 of the chessboard squares. But now suppose we cut off two squares at diagonally opposite corners of the board [see Fig. 13] and discard one of the dominoes. Is it possible to place the 31

dominoes on the board so that all the remaining 62 squares are covered? If so, show how it can be done. If not, prove it impossible. [Gardner, 1988, 24]

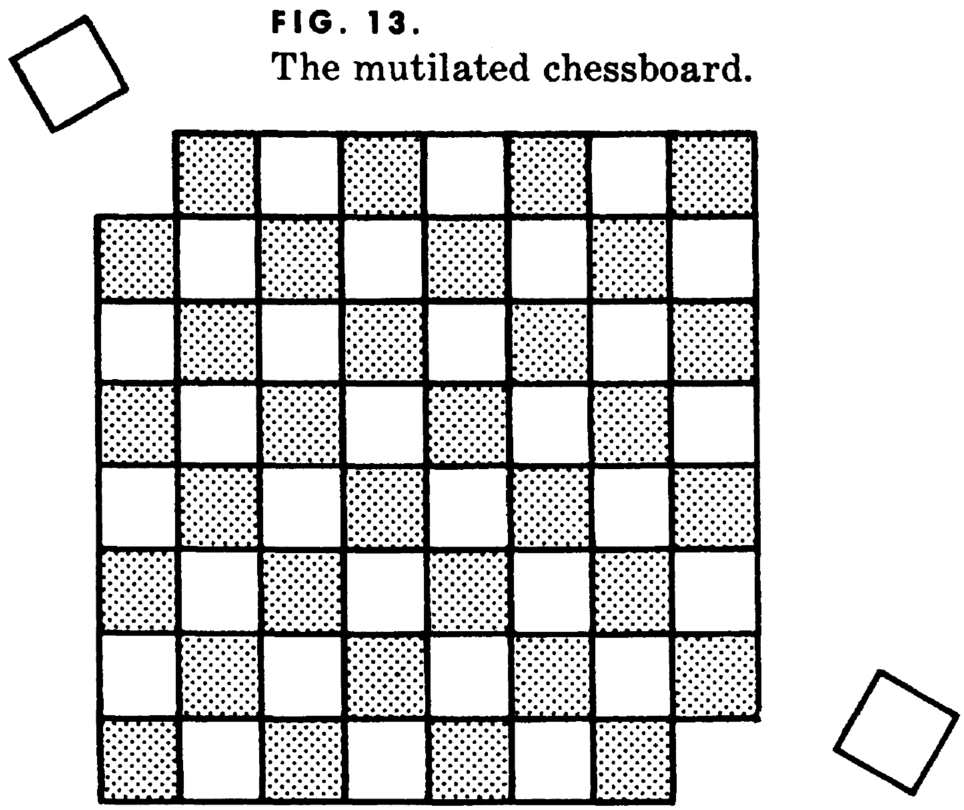

Figure 3.1: The mutilated chessboard image given in Gardner [1988].

It is impossible to cover the mutilated chessboard (with two opposite corner squares cut off) with 31 dominoes, and the proof is easy. The two diagonally opposite corners are the same color. Therefore their removal leaves a board with two more squares of one color than of the other. Each domino covers two squares of opposite color, since only opposite colors are adjacent. After you have covered 60 squares with 30 dominos, you are left with two uncovered

> squares of the same color. These two cannot be adjacent, therefore they cannot be covered by the last domino. [Gardner, 1988, 28]

[Robinson, 1991, 271] claims that it is "not easy to think of an improvement on Gardner's statement of the impossibility proof. It is simple, clear, elegant and convincing." Following Robinson [1991], the proof itself has been discussed in the literature about formalization and rigor, specifically by Tanswell [2017, 2015]. There are formal proofs of the impossibility of covering the mutilated chessboard problem in both Mizar and Lean.

Both of the examples given in this section are widely regarded as examples of rigor in informal proofs. The rest of this paper is dedicated to exploring what features the proofs have that make them rigorous. Recall that both proofs can, and have been, formalized by providing formal portrayals of the proofs in Mizar and Lean. The existence of formal counterparts is not in istelf evidence for the standand view nor is it an objection to the audience view. As mentioned in Chapter 1, proponents of the standard view have argued that, in addition to the fact that formal proofs exist, the proofs are informally rigorous in virtue of some relationship to the formally rigorous proofs. The existence of a few formal proofs does not guarantee that (a) all proofs can be formalized or (b) that the informal proofs are judged rigorous in virtue of some relationship to the formal proofs. The audience view presented in Section 3.3 claims that the proofs are rigorous in virtue of something other than their formally rigorous counterparts.

3.2 What is a Theory of Rigor?

Before giving the audience view of rigor, note that there are a few types of account that one may aim to develop. Mathematicians engage in an evaluative practice when they judge proofs to be rigorous. Burgess [1992] discusses the difference between a descriptive and prescriptive theory of an evaluative practice. A descriptive theory aims to explicitly describe the implicit standards of the practice. Descriptive theories are evaluated according to how well they conform with the practice. A prescriptive theory aims to provide what the practice's standards should be. The practice is evaluated with respect to the prescriptive theory.

Data for descriptive theorizing involves spontaneous evaluations of particular instances made by practitioners. Even when we assume the practitioners are following general rules, they may not be able to state the rules they are following. Descriptive theorizing aims to make these rules explicit. Once the theorist insists on a set of rules which do not match with conflicting data, the theorist is engaged in prescriptive theorizing. According to Burgess [1992], linguistic descriptive theorizing takes the form of judgments like "that's not good English" or "that isn't said" while prescriptive theorizing takes the form of judgments like "people talk that way all the time, but it's wrong." Burgess [1992] then applies the distinction to classical logic. A descriptive logic is a branch of naturalized epistemology where one aims to examine one's own practices and reveal the implicit rules. A prescriptive logic is an 'alienated epistemology' where the logician evaluates practices from without. The prescriptive logician often appeals to privileged intuitions and fundamental critical insights.

Burgess [1992] was interested in grammatical practices and logical practices. Here, I'm interested in expanding to mathematical practices. In the case of rigor, a descriptive account of rigor might explore what mathematicians mean when they say they "know it when they see it" about rigor. It can do so without offering a further claim that what mathematicians are doing is the *right way* of determining a proof is rigorous. In addition to the descriptive and prescriptive distinction, I want to add that a descriptive theory may look at different levels. One may aim to give a descriptive theory of the implicit rules of a community or of a single agent.

Using the mutilated chessboard example, then, an agent-level descriptive theory explains the standards by which an individual mathematician judges Gardner's proof to be rigorous. A community-level descriptive theory could explain the standards by which Gardner's proof is accepted as rigorous within the community. Finally, a prescriptive theory tells us whether the mutilated chessboard proof ought to be rigorous according to some standard. It is possible for the prescriptive theoretician to say "everyone thinks the Gardner proof is rigorous but that's wrong." It is not possible for the descriptive theoretician to say that. The practice has judged the proof to be rigorous.

A last important note about theorizing is that descriptive and prescriptive grammar are related. Almost no prescriptive grammar is devoid of descriptive theorizing since native judgments guide prescriptive theorizing. Likewise, native judgments are influenced by some prescriptive pedagogical grammar. Language learners are told which rules are good and bad. In other words, the descriptive and the prescriptive are often linked through education. Descriptive and prescriptive theories of rigor will also most likely be linked and difficult to separate.

These distinctions are useful for identifying the goals of providing an account of rigor. For example, according to Hamami [2019], some objections to the standard view fail because they're objections to the standard view as a descriptive theory which is not the typical aim of the standard view. The audience view described in Section 3.3 has three components: an agent-level descriptive theory, a community-level descriptive theory, and a prescriptive theory.

3.3 The Audience View

The purpose of this section is to introduce a new account of rigor called the audience view. In broad strokes, the audience view claims that a fully rigorous proof is one where an idealized universal audience would assent to every inference. The universal audience is a fictional audience meant to represent all reasonable people and it is constructed from experiences with real audiences. According to the audience view, proofs are rigorous without reference to formal proofs. The audience view does not imply that formalizability is impossible. Nor does it imply that formal proofs have no influence on mathematical practices. Rather, I argue for an account of rigor which focuses on the social components of practice which may be orthogonal to formalization.

I begin this section by giving an overview of the requisite rhetorical groundwork. I'll first identify multiple kinds of audiences that an arguer might be interested in. Following Ashton [2021], I identify the *universal audience* as the most important of these audiences for the purposes of mathematical proof. I'll then detail three claims: a descriptive claim about how proofs are judged to be rigorous by agents in practice; a descriptive claim about how proofs are judged to be rigorous within a community of practitioners; and a prescriptive claim about what feature proofs ought to have in order to be rigorous. The benefits of the descriptive claims will be explored in Section 3.4.

3.3.1 Rhetorical Groundwork

The philosophy of mathematics and argumentation theory have traditionally been divorced from one another. It was Aristotle who wrote "it is evidently equally foolish to accept probable reasoning from a mathematician and to demand from a rhetorician scientific proofs" (Aristotle, NE I.3). In recent years, argumentation theory has been successfully applied to mathematical practice. An overview of that work is available in Aberdein and Ashton [2024]. But the divide has been especially strong with respect to the study of audience-focused argumentation like rhetoric. Rhetoric is often conceived of as dealing with persuasion and uncertainty. The traditional division between mathematics and rhetoric continued in modern texts. Perelman and Olbrechts-Tyteca [1969] aimed to provide a *New Rhetoric* according to which persuasion was just one way of moving an audience. They gave an account aimed at convincing audiences which was appropriate in legal and philosophical contexts. More recently, Tindale [2015] gives an updated account of argumentation to audiences which relies heavily on Perelman and Olbrechts-Tyteca [1969].

Both Perelman and Olbrechts-Tyteca [1969] and Tindale [2015] separated the realm of rhetoric from the realm of deduction. The problem with most of these arguments, discussed by Dufour [2013] and Ashton [2021], is that formal proof is what Perelman and Olbrechts-Tyteca [1969] and Tindale [2015] claimed to be audience-free. But mathematical proofs are rarely fully formal proofs. Mathematical proof is gappy and various audiences are considered in presentation, axiom choice, and even development. Ashton [2021] argued that proofs are arguments constructed with audiences in mind and that the core audience in proof development is the *universal audience* which is a useful, fictional audience. The goal of this subsection is to outline the concept of the universal audience and explain some of its important features. In following subsections, I'll use this concept to characterize rigor.

Let us distinguish among a few kinds of audiences. Perelman and Olbrechts-Tyteca [1969] are not entirely clear about how to differentiate between types of audiences. In the following, I adopt Sigler [2015]'s distinction, as they synthesize the views and comments Perelman makes in later work as well. According to Sigler [2015], there are primarily three types of audiences which arise in Perelman's work. The first kind of audience is the *real audience*. These are physical audiences which receive arguments and can react to them.

They are the real people whose assent the arguer hopes to gain. Real audiences can receive arguments in person or through media like text. Each of these real audiences belong to different groups and react differently to different argumentative strategies.

It is not possible to access the real audience every time one constructs an argument. To overcome the issue of audience access, arguers invoke fictional, constucted audiences, which Perelman calls *particular audiences*. They are particular audiences because they are created by abstracting from some particular feature of real audiences. But many of our arguments are to a group, not a single person. One may have experience with how a number of real mothers react to real arguments. When the arguer tries to develop an argument for *mothers* as a group, then the arguer focuses on the particular audience of mothers. It is by arguing to real mothers that the arguer forms an idea of which arguments mothers assent to. For example, an arguer might notice that his real mother is persuaded to buy a product if it will help her son avoid getting sick. He might argue directly to his mother that he needs this product to avoid an illness. But a marketer, having experiences with multiple real mothers, forms an idea that the audience of *mothers in general* are swayed to buy products that protect children from illness. The marketer may even be confident enough in his generalization to say "all mothers are willing to buy products to protect their kids from illness." He is, strictly speaking, incorrect. There are some mothers who do not care about their childrens' health. The "all" that the marketer invokes makes sense when he's never met uncaring mothers. If pushed with evidence of uncaring mothers, the marketer will drop the "all" and adopt "most" or "all the mothers we care about." In this sense, sentences invoking particular audience assent function like a generic statement. Psychological research like Leslie et al. [2011] indicates that people often default to a generic interpretation when faced with a quantified statement. Just as adults accept "all ducks lay eggs" using a generic interpretation, a marketing executive might accept "all mothers would buy a product if it protected their child from illness."

Arguers construct fictional, particular audiences in their minds to test arguments out before approaching real audiences. Arguers can check the quality of both the argument and the fidelity of their fictional, particular audience by arguing to real audiences which closely instantiate that particular audience. For example, our marketing executive may run a focus group selecting only mothers in order to test his marketing campaign and verify that mothers, in general, are persuaded by arguments that invoke protecting children from sickness. The focus group is a real audience. It is a real group of people, all of whom are mothers, who receive and react to the marketing executive's arguments. The particular audience is a mental concept employed by the marketing executive, based on his experiences with real moms, to represent mothers as a whole. The marketer obviously can't test his argument with all real mothers because he can't collect all mothers from all of time in one room. Nevertheless, he has an idea of what is required to be a mother in his current society and he's had experiences arguing with those mothers. So he can construct an imagined audience which is representative of mothers based on his experiences with real mothers.

Sometimes, arguers want to convince *everyone*. When an arguer aims to convince every reasonable person, they are arguing to the universal audience. Just as mothers comprise a particular audience defined by motherhood, the universal audience is a particular audience defined by reasonableness. The universal audience is called universal because reasonableness is assumed to be a feature of all real audiences whereas motherhood is not a feature of all real audiences.

What is a reasonable person? Perelman and Olbrechts-Tyteca [1969] and Tindale [2015] both seem to take reasonableness as a feature people implicitly grasp over time. They give very little theoretical definition of reasonableness by which arguers can judge audiences or people. To be a reasonable person means, at minimum, that most of the time this person takes reasons into account when forming beliefs and making decisions. Reasonable people are receptive to evidence and argument. There are multiple ways of being reasonable. A reasonable person might follow either of two opposing schools of thought since both “command equal respect because they express thoughtful and recognized ways of thinking and in this they are both reasonable” [Perelman, 1979, 113]. Reasonableness is something people exhibit by engaging in discussion and argumentation in recognized ways. These recognized ways change over time and can be different in different societies. But people who continuously give and take reasons, evaluate evidence, and form judgments on the basis of reasons and evidence are called reasonable. Those are the people we wish to convince by constructing audiences to the universal audience. This characterization of reasonableness is open-textured. There may be people in the future for whom a decision must be made even though there is no fact of the matter as to whether they’re reasonable.

Given a certain set of experiences, one may come to bar unreasonable people but this should be based on an overall pattern of behavior, not a single instance of non-adherence to an argument. A dialethist, for example, may still be a reasonable person since, on the whole, they exhibit a willingness to share and evaluate reasons for their beliefs. On the other hand, someone with Cotard’s syndrome who believes themselves to be dead and refuses to consider any evidence of their existence will be unreasonable. Their pattern of behavior is not due to a recognized pattern of thinking but to an insurmountable neuropathology. Moreover, a group of people who consistently follow a different, but recognizable, pattern of thought are not unreasonable. This seems clear in Perelman’s discussion of opposing schools of thought. He says that “both are seen as equally reasonable; we will choose, but not on the basis of the falsity or irrationality of the one or the other” [Perelman, 1979, 113]. So we may disagree with whole groups of people while also recognizing them to be reasonable. Reasonableness seems to be a virtue that a person can possess. We know people are reasonable because they exhibit recognized ways of thinking within a community.

We now have a loose characterization of reasonable person. Again, this characterization is meant to be context sensitive. Children learn who counts as reasonable from the parents, educators, etc. They learn to apply the term "reasonable" in a similar way that they learn to apply terms like "mother." The ability to imagine an audience is also taught to children when they're asked to consider explaining a concept to a sibling, writing a paper for their mom, or trying to persuade their friends to do something. Each particular audience is an abstraction from the real audiences that an arguer encounters. When the arguer wants to try to generalize his argument so as to convince everyone, he argues to his universal audience. It's a fictional audience, influenced by his experiences and his communities, that can be used to construct, and test, arguments. We can imagine how "any reasonable person" might react to an argument.

Again, this "any" should not be strictly interpreted as every real reasonable person. As with mothers, it's more likely that the "all" in "all reasonable people" is interpreted by-default like a generic. There is no real counterpart to the universal audience which an arguer can access, since no arguer can interact with all reasonable people over all time. Arguers only experience a finite, temporally limited number of real audiences and construct their universal audience by abstracting from those real audiences. Because of this, a universal audience will not perfectly represent all people over all time. But by experiencing a wide variety of real, reasonable perople, and learning which arguments they accept, an arguer can construct a better concept of *all reasonable people*, and thus, a more representative universal audience. In other words, the universal audience is a type of mental fiction we use to develop arguments which appeal to reasonable people. To the arguer, the universal audience *will seem to be truly universal* since reasonableness is the most general trait of audiences. From a theoretical perspective, it's clear that the fictional universal audience is affected by real audiences and so the arguer's universal audience will not approximate everyone forever.

Before closing this sub-section, I want to address one immediate concern. One might worry that any person who does not assent to an argument to the universal audience is immediately unreasonable and excluded. In other words, if an arguer believes they have created a successful argument to the universal audience, then there is no possibility to engage in genuine disagreement with reasonable people. This concern is rooted in a misconception about the universal audience. The universal audience is a synthetic concept meant to represent all reasonable people. As is common in synthesis, it privileges a whole over each individual part. As such, an argument to the universal audience may not, in fact, convince every reasonable person. Moreover, no actual person is an instantiation of the universal audience alone. This is because no real person is *just reasonable*. People belong to multiple, composite audiences and so their dissent is often grounded in their inclusion in a different audience. This can be a source of genuine disagreement.

This completes the overview of types of audiences. The universal audience is one of the rarer audiences that one aims to convince. Perelman and Olbrechts-Tyteca [1969] believed that it was primarily for philosophers and lawyers. By mistakenly limiting all of mathematical argument to formal proof, they failed to consider the possibility that mathematicians might be interested in convincing all reasonable people. Ashton [2021] takes this consideration seriously and argues that the goals of arguing to a universal audience are similar to the goals of producing a proof. Thus, proof development involved arguing to the universal audience, as opposed to particular audiences or specific, real mathematicians. Moreover, the universal audience was a useful framework for explaining features of proofs like the invention of new techniques as in Alexander [1928]'s knot diagrams. All of this is good evidence that universal audiences are important to proof development. In the following subsections, I will explicitly apply the universal audience to the concept of rigor.

3.3.2 Agent-Level Descriptive Account

The first claim of the audience view is an agent-level descriptive account. An agent-level descriptive account explains how a single mathematician judges a proof, like those provided by Hardy & Wright or Gardner, to be rigorous. I'll begin by giving the claim that the audience view makes for the agent-level descriptive account. Then, I'll concretize what that looks like for the examples from Section 3.1. Finally, I'll give some reasons to think that mathematical training supports the audience view.

The rigor of a proof seems to derive from its inferential moves. We'll begin with the most extreme case – when a proof is judged to be rigorous by looking at every inferential move. This is, as Andersen [2020] and Geist et al. [2010] have discussed, not typically how proofs are evaluated. But it does happen in some cases. For example, a new graduate student might check the rigor of a proof by looking at every inferential move. Or a conscientious reviewer might examine every inference in a proof of a surprising theorem. The audience view characterizes an agent's judgment of full rigor as follows.

> **Agent-Level Full Rigor**: a practitioner judges a mathematical proof to be fully rigorous when every inferential move in the proof is one that the practitioner's universal audience assents to.

When a mathematician checks to see if a proof is rigorous, she asks herself, at each inferential move whether every reasonable person would accept such a move. But Andersen [2020]'s interviews with mathematicians indicate that proofs are rarely checked move-by-move. Interviews suggested that referees use two methods of validation when reading proofs. Andersen calls the first method Type 1 validation, where mathematicians check whether "the subresults of the proof seems reasonable in light of what she knows and, at least, for most of the subresults, whether it seems reasonable that this type of result can be proved in this type of way" [Andersen, 2020, 238]. If the proof passes the tests of Type 1 validation, then the mathematician will usually not go on to check a subproof line-by-line. This empirical data indicates that mathematicians rarely check every inference of a proof. The audience view is compatible with this empirical data in two ways. First, we may bracket the rigor-check to just Type 2 validation. In other words, the agential-level full rigor account is what happens when a mathematician engages in Type 2 rigor validation. But I think the audience view is compatible with even more of the data. When a mathematician performs Type 1 validation, she is still judging whether the proof is rigorous using her universal audience. She looks at a number of sub-results and applies her version of the reasonable person standard to them. The mathematician has a sense of whether a result is correct and can be reached using the methods mentioned in the proof. In performing Type 1 validation and checking the subresults to see if they're surprising, she's asking whether an audience of reasonable people would assent to those subresults. So she does judge the rigor of a proof through both Type 1 and Type 2 validation using her concept of the universal audience.

Let's examine what this would look like for the Hardy & Wright proof of Section 3.1. The first inferential move is from the assumption that the $\sqrt{2}$ is rational to the fact that $a^2 = 2b^2$ is soluble in integers a, b. The audience view posits that the mathematician would think something like "yeah everyone would agree there, so that's good." The fact that integers a, b can be assumed without loss of generality to be co-prime is also quickly checked. For an advanced mathematician, these steps might be Type 1 verifications. The mathematician might proceed fairly quickly until the claim that "a^2 is even and therefore a is even." The mathematician asks herself "hold on, would everyone accept that claim?" This inference requires more verification work than earlier claims. The mathematician will essentially need to provide a sub-proof that "for any integer n, if n^2 is even, then n is even." This sub-proof fills in the gaps of the original proof and decomposes inferential moves of the original proof into new inferential moves. The mathematician continues through the sub-proof verifying, at each step, whether it seems as though everyone would assent to that move.

Return to the mutilated chessboard problem. One of the first inferences to be verified in Gardner's proof is the one that concludes "the two diagonally opposite corners are the same color." This information is not in the statement of the problem. It is inferred from other information, namely the diagram which represents facts about chessboard coloring. To verify the inference, the mathematician only needs to ask "given the diagram representing the mutilated chessboard, would everyone assent that the two cut corners are the same color?" If the answer is 'yes,' then the inference is verified by the practitioner. At this level, the mathematician might not think everyone would assent to this inference. For example, they might ask themelves whether the diagram is general enough. What happens when we cut off the other corner diagonals instead? This would involve further decomposition processes: produce a similar diagram for the other diagonal, ask again whether everyone would have to assent to the claim that both corners are the same color given the diagram. If the answer is 'yes,' then the inference is verified. There is no need to find some suitable formal system and method of translating the inferential move.

Sometimes, though, not every inference in a proof is judged to be one which the mathematician's universal audience would assent to. For example, a mathematician might write a proof where she thinks most inferences are ones which everyone would assent to, but there's one step that only a particular group of people would assent to. Or the mathematician might validate a proof using Type 1 validation and think that all the subresults seem reasonable, but only for mathematicians in her own field. In those cases, she may judge the proof to be *somewhat rigorous*. The audience view can also account for these judgments. The audience view characterizes an agent's judgment of partial rigor as follows.

> **Agent-Level Partial Rigor**: a practitioner judges a mathematical proof to be somewhat rigorous when most inferential moves are ones that the practitioner's universal audience assents to while some inferential moves are ones that a practitioner's particular audience of some mathematicians would assent to.

Proofs are more or less rigorous based on the number, and relative importance, of inferential moves that the universal audience assents to. A proof which relies heavily on inferential moves that only some audiences of mathematicians would assent to might be judged as somewhat rigorous or barely rigorous. For example, the Hardy & Wright proof might be made more detailed in a classroom by adding the subproof of the claim that "a^2 is even and therefore a is even" by utilizing the contrapositive. The mathematician, presenting this more detailed proof, might naturally tell her class that "the proof I'm giving you here is more rigorous than in the book since you did not already know or accept that if a^2 is even then a is even." The professor might even hold some reservation that her classroom proof was maximally rigorous, since some of her inferential moves are such that only modern mathematicians (not Pythagoras, for example) would assent to. Our example mathematician is entrenched in that practice, though, so she would most likely not worry about Pythagoras. So she may think that her classroom proof is *about as rigorous as it gets*. Rigor is a gradable concept that depends on the audience being addressed.

In addition to making sense of gradable determinations, the agent level descriptive view is supported by mathematical training. Some versions of the standard view, like Tatton-Brown [2019] and Hamami [2019], seem to assume that mathematicians enter their training as blank-slate logicians and build up from the axioms. The audience view assumes that mathematicians enter their training as social arguers. A practitioner's universal audience depends on their training. Young mathematicians rarely enter their training with a logic course, much less a clearly outlined set of axioms. But the audience view is continuous with a mathematician's life. They've argued to particular audiences their whole lives. It's not far-fetched to think that they've formed some idea of what all reasonable people accept. For example, they might learn that any reasonable person accepts that 2 is an even number. They also might think that reasonable people assent to conjunction elimination. They believe this without choosing a formal deductive system in which it is valid. They simply recognize that all reasonable people adhere to the conclusion Q when given "if P then Q" and "P." But they've also built an idea of all reasonable people which is probably too permissive for mathematical practice. This is where some enculturation, as Tanswell and Rittberg [2020] discusses, is necessary. The mathematician refines the universal audience by encountering new types of audiences. It's important to note that such enculturation extends beyond the mathematician's proof practices. The mathematician doesn't think that only mathematicians reject "if P then Q, Q, therefore P." He'll refrain from using such inferential moves no matter who he is arguing to. This is compatible with the audience view's agent-level descriptive account. Compatibility with mathematical training is an important part of evaluating a descriptive account.

The agent-level descriptive account aims to provide the procedures by which proofs are judged to be rigorous by individual practitioners. The way mathematicians judge proofs to be rigorous under the audience view is by considering whether her universal audience would assent to it. I've argued that this method of judging rigor makes sense of gradable rigor judgments and is supported by mathematical training. Next, I turn to community judgments of rigor.

3.3.3 Community-Level Descriptive Account

So far I have described how individual mathematicians judge proofs to be rigorous. But an individual may judge a proof to be rigorous while still getting something wrong. The first level at which a mathematician might be wrong about the rigor of a proof is that the community may have a different standard of rigor than the individual. The audience view also explains how communities judge proofs to be rigorous. I'll begin by stating the audience view for the community-level descriptive account. Then I reflect on how that account explains the examples of Section 3.1. I then discuss the close connections between the agent-level and the community-level descriptive accounts.

A single mathematician has a universal audience in her mind, her conception of which has been influenced by her training in a community. She must satisfy her own universal audience when judging a proof to be rigorous. At a community level, the inferential moves must satisfy the community's universal audience. So, for a mathematical practice *M*, we have the following community-level descriptive account of rigor:

> **Community-Level Full Rigor**: A mathematical proof is judged fully rigorous by a practice *M* when every inferential move in that proof is one to which the universal audience of each M-practitioner would assent.

As in the agent-level descriptive account, we also have a gradable notion of rigor.

> **Community-Level Partial Rigor**: A mathematical proof *P* is judged somewhat rigorous by a practice *M* when most inferential moves in the proof are ones to which the universal audience of each M-practitioner would assent. Some inferential moves would be assented to by a recognizable, particular audience of *M*.

Each arguer faces different real audiences which means that each arguer could have different universal audiences. But people belonging to a certain practice will have similar real audiences and similar universal audiences. This is especially true in modern mathematical training. These similarites result in universal audiences common to a practice. Moreover, practice-specific rules about what is reasonable can alter this common universal audience. At this point, a common universal audience might appear to be just a particular audience of mathematicians. But, as Ashton [2021] argued, mathematicians rarely believe their arguments are convincing only to mathematicians trained like them. They often think anyone who could be brought up to speed on the facts, should be convinced. They also think proofs persist – a proof that is rigorous today should also be rigorous to the mathematicians of tomorrow. All of this indicates that mathematicians intend their arguments to be accepted universally, even if that never happens.

An example of a community judging a proof to be partially rigorous is the computer-assisted proof of the Kepler conjecture. The Kepler conjecture states that the densest packing of congruent balls in 3-dimensional Euclidean space is the face-centered cubic packing. Ferguson and Hales first posted a proof for the conjecture on arXiv in 1998 but that proof wasn't published fully until 2006. The delay was due to the fact that the proof was computer-assisted. A team of 13 reviewers worked to verify the proof. But it was difficult to check every step reliably because of the "nature of this proof, consisting in part of a large number of inequalities having little internal structure, and a complicated proof tree" [Lagarias, 2011, 17]. The proof involved a computer verification of 5,000 configurations of spheres and over 100,000 linear programming problems. Reviewers for the *Annals of Mathematics*, and some mathematicians in the community, were overall quite confident about the proof but held some uncertainty. The proof, as first presented, was somewhat rigorous for that practice as it was known that there were groups of the community who did not fully assent since they couldn't fully survey the computer portion. We can infer the proof is rigorous to the community because the *Annals of Mathematics* published the proof. The *Annals of Mathematics* decided to publish the original proof. That is clear evidence that the proof was rigorous by community standards since the journal is well-respected. The editor's note claims that the human part was "refereed for correctness in the traditional manner" and the computer part was "examined for the methods by which the authors have eliminated or minimized possible sources of error." Nevertheless, Hales launched the *Flyspeck* project to provide a complete formal verification. He says that this decision was partly "in response to the lingering doubt about the correctness of the proof" [Hales, 2006, 1]. The *Flyspeck* project was completed and published in Hales et al. [2017] which, again, indicates the formal verification is rigorous. Under the audience view, the original proof was rigorous and the formally verified proof is rigorous. The formally verified proof may be more rigorous, assuming it eliminated the doubts of the same communities which judged the original proof to be somewhat rigorous.

Let's return to the examples of Section 3.1. The Hardy & Wright proof is rigorous for most modern mathematical practices because those practices have mathematicians whose universal audiences are common to the practice. Likewise, the practitioners of most modern mathematical practices judge the Gardner proof to be rigorous. Imagine, for example, a community that completely eschews any use of diagrams or visual props in a proof. The individuals engaged in that practice would not judge Gardner's proof to be rigorous since it involves reference to a diagram or 'prop'. The community-level descriptive account depends on the judgments of individuals in the practice.

When the mathematician learns from their training and remains in contact with that practice, then, agent-level judgments usually match with community-level rigor judgments for that practice. Through training, the mathematician brings her universal audience in line with the universal audience common to M. Still a mathematician might judge a proof to be rigorous and discover that, after presenting it, the proof is not rigorous for the community. At this point, the mathematician must re-examine her argument or her universal audience. Either she mistook an inference as one that the common universal audience would accept or her universal audience does not match the universal audience common to M.

The audience framework could help diagnose cases of recalcitrant disagreement like the one De Toffoli and Fontanari [2023] discuss. De Toffoli and Fontanari argue that there are six species of mathematical disagreement. They claim that disagreement over the correctness of a putative proof can sometimes be recalcitrant. Recalcitrant putative proof disagreements arise, according to them, in communities where the abstract criterion of rigor and the local criteria of acceptability are not aligned. The authors identify the abstract criterion of rigor with formalizability. They adopt a version of the standard view. They then present a case study from the Italian school of algebraic geometry. Many incorrect putative proofs were given and sometimes accepted by the community, although there was significant controversy between mathematicians Federigo Enriques and Francesco Severi.

Although De Toffoli and Fontanari adopt a version of the standard view, their discussion of Severi's faults is more compatible with the audience view. [De Toffoli and Fontanari, 2023, 23] argue that Severi's putative proofs lacked rigor because he developed a kind of autoreferential trust. But autoreferential trust is more closely tied to convincing an audience than formalizability. Under the audience view, Severi's autoreferential trust in an example of a mathematician who has conflated the audience of himself with his universal audience. Severi had great success which led him to believe he was the ideal audience for his proofs. He fails to re-calibrate his universal audience by arguing with other real mathematicians. Instead, he determines his universal audience in line exclusively with what convinces him. Enriques, as quoted in [De Toffoli and Fontanari, 2023, 38], wrote of Severi's work that the “exposition seemed obscure to me and therefore dubious” which indicates that Severi failed to truly account for his fellow mathematicians in his fictional universal audience. Examples of how the audience view handles disagreement like this deserves serious study in further work. But for now, this example can be used to show what happens when a single mathematician like Severi fails to recalibrate his audience to one common to the community.

The audience view does not guarantee that different mathematical practices will agree as to whether a proof is rigorous. This is especially important when considering historical practices. Some historical proofs may have once been judged to be rigorous but are not rigorous with respect to current practices. An example of this, discussed further in Section 3.4, would be Euclid’s proof of Proposition 1 in Book I. This is because the universal audiences common to a practice vary across practices. The universal audience is a mental fiction which arguers create by abstracting from experiences with real audiences. Historical practices had different encounters with real audiences. So they had different universal audiences because they had formed different concepts of what all reasonable people would be like.

The audience view has a descriptive theory for both agent-level rigor judgments and community-level rigor judgments. The judgments will usually line up, assuming a mathematician is in touch with her practice and receives feedback from the real audiences that comprise the practice. By combining the two accounts, we can make sense of some instances when individual mathematicians erroneously judged a putative proof to be rigorous.

3.3.4 Prescriptive Account

We now have two theses from the audience view. The first is a claim about how agents judge proofs to be rigorous. The second is how communities judge proofs to be rigorous. We discovered one way that a mathematician might err with respect to rigor: she might judge a proof to be rigorous that the community would not judge to be rigorous. In these cases, we think the single mathematician has made the error. But discussions about error lead us into prescriptive theorizing. The purpose of this sub-section is to sketch the audience view's prescriptive account. A complete defense and discussion of the prescriptive theory will be taken up in further work.

Let's return to our 'erroneous' rigor judgment. A mathematician reviews her proof and judges it to be rigorous since each inferential move is something her universal audience assents to. She submits it for publication. Reviewers in her practice review the proof and judge that it is not rigorous, since their universal audiences do not assent to all the inferential moves of the proof. We have a disagreement about whether the proof is rigorous. But we usually privilege the reviewers' judgments over the single mathematician. Perhaps, upon reflection, the mathematician will discover that her universal audience wouldn't actually assent to one of the steps. But perhaps she must revise her universal audience to better match the one common to the practice.

But why should we privilege the community? The answer is that the goal of building a universal audience is to try to approximate a genuinely universal audience. Each arguer constructs her own universal audience which results in idiosyncracies, but she believes it to be universal. The intuition behind claiming that the single mathematician is in error, not the group, is that a single mathematician's universal audience is less likely to approach genuine universality than multiple mathematicians' universal audiences. Having more reasonable people thinking about what all reasonable people would accept is more likely to approach its goal than having a single person do so. This is related to the sheer number and variety of real audiences that each mathematician has encountered.

Of course, we still can't access every real person. And demanding that every real, reasonable person assent to every inferential move in a proof is too high a standard. Thus the prescriptive account of rigor shouldn't appeal to every real reasonable person. Instead, it appeals to a universal audience constructed under ideal conditions or in an ideal practice. So, the prescriptive theory of rigor can be sketched as this:

> **Prescriptive Rigor**: A mathematical proof ought to be judged rigorous by a person or practice when every inferential move in the proof is one that a universal audience common to an ideal practice assents to.

Providing a complete characterization of an ideal practice is a topic for future work which will connect to issues in social epistemology. I will characterize some requirements on it now. First, an ideal practice would not be motivated by fame or self-interest. It would not be about gaining 'theorem credits' as Thurston [1994] puts it. It would also be a practice which was not plagued by epistemic injustice. An ideal practice might also be one that endorses reliability and consistency of results or even control as Manders [2008] and Wagner [2017] discuss.

This needs a number of refinements which will be the focus of future work. One obvious issue is that it's conceptually possible that a large group of people may come to a consensus about what counts as rigorous that leads to unreliable results. There's a sense in which that happened with Euclid's *Elements*. Many communities judged those proofs to be rigorous. But later developments showed that the proofs were not always reliable since some of the hidden assumptions were not necessarily true. Similarly, we might have equally sized competing groups who have conflicting results but seem otherwise equally reasonable. A possible example of similarly sized competing groups could be the camps of Enriques and Severi mentioned earlier and examined by De Toffoli and Fontanari [2023]. In that case, Severi seemed less reasonable once he developed an autoreferential trust.

The primary focus of this sub-section is to sketch how all three components of the audience view work together. Note that one need not adopt the prescriptive view in order to adopt the descriptive views. But it's worth having a prescriptive audience view on the table since it helps make sense of our original motivation – a rigorous proof is that which convinces the right kinds of people. The next section focuses on the benefits derived from accepting the descriptive views.

3.4 Benefits of the Audience View

To close this chapter, I'll highlight the benefits of adopting the audience view and I'll also discuss some common concerns. I begin by discussing the motivating tagline and its original concerns. Then I turn to the benefits of gradable rigor judgments. Then I look at the relationship between the audience view and a mathematician's life as an arguer. I wrap up by discussing the relationship between reliability and rigor under the audience view.

3.4.1 The Motivating Tagline

First, recall that I motivated the audience view using the tagline that an informally rigorous "proof is that which convinces." The audience view has clearly lived up to that tagline, since the descriptive and prescriptive components are inherently tied to convinction. We've refined the tagline by introducing which audience needed to be convinced. But, as we saw, this tagline was often quickly dismissed because some convincing things are not proofs and some proofs are not convincing. The audience view does not fall prey to these objections as easily as the tagline does.

Not all proofs will be convincing immediately since every real reader of a proof is a real audience, not the idealized universal audience. A student might not be convinced on first pass because he's a real person who possesses many other qualities besides reasonableness that may interfere with his judgment. The student can be a reasonable person but not be convinced immediately because he simply can't remember what inferences came before or what was discussed in the class before. Note also that it would not help that student to see a formal proof. There are many issues that real audiences encounter which can inhibit their ability to be convinced. The issues of a real audience are compatible with the belief that an idealized, fictional audience would be convinced.

The more interesting objection is that some convincing things are not rigorous proofs. For example, the teacher could say "look I know you're not convinced by the proof but trust me, the theorem is true." By this testimony, the student is fully convinced that the theorem is true. But no mathematician would say that testimony should constitute rigorous proof. Broadly, the audience view stated that an argument would be judged rigorous when the mathematician's universal audience assents to every inference in it. Under that description, testimony would not be judged rigorous by either the student or the teacher. This is because the experthood and authority required for testimony are both very particular. The teacher is an expert with reference to a real audience of pupils. He is not an expert for all reasonable people. We don't seem to consider reliable authorities to have an atemporal quality to them. We usually recognize them as authorities in a local community. This helps dispel the classic objection to the tagline "a proof is that which convinces" since we don't believe appeals to authority convince our universal audiences.

3.4.2 Benefits of Gradable Rigor

In subsections 3.3.2 and 3.3.3, I discussed descriptive accounts that allowed for both full rigorous and somewhat rigorous proofs. This helped make sense of some linguistic data about rigor judgments. This is crucial because we want to characterize and evaluate mathematical rigor as it is used by mathematicians. And they seem to be able to compare and rank proofs as more or less rigorous. To my knowledge, there is no standard view in the literature that has a gradable notion of rigor. In this subsection, I'll discuss some other benefits of having a gradable notion of rigor.

We can use the gradable notion of rigor to make sense of historical rigor judgments. Take, for example, Book I Proposition 1 of Euclid's *Elements*. We begin with a straight line AB and describe two circles of radius AB, one with A at its center and one with B at its center. The proof assumes that the two circles will intersect at a point C from which we can construct the rest of the triangle. But that inference is only warranted if there is such a point. To supply the point, in modern terms, we invoke the principle of continuity. This is the very first proof in the *Elements* which was long regarded as a paragon of mathematical rigor. Under the standard view, either the proof of Proposition 1 was not rigorous, and all the mathematicians who judged it to be rigorous were wrong, or the proof is fully rigorous in virtue of its relationship to a formal counterpart. But with the audience view, we have more coherent options. In Euclid's time, and according to de Risi [2020] up until the Early Modern period, that the intersection point existed was taken for granted. In my analysis, the mathematicians of those periods had universal audiences who assented to the inference – they thought any reasonable person would have to admit a point existed at the intersection. So for both Euclid and a number of communities before the Early Modern period, the proof of I.1 is fully rigorous. In modern times, most mathematicians would say the proof is only somewhat rigorous. Most of its inferential moves retain their impressive, universally assentable nature. But the inference involving the intersection point is only admissible for some particular audiences of the past. In other words, the gradability of the audience view allows us to make sense of seemingly conflicting rigor judgments where a proof might be a paragon of rigor historically but is only somewhat rigorous by modern standards.

We can also use the audience view's gradable account of rigor to make sense of field-specific standards of rigor. For a community of knot theorists, a diagrammatic proof may be fully rigorous while an algebraist might consider the same proofs to be only somewhat rigorous. And we can make sense of rigor comparisons between two proofs that are vastly different. Such comparisons, as well as historical and field-specific considerations are not as straightforward under the standard view, if they're possible at all.

One final benefit of the gradable notion of rigor is that it can help us make sense of the intuition that mathematicians hold formal proof as a gold standard. The audience view is compatible with the intuition that formal proofs are a gold standard. Formal proofs contain only immediate inferences within a formal deductive system. In other words, anyone who assents to the formal deductive system must assent to the formal proof. Formal proofs are a gold standard because it is presupposed that all reasonable people should assent to the inferences therein. So a formal proof can still be a gold standard because it's a fully rigorous proof under the audience view. This doesn't reduce the audience view to a standard view, though, because rigor is not determined by formalizability. Rather, formalization sometimes matches the overarching desire to produce an argument that all reasonable people would assent to. Thus the audience view can, and does, accommodate the intuition that formal proof is a gold standard since it is a fully rigorous proof.

3.4.3 Mathematicians are Arguers, Proofs are Arguments

There's a broad reason to adopt the audience view which is related to something in subsection 3.3.2. The audience view gives an account of mathematical rigor where mathematical arguments are connected to, but relevantly different from, everyday argumentation. Mathematicians are people engaging in a practice which centers around a particular kind of argument. The standard view asks us to reduce these arguments to a formalization thesis. The audience view respects the natural place of informal proof as argument. Thus, rigor judgments are not mysteriously tied to formal proofs in formal deductive systems that mathematicians may not know about. Rigor judgments are, instead, centered on argumentative practices which are continuous with all of rational life. Moreover, given the importance of the universal audience, rigor is still a special property of mathematical proof. People argue all the time, but they rarely argue to a universal audience. Rigor is special because it aims at the assent of all reasonable people. Jurisprudential rigor and philosophical rigor also aim to do a very similar thing. However, jurisprudential and philosophical rigor are not nearly as successful at achieving long-term assent.

This raises a natural question: what is special about mathematical rigor under the audience view? We started this paper by claiming that one of the special features of modern mathematical proof is that it is rigorous. Nothing about the audience view of rigor is bound to mathematical arguments. We could judge the rigor of any argument containing inferential moves using the same theory. Mathematical proofs are different because mathematical practice demands rigor for justification of a theorem. One can provide rigorous arguments in other fields but it's rarely required or useful. Consider, for example, legal arguments. Legal arguments are occasionally rigorous in the sense described by the audience view. Sometimes we craft legal arguments that are meant to appeal to every reasonable person.

But current legal practice doesn't always demand rigor for justification. Although, perhaps this could have been different if legal formalism were sustained. Policy arguments, for example, are often grounded in notions of public welfare that are not considered universal, even by the people making those arguments.

It seems to me that one of the key reasons that we can demand rigorous arguments in mathematics, but not in the law, is that the nature of the facts, definitions, and values at stake are very different. Contingent facts guide policy considerations. But mathematics is about necessary facts. The contingency of the facts involved in law, along with its definitions and values, are seen as particular to a community. Since the very content of the arguments are particular and contingent, there's little demand for universality, and hence for rigor as described here. But in mathematics, the content of the arguments have a universality of their own and the demand for rigor can be met. This might help explain why we sometimes see rigorous arguments in philosophy or law. But legal and philosophical contexts don't demand rigorous argument since they're not about universal topics. This paper on rigor, for example, is written to my particular audience of philosophers of mathematics.

To sum, mathematical practice is special since it demands rigor but rigor itself can be exhibited by arguments in any field. Mathematical arguments are more disposed to rigor since they involve necessary facts and explicit definitions. That rigor can be exhibited by some non-mathematical arguments is a feature of the audience view. The view characterizes mathematicians as arguers and proofs as arguments in a way consistent with other justificatory social practices.

3.4.4 Reliable, Reasonable People

Historically, an increase in rigor has led to greater reliability. Rigorization is often driven by the desire to have more reliable results. Additionally, mathematical arguments, specifically proofs, are more reliable than arguments in other fields. An account of rigor must explain how rigor is tied to reliability. The standard view does not succeed at explaining this connection. As Avigad [2021] points out, formal proofs are quite fragile. They're fragile in the sense that they are long and so have many steps and if any step is incorrect, the quality of the whole proof decreases. These formal proofs are difficult for humans to grasp, survey, or verify. It is unclear why standard view-rigorous proofs would be more reliable if the most rigorous proofs, i.e., formal proofs, are fragile. So, the standard view alone does not explain why rigorous proofs would be more reliable.

The audience view is in a slightly better position than the standard view with respect to reliability. Rigorous proofs are reliable because they inherently try to avoid and overcome disagreement from a seemingly universal audience. Giving and receiving feedback on arguments helps us avoid mistakes and unclarities. The more people we can convince of something, the less likely there is to be a genuine error in it. So, under the audience view, we still have a theoretical connection between reliability and rigor. Of course, it doesn't guarantee reliability in practice. We might be systematically mistaken about which axioms are true. Then even our most rigorous proofs will be unreliable. Or, as in the case of the pre-Early Modern mathematicians, we might take for granted principles which ought to be expressed and examined. So rigor under the audience view is reliability-conducive even if it doesn't guarantee reliability. This is similar to the correctness-conducive quality of rigor that De Toffoli and Fontanari [2023] discusses.

Reliability is most likely ensured by methods other than rigor alone. Although Avigad [2021] is interested in explaining reliability in conjunction with a standard view, the reliability-increasing features that Avigad describes are not inherently tied to the standard view. For them to be intrinsically tied to the standard-view, Avigad needed to show that there were features in informal proofs that (a) increased reliability and (b) indicated or were routinely translatable into formal proofs. I have no doubt that Avigad shows (a) is true. The problem is that there is nothing about these techniques that indicate formal proofs. To show why, consider modularization which occurs when something can be broken down into smaller components which have limited interaction with each other. Modularizing increases reliability in mathematics, but it also increases reliability in engineering and design. Modularization makes it less likely that my plane, or my software, will crash. But modularized plane designs do not indicate the existence of some formal plane design.[1]

Another example is reasoning by analogy. Reasoning by analogy helps to isolate critical information. If I refer you to a similar proof, then I don't have to go through every step and can focus on the new information. Moreover, casting the proof as analogous allows one to focus on ways that the differences may *make a difference*. This supports error checking. Modularizing occurs when something can be broken down into smaller components which have limited interaction with each other. When one component fails, it only minimally affects the others. This is a common technique in engineering and design. Avigad [2021] also argues that modularity in proof minimizes the required information, helps keep the flow of information clear, and supports error checking. Collecting examples is also an important technique since it often helps to prove a particular instance before moving to a more general setting. So collecting a number of various examples to test and keep in mind when proving

[1]The fact that this category error is so painful and unparseable ought to be strong evidence against (b). What could possibly be indicated in the plane case?

helps the mathematician find out what works and what will not generalize.

The fact that Avigad's methods for increasing reliability occur in other types of arguments across disciplines is very natural under the audience view since rigor is not about formalization. While neither the standard view nor the audience view immediately explain the reliability of proofs, the audience view provides an account of rigor that is reliability-conducive since it naturally accommodates rhetorical maneuvers like those cited by Avigad [2021].

To sum up, the audience view provides an important new development in accounts of rigor, namely gradable rigor judgments. It is consistent with the idea that mathematicians are arguers while maintaining that mathematical practices are special in their demands for rigor. And the audience view maintains a conceptual relationship between rigor and reliability. Given all of this, one should prefer the audience view over the standard view at least for any descriptive account of rigor. I've also sketched some reasons for adopting the audience view's prescriptive account of rigor but reserve a fuller discussion for future work.

3.5 Concluding Remarks

The goal of this chapter was to outline a new theory of rigor in terms of rhetorical concepts. To do so, we first had to delimit what a theory of rigor could be about. A theory could be descriptive or prescriptive. In line with these dimensions, I characterized the audience view of rigor through three components: the agent-level descriptive account, the community-level descriptive account, and the prescriptive account. I argued that there are numerous benefits to adopting the descriptive accounts over a standard view of rigor. Note that the descriptive accounts of the audience view may be adopted without adopting the prescriptive view. Indeed, it may be possible to create a new prescriptive formulation of the standard view that is consistent with the descriptive accounts outlined in this paper. I do not think this is the best future direction, though, since it involves disconnecting the mathematical practices of rigor (the descriptive) from what rigor ought to be (the prescriptive). It will be more fruitful and interesting to see how the new audience view, which was designed from the descriptive practices up to an ideal prescriptive view, fares with respect to current issues in the epistemology of mathematics. One rich direction for such research is to examine how work on epistemic injustice in mathematics might seep into the universal audience and rigor judgments. This is the direction taken next in Chapter 4.

Chapter 4

Rigor, Conviction, and Injustice

Recent work in epistemology has highlighted the importance of the social and ethical dimensions of knowledge. Mathematical knowledge has often been treated as immune to these issues. This is due to the belief that mathematical proof is objective, certain, and formally verifiable. In the last chapter, and in Ashton [2021] and Aberdein and Ashton [2024], I argued that a rigorous mathematical proof is more social than previously recognized. I argued that rigorous proofs are arguments that convince the right kind of audience. The right kind of audience is the universal audience. The universal audience is a fictional audience meant to represent all reasonable people. This audience is a mental construct and functions like a generic in some evaluations. Universal audiences were constructed by experiences with the real audiences that a mathematician encountered. Thus the account of rigor I proposed was deeply social and influenced by the real people that a mathematician encounters throughout their life. The audience view situates rigor among topics in social epistemology as opposed to topics of formalizability and computability.

A question naturally arises after connecting mathematical rigor and social epistemology. Does epistemic injustice occur in rigor judgments and their associated practices? I think the answer is yes. We'll see that there are genuine cases of epistemic injustice related to mathematical rigor. As elsewhere in discussions of epistemic injustice, I'll suggest that we leverage a virtue-theoretic approach to correct epistemic injustices related to rigor.

This chapter has two goals. The first is to discuss two examples of epistemic injustice in mathematics that relate to rigor. Along the way, I'll demonstrate the usefulness of the audience view in characterizing these instances of injustice. The second goal is to suggest an epistemic virtue that will begin to deepen the prescriptive audience view of rigor sketched in Chapter 3. In Section 4.1 I'll review some varieties of epistemic injustice and highlight the core requirements. I'll then discuss two examples of epistemic injustice in mathematics related to rigor in Section 4.2. I'll also review the audience view of rigor and argue that it helps reveal the key epistemic capacity needed to identify participatory injustice in math. Finally, in Section 4.3, I'll argue that a necessary virtue for a well-constructed universal audience is the virtue of participatory justice.

4.1 Background: Epistemic Injustice

To begin, it's worth reviewing what an epistemic injustice is and how it differs from other injustices that one might face. The purpose of this section is to review three kinds of epistemic injustice drawing on discussions from both Fricker [2007] and Hookway [2010]. We'll also review the current literature on epistemic injustice in mathematics to motivate the idea that mathematical practice is not immune to the phenomenon.

4.1.1 Varieties of Epistemic Injustice

According to Fricker [2007], an epistemic injustice occurs when someone is harmed in their capacity as a knower because of some identity prejudice. There are two key components of this account. First, the harm involved is genuinely epistemic. The harm undermines a person's ability to possess knowledge. Second, the relevant identity prejudice must be one that tracks someone through multiple dimensions of social activity. For example, race can be involved in a tracker prejudice since the prejudices associated with one's race can occur in many areas of their social life. On the other hand, a prejudice against continental philosophers would not be a tracker prejudice since it is limited to one relatively small social domain. The harm of a localized prejudice may still be ethically devastating. But it is not, according to [Fricker, 2007, 29], the main focus of epistemic injustice. Both of these features are made clearer by looking at specific breeds of epistemic injustice. Fricker [2007] introduced the concept of epistemic injustice by identifying two varieties – testimonial injustice and hermeneutical injustice. Further, Hookway [2010] has argued that there are also participatory injustices which fall under the phenomenon of epistemic injustice.

The first form of epistemic injustice that Fricker [2007] discusses is testimonial injustice. A testimonial injustice occurs when an audience assigns lower credibility to a speaker's testimony than they deserve on the basis of an identity prejudice. Fricker's central example of this is Tom Robinson's case in *To Kill a Mockingbird*. The book is set in 1930's Alabama where Robinson is a black man on trial for the rape of a white girl. Robinson is innocent. Given the milieu, it's obvious that Robinson faces an identity prejudice which tracks him throughout his social activities. As a result of this identity prejudice, Robinson faces a testimonial injustice. His testimony in the courtroom is not assigned the appropriate credibility because of harmful stereotypes about his race. Robinson is wrongly prevented from conveying his knowledge due to his race. Prejudicial credibility deficits harm the speaker in their capacity as knowers and are thus epistemic wrongs. This kind of epistemic harm can also produce an obstacle to truth and can limit the circulation of critical ideas. Moreover, a recipient of testimonial injustice may lose confidence in their own reasoning. In doing so, they may reduce her own confidence to the point that they fail to meet the conditions required for knowledge. This is, again, best characterized as an epistemic harm since it wrongs the speaker as a knower.

The second type of epistemic injustice discussed by Fricker [2007] is hermeneutical injustice. A hermeneutical injustice occurs when "some significant area of one's social experiences [is] obscured from collective understanding owing to a structural identity prejudice in the collective hermeneutical resource" [Fricker, 2007, 155]. An example of hermeneutical injustice is women who suffered unwanted sexual advances in the work-place before the term sexual harassment. The hermeneutical resources lacked a term for the distincive social experience that women were facing in the workplace. That lack of terminology prevents a sufferer from understanding a significant part of her own social experience. Again, the injustice relies on the concept of social power – women had less power in the workplace and their gender tracked them throughout many social experiences. And the harm done is epistemic – part of someone's social experience is obscured from them in a way that leaves them confused and doubtful. Unlike testimonial injustice, there is no particular culprit for hermeneutical injustice. In testimonial injustice, the hearer inflicts harm by not affording the speaker an appropriate amount of credibility. But in hermeneutical injustice, there is a systematic hermeneutical failure.

Hookway [2010] cites a third type of epistemic injustice which occurs without credibility deficit or conceptual impoverishment. Hookway argues that there is a type of epistemic injustice which harms someone as a potential knower by excluding them participating. While testimonial and hermeneutical injustice harm someone as a knower, participatory injustice harms them as an epistemic participant. The core idea is that the ability to know isn't the only good epistemic activity – they also include the ability to inquire and deliberate. In testimonial injustice, the person is harmed from an informational perspective; they're not treated as a reliable source of information. But in participatory injustices, the person is harmed as a participant in an activity. They're not treated as "competent to carry out some particular activity that has a fundamental role in carrying out inquiries into the solution of some problem" [Hookway, 2010, 7]. One of Hookway's examples of participatory injustice is a woman, perceived stereotypically to be, shy who attempts to engage in philosophical discussion. She's viewed as an incompetent discussant because of her shyness. Since discussion is a mark of success for a philosopher, prejudicial stereotyping leads her to be perceived as a poor performer along a dimension of her epistemic life. Participatory injustices are epistemic since they call into question whether the person is capable of participating in certain epistemic activities.

In this sub-section, we've surveyed three types of epistemic injustice. All three require that the injustice is related to an identity prejudice and that the injustice harms the person in an epistemic way. The main goal in surveying these three types of epistemic injustice is to set the stage for applying them to mathematical practices. Historically, certain groups have been excluded from mathematical education, training, and discussion. In other words, mathematical practice can be part of one's social world where identity prejudices were recognized and enforced. The question naturally arises: do these injustices occur along an epistemic dimension? Some have argued that the answer is yes. We'll turn in the next sub-section to an overview of the literature on epistemic injustice in mathematics.

4.1.2 Epistemic Injustice in Mathematics

The main goal of this section is to review the current literature on epistemic injustice in mathematics. The proposed injustices are not examples of epistemic injustice related to rigor. But this discussion will set us up to explore that possibility. The most salient type of epistemic injustice relied on in this literature is participatory injustice.

Rittberg et al. [2020] investigate the potential for epistemic injustice in mathematics via ghost theorems. A ghost theorem is a type of folk theorem. Folk theorems are theorems which everyone believes to be true and proved but they don't know where or who proved it. Ghost theorems are folk theorems where the proof cannot be traced back to the literature. Their main focus is participatory injustices in the sense that Hookway [2010] defined. They look at a particular example of a ghost theorem affecting Olivia Caramello. The authors are clear that they "do not pass judgment on whether Caramello suffered epistemic injustice" but hope that it will function as an "indicative story" that raises our awareness for the possibility of epistemic injustice. But they also provide a more general argument that ghost theorems can be a potential source of epistemic injustice. In Caramello's case, she was not able to publish a number of results since they were deemed "well-known" or "folklore." Thus the ghost theorem acts as a barrier to Caramello's participation in a specifically epistemic practice. The authors are clear that they couldn't find any evidence of a gender-based bias against Caramello. But, they argue, there could be a different kind of epistemic injustice at play. There are a group of experts in the field that shape what the "existing body of knowledge" is. But in shaping this body of knowledge, they come to possess a "secret knowledge" that Caramello doesn't fully possess. The authors claim that "to judge the originality of Caramello's work on the basis of a conception of the 'existing body of knowledge' which comprises both secret and possible knowledge is intellectually callous" [Rittberg et al., 2020, 3888]. The authors don't specifically claim that Caramello suffered this kind of epistemic injustice, but that secret knowledge can create epistemic injustice in cases like Caramello's.

Tanswell and Kidd [2020] look at epistemic injustices in mathematics education. They're specifically interested in how the ethical components of teaching proofs can be used to refine the apprenticeship model of mathematics education. Under the apprenticeship model, classroom mathematical practices ought to reflect professional practices – the students should be enculturated into the professional practices. According to Tanswell and Kidd [2020], the cultural and class backgrounds of learners clashes with the apprenticeship model. Ultimately, they argue that the apprenticeship model "encourages teachers to enforce the norms of mathematical proving practices ... that strict enforcement of the norms of mathematical proving practices in classrooms can be a source of epistemic injustices" [Tanswell and Kidd, 2020, 1207]. One of the features of mathematical proving practices that allows for epistemic injustice, in their view, is the acontextuality of proof. This involves separating proofs from their authors, histories, and communities. Learners from certain backgrounds, like working-class children and Pasifika learners, face barriers that conflict with the acontextuality. Working-class children have less exposure to the rigid use of language required for these proofs. And Pasifika children view mathematics as a group enterprise which makes it difficult to separate the proof from the community. The authors concede that this alone is not an epistemic injustice but merely a potential source of injustice. They write that "by strictly enforcing the ethical order of mathematics in the classroom one risks disadvantaging some learners in their participation in an epistemic practice. The ethical order of mathematics can thus be a source of epistemic injustice" (1208).

Both of the papers discussed in this sub-section aimed to explore areas where epistemic injustice might occur in mathematics. Although both make important strides in the discussion, neither are related to the concept of mathematical rigor. In the rest of this chapter, I'll discuss two such examples and explore the epistemic harm through the audience view of rigor.

4.2 Epistemic Injustice in Mathematics

In Chapter 3 of this dissertation, I introduced an account of mathematical rigor called the audience view. This account is driven by the idea that an argument is judged to be rigorous when it convinces the right kind of audience. The main goal of this section is to examine some examples of epistemic injustice related to rigor through the lens of the audience view.

To do so, I'll begin by reviewing the audience view. Then, in subsection 4.2.2, I'll examine two examples of epistemic injustice using the audience view to define rigor through capacities. We'll see that exploring potential instances of epistemic injustice leads us to a clarification of the audience view in Section 4.3 which is a legitimate contribution to the philosophy of mathematics.

Before we begin, let's lay out some starting assumptions. A mathematical proof is an argument from premises to a conclusion. Its content concerns the structures and/or objects of mathematics. Although they contain mathematical symbols and notation, proofs are primarily written in a natural language like English or French. Moreover, proofs form an important part of mathematical epistemology. They're a core component of mathematical justification. Setting aside the phenomenon of ghost theorems, one typically can't claim to know a conjecture is true without a proof. In modern mathematics, an argument must be rigorous in order to qualify as a proof. So two necessary conditions for having a mathematical proof is that the argument must be about mathematics and it must be rigorous.[1] I mention all of this because it sets the foundation for our epistemological journey. Credibility is a crucial component of testimony so an unfair credibility deficit can lead to testimonial injustice. Concepts are a crucial part of understanding our experiences so a conceptual lacuna can lead to hermeneutical injustice. Some activities like conjecturing, discussion,

[1]These are most likely not sufficient conditions on proof. Also note that published mathematical proofs must also be interesting, worthwhile, etc.

etc are crucial parts of participation in certain epistemic domains so an unfair capability judgment can lead to participatory injustice. Evaluating the rigor of proofs and making rigor judgments are mathematical epistemic activities. If a mathematician is made incapable of making or giving rigor judgments due to identity prejudice, then we'll have examples of epistemic injustice in mathematics.

4.2.1 The Audience View

The first step to is to explore what constitutes a rigor judgment at all. I've argued in Chapter 3 of this dissertation that the best account of rigor is the audience view.[2] This view is motivated by the idea that a rigorous proof is one that convinces the right kind of audience. This is an intrinsically social view and should allow for a clear connection to social epistemological issues like epistemic injustice. I'll briefly sketch the view in the rest of this subsection. I'll then outline three types of rigor judgments in terms of the audience view.

To begin, we need a partial taxonomy of audiences following Perelman and Olbrechts-Tyteca [1969] and Sigler [2015]. First, there are real audiences. These are people we actually argue with and who can receive and question our arguments. Of course, we don't have access to real audiences all the time. We often have to imagine some version of our real audiences in order to build and evaluate arguments. When we imagine audiences, we use a mental heuristic built on our previous experiences with real audiences. For example, a PhD student in philosophy can't actually argue to her advisor as she writes a paper. Instead, she has an imagined audience of philosophers that she tests arguments out with while she's

[2]The audience view is an alternative to the standard view of rigor. The standard view claims that rigor is due to the relationship between an informal proof and a formal proof. For an overview of arguments against the standard view, see the first chapter of this dissertation.

writing. That same philosophy student might teach her own students to write papers as if they were writing them to explain an idea to their mom. These are instances of imagined particular audiences. They're particular because they're defined along a dimension that only some real audiences have. Not every real person is a philosopher or a mother. But we can imagine audiences that are defined primarily by being mothers or philosophers. Those imagined audiences are informed by real experiences arguing with philosophers and mothers. One of the crucial concepts of Perelman and Olbrechts-Tyteca [1969] is that there is another kind of imagined audience which is defined by a feature we think is universal. This is the universal audience and is an imagined audience built out of experiences with reasonable people. Unlike motherhood and philosophical training, reasonableness is a seemingly universal aspect of real audiences. When an arguer wants to convince everyone, she'll invoke the universal audience in building and evaluating an argument. Every real arguer encounters different real mothers and thus has different imagined audiences of mothers. Additionally, every arguer will have encountered different real reasonable people which leads each arguer to have different universal audiences. In Ashton [2021] I've argued that the universal audience is the main audience involved in developing proofs. Further, in Chapter 3 of this dissertation, I've argued that it plays a core role in rigor judgments. In what follows, I'll describe how rigor can be evaluated using the universal audience.

The main claims of this section will be at the level of descriptive account. The descriptive component of the audience view aims to explain the rigor judgments that mathematicians make. The descriptive accounts of rigor are always relative to an agent or a community. Proofs are made up of many inferential moves. The descriptive-agential audience view of rigor claims that a mathematician judges a proof to be fully rigorous when every inferential move in the proof is one that the practitioner's universal audience assents to. Mathematicians often give "more or less" rigorous proofs though, so there's also an account of partial rigor. A practitioner judges a mathematical proof to be mostly rigorous when most inferential moves are ones that the practitioner's universal audience assents to while some inferential moves are ones that a practitioner's particular audience of some mathematicians would assent to. The number and importance of the inferences which fail the universal audience standard will determine how rigorous the proof is in comparison to other proofs. Mathematicians engaged in a practice will have similar universal audiences since they'll have encountered similar reasonable people. So a proof will be judged fully rigorous for a community when every inference is one that the universal audiences common to that community assent to. Similarly, a mathematical proof is judged somewhat rigorous for a practice when most inferential moves in the proof are ones that the universal audiences common to that practice assent to while some inferential moves are such that some particular mathematical audiences would assent to. Usually, a proof is rigorous for a community when it's passed the judgment of a few peer reviewers. Proofs are not usually judged by every member of a community. But peer review works in part because we take it that peer reviewers have a concept of what is acceptable to the whole community. In our terminology, that means they have a concept of universal audience that matches the universal audiences common to a practice. In practice, then, judgments of rigor depend on whether the inferences in a proof would be assented to by the practitioner's imagined universal audience.

In addition to a descriptive account of rigor, a philosopher of mathematics will want a prescriptive account of rigor. A prescriptive account tells us when a proof really is rigorous, not just when we can expect a mathematician or a community to judge it to be rigorous. This can help us adjudicate between debates in rigor and to help us reveal the underlying epistemic norms of rigor that help ensure reliability and correctness of results. In Chapter 3 of this dissertation, I sketched the start of such an account. A mathematical proof is rigorous when every inferential move in the proof is one that a well-constructed universal audience assents to. Minimally, a well-constructed universal audience reflects a large group of real, reasonable people. This is essential for securing the connections between reliability and rigor. Rigorous arguments are more reliable because they try to avoid counterexamples from any reasonable person; the proofs that stand the test of time are those that can appeal to a majority of reasonable people across time and areas. In Section 4.3, I plan to spell out one ideal construction through the lens of virtue epistemology. This is a natural move in the literature on epistemic injustice – resolve sources of epistemic injustice by developing epistemic virtues.

4.2.2 Injustice in the Descriptive Account

The goal of this subsection is to identify two cases of epistemic injustice related to rigor and characterize them by using the audience view. Our first example involves Sophie Germain's contributions to elasticity theory. In the first case, the purported proofs did lack rigor by contemporary and modern standards. But examination of why the proofs lacked rigor reveals an important dimension of rigor acquisition. The second is drawn from Grete Hermann's work on quantum theory. In this second case, her work was rigorous by modern and contemporary standards. But that rigor was not recognized due to participatory injustice.

The most obvious way that epistemic injustice might occur in rigor judgments is through stereotyping a social group as non-rigorous. Consider, for example, the prejudicial stereotype that women are guided by emotion, not rigorous logic. Being a woman is a feature of one's social identity and this stereotype can track them throughout many aspects of their social life. In mathematical practice, the prejudicial stereotype can be exhibited through rigor judgments. For example, a prejudiced mathematician may judge a proof to be lacking in rigor if he knows it's written by a woman. This can harm the female mathematician as a knower since she's judged unfairly to lack the capacities required for rigorous argument. This is a straightforward participatory injustice in Hookway [2010]'s sense. Providing rigorous arguments is an important epistemic activity in mathematics that she's not able to fully participate in due to prejudice about her capacities. While this example of epistemic injustice causes clear harm, it doesn't require any particular theory of rigor or mathematical practice to understand it. Moreover, the issues involved in this kind of epistemic injustice can be solved via blind review. The ease with which a harm can be resolved is not a sign that the harm is less important. But I think there are more subtle, insidious spaces for epistemic injustice and rigor judgments to interact.

Sophie Germain

Sophie Germain was a female mathematician, physicist, and philosopher around the turn of the 19th century in France.[3] She taught herself mathematics during the French revolution using the books in her father's library. Her parents were not supportive of her studies so she had no formal education in mathematics. After the revolution, she could not attend the École Polytechnique because she was a woman. But lecture notes were made available to any man who asked. So she wrote in to obtain Legendre's lecture notes under the pseduonym Le Blanc. She sent her thoughts on the lecture notes to Legendre under the same pseudonym. Eventually, Legendre was so impressed that he insisted on meeting the young Le Blanc. After discovering Germain was actually a woman, Legendre was largely supportive and acted as an occasional mentor.

[3] The story of Germain's life described here is drawn from Musielak [2015]'s biography. Information about her correspondence with Gauss is drawn from Del Centina and Fiocca [2012].

Under her assumed identity, Germain corresponded with a number of other mathematicians including Gauss and Poisson. Her correspondences with Gauss regarding his *Disquitiones Arithmeticae* are a useful example of both Germain's approach to mathematics and how other mathematicians treat her. In the first four letters to Gauss, she presents herself as Le Blanc. All of her letters contain extensive mathematical notes and proofs. Many of these mathematical arguments involve lengthy, unnecessary calculations. Gauss responds to all of these letters with praise. He is impressed by the young mathematician. In her fifth letter to Gauss, she is forced to announce her true identity. This is because she had a family friend in the military check in on Gauss during the Napoleonic wars. Gauss was safe but he swore he didn't know anyone named Germain. Gauss's letter to Germain upon this discovery is rife with praise. He exalts her for being able to overcome the barriers women face to participating in mathematics. He then presents a single counterexample to one of her previous proofs.[4] The mathematical content of Germain's sixth letter is ignored. Gauss writes back only to say he will be too busy to correspond further. Over the next decade, Germain sends a further four letters which go unanswered.

After her correspondence with Gauss, Germain took up elasticity theory in response to a contest held by the Paris Academy of Sciences. The goal of the contest was to provide a theory of the vibration of elastic surfaces that would match empirical evidence from Chladni's experiments with vibrating plates. The Academy held the contest three times. Each time, Germain was the only entrant. The first time, she was supported by Legendre who provided her with the current research. The second time, Legendre stops supporting her and she's all alone in developing her work. The third time she consults with Poisson.

[4]This counterexample posed a curiosity for modern mathematicians since it's unusually large when smaller counterexamples will do. Some like Mackinnon [1990] assumed that Gauss purposely chose a difficult counterexample so as not to offend Germain. Others like Waterhouse [1994] believe the counterexample is the most obvious within Gauss's frame of mind.

Poisson would eventually publish their work together without any reference to Germain. The first two times, her entry is rejected because it's not rigorous enough. She wins the prize on the third attempt but is forced to self-publish because the Academy still judges it to be insufficiently rigorous. Poisson would go on to rigorously prove many of the results Germain tried to prove using a similar methodology.

Germain has access to all the physical materials she needs to succeed, she had the occasional support of other mathematicians, and she had the right ideas for solving. No other mathematicians were even willing to attempt the problem. She consistently evaluates her own proofs as rigorous while the community correctly recognized them as lacking in rigor. So why does Germain think she's rigorously proved something that's seemingly impossible at the time? I think the answer lies in her own concept of rigor. Germain cannot possess a well-constructed universal audience because of the severe limitations she faces with real audiences. She has no consistent, honest interlocutors. In her non-mathematical life, she's limited to discussions with people who are deemed incapable of the reasonableness required for mathematics. In her mathematical life, she receives largely empty praise. She's treated as an oddity, not a genuine interlocutor. This is why the Gauss correspondence is of such importance to her. And even Gauss did not correspond with her in any significant way after discovering she's a woman. These two features result in a universal audience that does not represent all reasonable people. So she can still make rigor judgments but they won't reliably match those of the community or genuine constraints on rigor.

This is clearly a sad story. But it's also a genuine case of epistemic injustice in mathematics. Germain is harmed as a knower because she's blocked from forming an adequate universal audience. She's blocked because of identity prejudices. There were rules in place to stop women from gaining a mathematical education that were based on harmful stereotypes about women's capacity for reason. This is an identity prejudice that stops her from gaining an education. The praise that Germain receives demonstrates that her contemporaries think she might be the exception to the stereotype. But even her contemporaries seem to fear 'upsetting her' which indicates that they judge her to be so sensitive that she's incapable of reasonable argument. These are participatory injustices – she's unfairly judged to be incapable of reasonable argument on the basis of a prejudice against her sex. These participatory injustices fuel another injustice. By not treating her as a genuine interlocutor, she's further harmed because she's not allowed to develop a functional concept of rigor. Again, this is blocked due to identity prejudice against her sex.

So there are two levels of epistemic harm that Germain faces – a participatory injustice against her capacity for reason and an epistemic injustice blocking her from acquiring a vital capacity. The audience view of rigor is implicitly useful in characterizing these harms. First, according the audience view of rigor, rigor judgments are tied to a universal audience which is constructed by one's experiences with real audiences. And so there's a clear connection between Germain's isolation and her impoverished concept of rigor. She writes proofs that lack rigor because she's been excluded from the mathematical interlocution necessary to align her concept of rigor with the community.

Compare, for example, how a proponent of the standard view might characterize Germain's case. For versions of the standard view, like Tatton-Brown [2019] or Hamami [2019], Germain's judgments should be based on the in-principle formalizability of her purported proofs. Moreover, the skills needed to verify whether or not a proof is in-principle formalizable were gained by training but not necessarily by interlocution. Both Tatton-Brown [2019] and Hamami [2019] think developing a concept of rigor begins by reading and proving with the most basic inference rules and axioms. Then the trainee learns to leave steps out because she has previously proved the sub-proofs related to common inferences. She learns through books on logic and mathematics. Under the standard view, interlocutive experiences with real people are not required to grasp the concept of rigor. Germain has all of the textbooks and notes she should need to learn to be rigorous under the standard view. And so, under the standard view, Germain's systematic faulty rigor judgments are an issue *separate* from the participatory injustices she faces. The audience view provides a much clearer connection between her faulty rigor judgments and the participatory injustices she faces. This helps to reveal the distinct epistemic harm that she faces by being isolated or treated as incapable of being rigorous.

Grete Hermann

One person's concept of rigor can be flawed by epistemic injustice. The concept of rigor will suffer with every injustice that causes a person to assign less of a capacity for reason to a group or person on the basis of a tracker prejudice. This is because the injustice blocks the required feedback loop between real, reasonable audiences and imagined, universal audiences. While one's universal audience can never fully represent all reasonable people, it should at least approximate it in some way. Since a community-level rigor judgment is based on representative agential rigor judgments, we should expect that epistemic injustice can arise at a community level as well.

Let's turn to an example involving rigor judgments in particular. In 1932, in his *Mathematical Foundations*, John von Neumann published a purported proof of the impossibility of hidden-variable theories in quantum mechanics. It was accepted as a rigorous proof until 1966 when John Bell published a devastating objection to the proof. The proof contains a problematic assumption – that of the linearity of expectation for all possible observables. But this assumption does not actually hold. There are well-known counterexamples involving eigenvalues in the quantum mechanical case. Bell is clear – von Neumann's "'very general and plausible' postulate is absurd" [Bell, 1982, 994]. There's clearly been a failure in rigor at a community level. In the 30 years between von Neumann's proof and Bell's criticism, the von Neumann proof was venerated and affected the entire study of quantum mechanics. It's interesting to note that criticisms of the rigor of von Neumann's proof involve a kind of reasonable-person standard. Mermin [1993, 806] writes that "von Neumann's no-hidden-variables proof was based on an assumption that can only be described as silly – so silly, in fact, that one is led to wonder whether the proof was ever studied by either the students or those who appealed to it." This kind of analysis is in-line with the audience view of rigor – we know the proof is not rigorous because any reasonable person ought to have seen the use of such a silly assumption.

On its own, this is an interesting story about how long it might take to correct incorrect rigor judgments. Von Neumann clearly thought his own proof was rigorous. And at least a few mathematicians also judged it to be so. Bell managed to point out an assumption that most reasonable people would not agree with which resulted in the community recognizing the von Neumann proof as non-rigorous. The core issue is that John Bell was not the first to point out the absurd use of linearity assumptions in von Neumann's proof. In 1935, Grete Hermann published an essay 'The Circle in Neumann's Proof.' She argues that von Neumann's proof is circular since it attempts to rule out the existence of dispersion-free states by assuming the additivity rule but the additivity rule only holds when there are no such states. Hermann's argument was almost entirely ignored in the years between von Neumann and Bell. People have made a number of assumptions as to why. The obvious examples are that the paper was not published in a popular journal, von Neumann was an almost prophet-like figure, she was young, she was a woman, she was a political outsider (read: openly anti-Nazi), she typed that part of her article in a small font, she did not ascribe significance to her own argument, and she "did not desire the status of a revolutionary or radical" [Seevinck, 2016, 116]. Paparo [2012] argues that since later editions of her work don't include the argument criticizing von Neumann, this shows that she thought the argument was insignificant. Again, that seems like an absurd conclusion to draw about a paper literally entitled 'The Circle in Neumann's Proof.' It seems fairly obvious that she did attribute importance to her argument since she went through the trouble of having it published and titled in a clear, non-conciliatory way. I think we can dismiss any arguments that Hermann's critique was ignored on the basis of her personality. The stronger argument is that the essay is not published in a popular journal. This would be a strong argument if the paper had not received further contemporary attention. As Seevinck points out, Carl Friedrich von Weizsäcker wrote a review of the 1935 essay so must have at least read her circularity objection. Likewise she composed the argument while working with Heisenberg at his institute in Leipzig. The journal she published in is not the strongest, but the arguments within the paper were clearly read by important physicists at the time.

Having ruled out the absurd arguments, combined with the evidence that Hermann's circularity argument must have been read by important, respected physicists independent of the publishing journal, we are left with the sad, but obvious conclusion. Her clear, correct argument against von Neumann's proof was ignored for a social factor. It doesn't matter for the purposes of this paper whether that social factor was her youth, sex, or political orientation. What matters is that each of those carry an identity prejudice that can track her throughout her social life. That identity prejudice is what allows mathematicians and physicist to exclude her from their universal audience and, in turn, to ignore her circularity objections and continue to make incorrect rigor judgments about von Neumann's proof. Grete Hermann suffered an epistemic harm by having her rigor judgment ignored on the basis of an identity prejudice.

There are a few core points to pull from Hermann's case. First, she has a prescriptively adequate concept of rigor. Unlike Germain, Hermann has sustained, stable interlocutors in the field who don't treat her with kid gloves. Second, her proof is published which indicates that she understands the community's working concept of rigor. So she doesn't face the secondary epistemic injustice that Germain faced. Hermann does still face an epistemic injustice. She's harmed by having her correct rigor judgment ignored on the basis of an identity prejudice. Since an entire community ignores her, this is an epistemic injustice occurring at the community level. Moreover, von Neumann's lapse in rigor propagates throughout the field and hinders quantum mechanical progress for over 30 years. So the epistemic injustice Hermann experienced was an epistemic harm to both her and the field itself.

Again, we may be tempted to ask what the standard view would say about Hermann's case. Hermann's critique of von Neumann was correct. According to the standard view, when disagreements about the status of a proof occur, they can be passed to the 'higher court' of formalization. That clearly is not what happens in Hermann's case. Her critique is ignored altogether. And when Bell's critique is accepted, it is no more formal than Hermann's. So the standard view doesn't help us explain why Bell's critique was taken seriously while Hermann's was not.

Both Germain and Hermann's cases are examples of genuine epistemic injustice. This is because both examples satisfy the core components of epistemic injustice – an epistemic harm and an associated identity prejudice. To best characterize the epistemic harm, we identified reasonableness and rigor as epistemic capacities which are vital to mathematical practice. This involved invoking the descriptive audience view of rigor. In the next section, I aim to characterize a necessary virtue for a prescriptive account of rigor that avoids epistemic injustice.

4.3 Righting the Wrongs: Virtue Argumentation

The purpose of this final section is to refine the prescriptive component of the audience view so as to avoid participatory injustices leading to lapses in rigor. Just as Fricker [2007] suggested cultivating the virtues of testimonial and hermeneutical justice, I'll suggest that we should cultivate the virtue of participatory justice. I'll define this virtue in terms of who should be included in the universal audience. I'll suggest that the virtuous mathematician (the one with good rigor judgments in line with the prescriptive view) works to adjust his capacity judgments in the face of identity prejudices.

There has been a recent resurgence in virtue-theoretic approaches to many areas of philosophy including ethics and epistemology. Virtue-theoretic turns involve shifting the focus to one's epistemic or ethical character instead of the rules or conditions of an act. A virtue is a certain disposition to value and act in characteristic ways. Virtue-theoretic thinking was characteristic of ancient Greek thought. Plato, Socrates, and Aristotle all endorsed studies of the virtues. Aristotle's virtue theory is perhaps the most influential on Western philosophical thought. Both his ethical works discuss the role of cultivating virtue in pursuit of the good life. In the *Nichomachean Ethics*, he discusses the idea that each virtue can be understood as the right degree of a property situated between the dual vices of excess and deficit. Virtue ethics and virtue epistemology both have varying potential lists of the important virtues. But they also diverge on which aspects of virtue are the most important. For example, virtue epistemologists tend to fall into either a responsibilist or a reliabilist camp each of which stresses a different component of virtue as being important for generating knowledge. The focus of this section is not to provide a comprehensive account of the epistemic virtues, but to focus on one type of epistemic virtue – those which arise in discussions on epistemic injustice.

Fricker identifies testimonial injustice as a kind of vice where hearers fail to correct for identity prejudice when hearing testimony. She argues that “testimonial responsibility requires a distinctly reflexive critical social awareness” [Fricker, 2007, 91] which a virtuous hearer uses to neutralize prejudicial impacts on her credibility judgment. She gives a characterization of the fully virtuous hearer as follows:

> The fully virtuous hearer, then, as regards the virtue of testimonial justice, is someone whose testimonial sensibility has been suitably reconditioned by sufficient corrective experiences so that it now reliably issues in ready-corrected judgements of credibility. She is someone whose pattern of *spontaneous* credibility judgement has changed in light of past anti-prejudicial corrections and retains an ongoing responsiveness to that sort of experience. Full posssession of the virtue, then, in a climate that has a range of prejudices in the social atmosphere, requires the hearer to have internalized the reflexive requirements of judging credibility in that climate, so that the requisite social reflexivity of her stance as hearer has become second nature. [Fricker, 2007, 97]

In singular instances, the virtuous hearer reflects on potential credibility deficits at play in her evaluation of a speaker owing to identity prejudices. She then upwardly corrects her credibility assignment to adjust for the effect of prejudice. But full possession of a virtue cannot occur in a single instance. We can’t determine whether a person is honest on the basis of a single instance. Likewise, a testimonially just hearer, reliably adjusts her credibility assignments to correct prejudicial deficits. Virtue comes in degrees so a person may be more or less virtuous. But for full possession of the virtue, she must issue in ready-corrected judgments spontaneously, in a similar way that an honest person is spontaneously disposed to tell the truth. Parallel to Fricker's introduction of testimonial justice, she also introduces a virtue of hermeneutic justice. We're provided with a characterization:

> The form the virtue of hermeneutical justice must take, then, is an alertness or sensitivity to the possibility that the difficulty one's interlocutor is having as she tries to render something communicatively intelligible is due not to its being a nonsense or her being a fool, but rather to some sort of gap in collective hermeneutical resources. ... The virtuous hearer, then, must be reflexively aware of how the relation between his social identity and that of the speaker is impacting on the intelligibility to him of what she is saying and how she is saying it. [Fricker, 2007, 169]

So again we see that to combat epistemic injustice, we need to cultivate the associated virtues. And the associated virtues themselves involve adjusting the hearer's interpretation of the speaker's credibility or hermeneutical struggle in light of the prejudices they face.

Given Fricker's use of the virtues, we'd expect Hookway [2010] to characterize a similar virtue of participatory justice. But he only points to the importance of a wide range of participatory activities which might also require reflective virtues. One obvious set of epistemic activities are those involved in argumentation. Aberdein [2010] argues that a virtue-theoretic approach to argumentation is fruitful and attempts to survey some of the argumentational virtues. He argues that the virtues of argument propagate truth and virtuous arguers are disposed to spread true beliefs around. Argumentative virtues fall into categories like a willingness to engage in argumentation, willingness to listen to others, willingness to modify one's own position, and a willingness to question the obvious. The category of willingness to listen to others includes intellectual empathy and fairmindedness.

Along the dimension of willingness to listen to others, I'd like to sketch a virtue of participatory justice as it relates to the construction of a universal audience. The universal audience represents what the arguer takes to be all reasonable people. Reasonableness is a capacity required for consideration when building the audience. In other words, if an arguer thinks that a group of people is straightforwardly unreasonable, or judges them to be incapable of giving, taking, and evaluating reasons, then that group will not be considered when the arguer constructs his universal audience. We saw in earlier sections that judgments of unreasonableness can occur on the basis of a prejudicial stereotype, like "women are too emotional to be properly reasonable." This example illustrates how certain kinds of particular audiences – like women – can be excluded from one's universal audience. The particular identity overshadows the potential for exhibiting reasonableness. When the particular audience is deemed unreasonable, then dissenting opinions won't be incorporated into the arguer's universal audience. And the arguer's universal audience will be less likely to represent real reasonable people. The virtue of participatory justice, then, involves reflecting on whether one's judgment of a reasonable person is based on the person's reasoning or based on an identity prejudice. When it is based on their identity, the virtuous arguer is aware that seeming unreasonableness is due to an identity prejudice and makes extra effort to incorporate those real audiences when constructing his universal audience. The virtuous arguer adjusts his willingness to change his conceived universal audience in the face of arguing with real audiences subject to participatory prejudices.

4.3.1 The Virtuous Mathematician

We've now sketched the virtue of participatory justice. The goal of this subsection is to discuss how that virtue affects mathematical practices. We'll discuss the descriptive and prescriptive accounts of mathematical rigor. Then I'll discuss an example where a particular audience is excluded in a vicious fashion, resulting in rigor judgments which fail to meet the prescriptive standard.

Let's return to the descriptive account of rigor. There, a virtuous mathematician will improve her rigor judgments by including real, reasonable audiences in the process by which she constructs her universal audience. Based on the account given in this dissertation, we can venture the following empirical hypothesis, if enough mathematicians in the community exhibit the virtue of participatory justice, then community-wide rigor judgments will be improved as well. For example, in Grete Hermann's case, her argument against von Neumann's purported proof was ignored on the basis of some feature of her identity. She encountered participatory injustice and the people who ignored her work failed to exhibit the virtue of participatory justice. In this case, people judged that von Neumann's proof was rigorous using their universal audience. If those mathematicians were virtuous, they would have put extra effort into recognizing and evaluating her dissent and then incorporating her into their universal audience. By incorporating that real audience in constructing their universal audience, the same mathematicians would come to judge von Neumann's purported proof as lacking in rigor. In this case, the virtue of participatory justice would improve rigor judgments, reaching an epistemic aim.

Now let's turn to the prescriptive account of rigor. Earlier in this paper, we said that a mathematical proof is fully rigorous when every inferential move in the proof is one that a well-constructed universal audience assents to. Minimally, a well-constructed universal audience reflected a large group of real, reasonable people. This is essential for securing the connections between rigor and reliability across time and contexts. We can now characterize a necessary condition for constructing a universal audience well – the arguer must exhibit the virtue of participatory justice. For a proof to really be fully rigorous, it must at least be an argument that a universal audience free of participatory injustice assents to. As it stands, this account allows for pluralism – there are many ways of being a virtuous arguer and, thus, many ways of having a well-constructed universal audience. So there might be many ways for a proof to be genuinely rigorous.

4.4 Conclusion

The main goal of this chapter was to examine whether, and how, epistemic injustice might be related to rigor judgments. We saw two examples where participatory injustices affected rigor judgments. Along the way, we showed that the audience view of rigor could help illustrate the *epistemic* harm involved in Germain and Hermann's examples. To avoid such injustices in mathematics, I sketched a new virtue of participatory justice that would refine the audience view.

Chapter 5

Conclusion

The focus of this dissertation was mathematical rigor as it is exhibited and judged in practice. In Chapters 1 and 2, I examined the standard views of mathematical rigor and argued that they did not make sense of mathematical practices. In Chapter 3 I outlined the audience view of rigor through three interrelated rigor judgments. For a proof to be rigorous, it must convince the right kind of audience. This is a deeply social account of mathematical rigor. In Chapter 4, I discussed examples demonstrating that the social and ethical dimensions of knowledge apply in mathematical rigor. To conclude this dissertation, I'll briefly mention some areas for future work that were not addressed here.

The audience view focuses on the inferences of a rigorous proof. But, as we saw in Chapter 1, concepts, definitions, and notions all play a role in informal rigor. For Kreisel [1967, 1987] and Rav [1999, 2007] the meaning of the terms involved was a crucial part of informal rigor. The audience view does not yet provide us with an account of *rigorous definition*. This is in line with current literature on the standard view, which also fails to give an extensionally adequate account of rigorous definition. Mathematicians in practice don't translate their definitions into a formal language in order to rigorize them and Burgess [2015] tells us that rigor demands any new concepts be defined in terms of the primitives or previously defined concepts. Moreover, proponents of the standard view often face the paradox of rigor: a fully rigorous treatment of a subject matter will ipso facto cease to be a treatment of the subject matter alone. So we should have serious doubts that the standard views on offer can strike a balance between meaningful definition and rigorous definition. The audience view will avoid the paradox of rigor and should be compatible with an account of rigorous definition. I leave it open to future work whether an account of rigorous definition can fruitfully involve the universal audience or whether we should identify a separate, complementary view of rigorous definition.

In addition to rigorous definition, there are other aspects of mathematical practices around rigor that may be worth examining using the audience view. For example, Tanswell and Rittberg [2020] have already begun to look at epistemic injustice in mathematics education. I think the audience view could also help to construct more dialogical methods for teaching proof in a more inclusive manner.

Another aspect worth exploring in closer detail is diagrammatic proof. Again the audience view does not rule out diagrammatic proof and a universal audience can assent to a variety of inferential moves, including those involving diagram manipulation. But further investigation into diagram use is needed to help identify how misleading diagrams are ruled out.

One final topic is the issue of deductive validity. Unlike the standard view, rigor does not guarantee validity (or vice versa) under the audience view. Nevertheless, there should be some reliable connection between the two. So I plan in future work to connect the dialogical account of deduction proposed by Dutilh Novaes [2020] to the audience view of rigor. This will help tighten the connection between rigor and validity from a social, argumentative perspective.

Overall, there are numerous topics left to be explored relating to mathematical rigor. There have been very few alternatives to the standard view of rigor and even fewer are sensitive to the social dimension of mathematical rigor. This dissertation provides an alternative account from which we can begin to explore those topics through a social lens.

References

Abbott, S. *Understanding Analysis*. Springer-Verlag, New York, 2015.

Aberdein, A. The Informal Logic of Mathematical Proof. In Hersh, R., editor, *18 Unconventional Essays about the Nature of Mathematics*, pages 56–70. Springer-Verlag, New York, 2006.

Aberdein, A. Virtue in Argument. *Argumentation*, 24(2):165–179, 2010.

Aberdein, A. Evidence, Proofs, and Derivations. *ZDM Mathematics Education*, 51(4), 2019.

Aberdein, A. and Ashton, Z. Argumentation in Mathematical Practice. In Sriraman, B., editor, *Handbook of the History and Philosophy of Mathematical Practice*, pages 1–23. Springer International Publishing, Cham, 2024.

Aberdein, A. and Dove, I. J., editors. *The Argument of Mathematics*. Springer, Dordrecht, 2013.

Alexander, J. W. Topological Invariants of Knots and Links. *Transactions of the American Mathematical Society*, 30(2):275–306, 1928.

Alexander, J. W. and Briggs, G. B. On Types of Knotted Curves. *Annals of Mathematics*, 28(1):562–586, 1927.

Andersen, L., Johansen, M., and Sorensen, H. Mathematicians Writing for Mathematicians. *Synthese*, 198:6233–6250, 2021.

Andersen, L. E. Acceptable Gaps in Mathematical Proofs. *Synthese*, 197:233–247, 2020.

Antonutti Marfori, M. Informal Proofs and Mathematical Rigour. *Studia Logica*, 96: 261–272, 2010.

Aristotle. *Nicomachean Ethics*. Internet Classics Archive.

Ashton, Z. Audience Role in Mathematical Proof Development. *Synthese*, 198:6251–6275, 2021.

Avigad, J. Reliability of Mathematical Inference. *Synthese*, 198:7377–7399, 2021.

Avigad, J., Dean, E., and Mumma, J. A Formal System for Euclid's Elements. *The Review of Symbolic Logic*, 2(4):700–768, 2009.

Azzouni, J. The Derivation-Indicator View of Mathematical Practice. *Philosophia Mathematica*, 12(2):81–105, 2004.

Azzouni, J. Is there still a Sense in which Mathematics can have Foundations? In Sica, G., editor, *Essays on the Foundations of Mathematics and Logic*, pages 9–47. Polimetrica International Scientific Publisher, Monza, 2005.

Azzouni, J. Why Do Informal Proofs Conform to Formal Norms? *Foundations of Science*, 14(1–2):9–26, 2009.

Azzouni, J. That We *See* That Some Diagrammatic Proofs Are Perfectly Rigorous. *Philosophia Mathematica*, 21(3):323–338, 2013.

Azzouni, J. The Algorithmic-Device View of Informal Rigorous Mathematical Proof. In *Handbook of the History and Philosophy of Mathematical Practice*. Springer, 2020.

Bair, J., Błaszczyk, P., Ely, R., Henry, V., Kanovei, V., Katz, K. U., Katz, M. G., Kutateladze, S. S., McGaffey, T., Schaps, D. M., Sherry, D., and Shnider, S. Is Mathematical History Written by the Victors? *Notices of the AMS*, 60(7):886–904, 2013.

Bell, J. S. On the Impossible Pilot Wave. *Foundations of Physics*, 12(10):989–999, 1982.

Boyer, C. *A History of Mathematics*. Wiley, New York, 1968.

Burgess, J. Proofs about Proofs: A Defense of Classical Logic Part I: the Aims of Classical Logic. In Detlefsen, M., editor, *Proof, Logic, and Formalization*. Routledge, London, 1992.

Burgess, J. *Rigor & Structure*. Oxford University Press, Oxford, 2015.

Burgess, J. and De Toffoli, S. What is Mathematical Rigor? *Aphex*, 25:1–17, 2022.

Burton, L. Moving Towards a Feminist Epistemology of Mathematics. *Educational Studies in Mathematics*, 28(3):275–291, 1995.

Carrascal, B. Proofs, Mathematical Practice and Argumentation. *Argumentation*, 29(3): 305–324, 2015.

Casella, G. and Berger, R. L. *Statistical Inference*. Duxbury, 2002.

Corcoran, J. Argumentations and Logic. *Argumentation*, 3:17–43, 1989.

Crocco, G. Informal and Absolute Proofs: Some Remarks from a Gödelian Perspective. *Topoi*, 38:561–575, 2019.

Davis, P. J. and Hersh, R. Mathematics and Rhetoric. In Davis, P. J. and Hersh, R., editors, *Descartes' Dream: The World According to Mathematics*, pages 57–73. Penguin, London, 1986.

de Risi, V. Did Euclid Prove Elements I, 1? The Early Modern Debate on Intersections and Continuity. In Beeley, P., Nasifoglu, Y., and Wardhaugh, B., editors, *Reading Mathematics in Early Modern Europe*. Routledge, 2020.

De Toffoli, S. Groundwork for a Fallibilist Account of Mathematics. *The Philosophical Quarterly*, 71(4), 2020.

De Toffoli, S. Reconciling Rigor and Intuition. *Erkenntnis*, 86:1783–1802, 2021.

De Toffoli, S. and Fontanari, C. Recalcitrant Disagreement in Mathematics: An "Endless and Depressing Controversy" in the HIstory of Italian Algebraic Geometry. *Global Philosophy*, 33, 2023.

De Toffoli, S. and Giardino, V. Forms and Roles of Diagrams in Knot Theory. *Erkenntnis*, 79(4):829–842, 2014.

De Toffoli, S. and Giardino, V. Envisioning Transformations—the Practice of Topology. In Larvor, B., editor, *Mathematical Cultures*, pages 25–50. Springer, Berlin, 2016.

Del Centina, A. and Fiocca, A. The Correspondence Between Sophie Germain and Carl Friedrich Gauss. *Archive for History of Exact Sciences*, 66(6):585–700, 2012.

Detlefsen, M. Proof: Its Nature and Significance. In Gold, B. and Simons, R. A., editors, *Proof and Other Dilemmas: Mathematics and Philosophy*, 2009.

Dove, I. J. Towards A Theory of Mathematical Argument. *Foundations of Science*, 14 (1–2):137–152, 2009.

Dufour, M. Arguing Around Mathematical Proofs. In Aberdein, A. and Dove, I. J., editors, *The Argument of Mathematics*, pages 61–76. Springer, Dordrecht, 2013.

Dutilh Novaes, C. Reductio Ad Absurdum from a Dialogical Perspective. *Philosophical Studies*, 173:2605–2628, 2016.

Dutilh Novaes, C. A Dialogical Conception of Explanation in Mathematical Proofs. In Ernest, P., editor, *The Philosophy of Mathematics Education Today*, pages 81–98. Springer, Cham, 2018.

Dutilh Novaes, C. *The Dialogical Roots of Deduction*. Cambridge University Press, 2020.

Editors for Annals of Mathematics. Front Matter. *Annals of Mathematics*, 162(3), 2005.

Epple, M. Knot Invariants in Vienna and Princeton during the 1920s: Epistemic Configurations of Mathematical Research. *Science in Context*, 17(1):131–164, 2004.

Fallis, D. Intentional Gaps in Mathematical Proofs. *Synthese*, 134:45–69, 2003.

Ferreirós, J. and Gray, J., editors. *The Architecture of Modern Mathematics*. Oxford University Press, Oxford, 2006.

Fricker, M. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford University Press, 2007.

Gardner, M. *Hexaflexagons and Other Mathematical Diversions*. University of Chicago Press, Chicago, 1988.

Geist, C., Löwe, B., and Kerkhove, B. V. Peer Review and Knowledge by Testimony in Mathematics. In Löwe, B. and Müller, T., editors, *PhiMSAMP. Philosophy of Mathematics: Sociological Aspects and Mathematical Practice*, pages 155–178. College Publications, London, 2010.

Giaquinto, M. *Visual Thinking in Mathematics: An Epistemological Study*. Oxford University Press, Oxford, 2007.

Goldberg, S. *To the Best of Our Knowledge*. Oxford University Press, Oxford, 2018.

Hadamard, J. *An Essay on the Psychology of Invention in the Mathematical Field*. Princeton University Press, Princeton, NJ, 1945.

Hadamard, J. Thoughts on the Heuristic Method. In Ayoub, R. G., editor, *Musings of the Masters*. The Mathematical Association of America, 2004.

Hales, T., Adams, M., Bauer, G., and et. al. A Formal Proof of the Kepler Conjecture. *Forum of Mathematics, Pi*, 5, 2017.

Hales, T. C. Introduction to the Flyspeck Project. In Coquand, T., Lombardi, H., and Roy, M.-F., editors, *Mathematics, Algorithms, Proofs*, volume 5021 of *Dagstuhl Seminar Proceedings (DagSemProc)*, pages 1–11, Dagstuhl, Germany, 2006.

Hamami, Y. Mathematical Rigor, Proof Gap and the Validity of Mathematical Inference. *Philosophia Scientiæ*, 18(1):7–26, 2014.

Hamami, Y. Mathematical Rigor and Proof. *The Review of Symbolic Logic*, 15(2):409–449, 2019.

Hardy, G. H. Mathematical Proof. *Mind*, 38:1–25, 1929.

Hardy, G. H. and Wright, E. M. *An Introduction to the Theory of Numbers*. Oxford University Press, Oxford, fourth edition edition, 1975.

Harrison, J. Without Loss of Generality. In Berghofer, S., Nipkow, T., Urban, C., and Wenzel, M., editors, *Proceedings of the 22nd International Conference on Theorem*

Proving in Higher Order Logics, TPHOLs 2009, volume 5674 of *Lecture Notes in Computer Science*, pages 43–59, Munich, Germany, 2009. Springer-Verlag.

Heath, T. L. *The Thirteen Books of Euclid's Elements*. Barnes & Noble, New York, NY, 2006.

Hookway, C. Some Varieties of Epistemic Injustice: Reflections on Fricker. *Episteme*, 7 (2):151–163, 2010.

Hume, D. *A Treatise of Human Nature*. Penguin Books, London, 1739. 1969 Reprint.

Hunsicker, E. and Rittberg, C. J. On the Epistemological Relevance of Social Power and Justice in Mathematics. *Axiomathes*, 32(3):1147–1168, 2022. ISSN 1572-8390.

Katz, M. G. and Sherry, D. Leibniz's Infinitesimals: Their Fictionality, Their Modern Implementations, and Their Foes from Berkeley to Russell and beyond. *Erkenntnis*, 78: 571–625, 2013.

Krantz, S. G. *The Proof is in the Pudding: The Changing Nature of Mathematical Proof*. Springer, New York, NY, 2011.

Kreisel, G. Informal Rigour and Completeness Proofs. In Lakatos, I., editor, *Problems in the Philosophy of Mathematics*, pages 138–186. North-Holland, Amsterdam, 1967.

Kreisel, G. Church's Thesis and the Ideal of Informal Rigour. *Notre Dame Journal of Formal Logic*, 28(4):499–519, 1987.

Kreisel, G. and Krivine, J. L. Appendix II: Foundations of Mathematics. In *Elements of Mathematical Logic (Model theory)*. North-Holland, Amsterdam, 1967.

Lagarias, J. C. The Kepler Conjecture and Its Proof. In Lagarias, J. C., editor, *The Kepler Conjecture: The Hales-Ferguson Proof*, pages 3–26. Springer New York, New York, NY, 2011.

Lakatos, I. *Proofs and Refutations: The Logic of Mathematical Discovery*. Cambridge University Press, Cambridge, 1976. Edited by J. Worrall and E. Zahar.

Lakatos, I. Part I: Philosophy of Mathematics. In *Mathematics, Science and Epistemology*, volume 2. Cambridge University Press, 1978a.

Lakatos, I. What does a Mathematical Proof Prove? In *Mathematics, Science and Epistemology*, volume 2. Cambridge University Press, 1978b.

Larvor, B. How to Think about Informal Proofs. *Synthese*, 187(2):715–730, 2012.

Larvor, B. Why the Naïve Derivation Recipe Model Cannot Explain How Mathematicians' Proofs Secure Mathematical Knowledge. *Philosophia Mathematica*, 24(3):401–404, 2016a.

Larvor, B., editor. *Mathematical Cultures: The London Meetings 2012–2014*. Birkhäuser, Basel, 2016b.

Larvor, B. From Euclidean Geometry to Knots and Nets. *Synthese*, 196:2715–2736, 2019.

Lavers, G. Benacerraf's Dilemma And Informal Mathematics. *Review of Symbolic Logic*, 2(4):769–785, 2009.

Leslie, S.-J., Khemlani, S., and Glucksberg, S. Do all Ducks Lay Eggs? The Generic Overgeneralization Effect. *Journal of Memory and Language*, 65(1):15–31, 2011.

Macbeth, D. Diagrammatic Reasoning in Euclid's Elements. In Kerkhove, B. V., Vuyst, J. D., and Bendegem, J. P. V., editors, *Philosophical Perspectives on Mathematical Practice*, volume 12 of *Texts in Philosophy*. College Publications, London, 2010.

Mackinnon, N. Sophie Germain: Or Was Gauss a Feminist? *The Mathematical Gazette*, 74(470):346–351, 1990.

Mancosu, P., editor. *The Philosophy of Mathematical Practice*. Oxford University Press, Oxford, 2008.

Mancosu, P. Measuring the Size of Infinite Collections of Natural Numbers. *The Review of Symbolic Logic*, 2(4), 2009.

Mancosu, P., Jørgensen, K., and Pedersen, S., editors. *Visualization, Explanation, and Reasoning Styles in Mathematics*. Springer, Dordrecht, 2005.

Manders, K. The Euclidean Diagram. In Mancosu, P., editor, *The Philosophy of Mathematical Practice*, pages 80–133. Oxford University Press, Oxford, 2008.

Mermin, N. D. Hidden Variables and the Two Theorems of John Bell. *Reviews of Modern Physics*, 65:803–815, 1993.

Musielak, D. *Prime Mystery: The Life and Mathematics of Sophie Germain*. AuthorHouse, 2015.

Paparo, G. Grete Hermann: Mathematician, Philosopher, and Physicist. Master's thesis, Utrecht University, 2012.

Paseau, A. What's the Point of Complete Rigour? *Mind*, 125:177–207, 2016.

Perelman, C. *The New Rhetoric and the Humanities*. Reidel, 1979.

Perelman, C. and Olbrechts-Tyteca, L. *The New Rhetoric: A Treatise on Argumentation*. University of Notre Dame Press, Notre Dame, IN, 1969.

Pólya, G. *How to Solve It: A New Aspect of Mathematical Method*. Princeton University Press, 1945.

Rav, Y. Why do we Prove Theorems? *Philosophia Mathematica*, 7(3):5–41, 1999.

Rav, Y. A Critique of a Formalist-Mechanist Version of the Justification of Arguments in Mathematicians' Proof Practices. *Philosophia Mathematica*, 15(3):291–320, 2007.

Reyes, G. M. The Rhetoric in Mathematics: Newton, Leibniz, the Calculus, and the Rhetorical Force of the Infinitesimal. *Quarterly Journal of Speech*, 90(2):163–188, 2004.

Reyes, G. M. Stranger Relations: The Case for Rebuilding Commonplaces between Rhetoric and Mathematics. *Rhetoric Society Quarterly*, 44(5):470–491, 2014.

Rittberg, C. J., Tanswell, F. S., and Van Bendegem, J. P. Epistemic Injustice in Mathematics. *Synthese*, 197(9):3875–3904, 2020.

Robinson, J. A. Formal and Informal Proofs. In Boyer, R. S., editor, *Automated Reasoning: Essays in Honor of Woody Bledsoe*, pages 267–282. Kluwer, 1991.

Robinson, J. A. Informal Rigor and Mathematical Understanding. In Gottlob, G., editor, *Computational Logic and Proof Theory*, volume 1289 of *LNCS*, pages 54–64. Springer, 1997.

Rudin, W. *Principles of Mathematical Analysis*. McGraw HIll, New York, 1976.

Seevinck, M. Challenging the Gospel: Grete Hermann on von Neumann's No-Hidden-Variables Proof. In Crull, E. and Bacciagaluppi, G., editors, *Grete Hermann - Between Physics and Philosophy*, pages 107–117. Springer Netherlands, Dordrecht, 2016.

Shapiro, S. Computability, Proof, and Open Texture. In Olszewski, A., Wolenski, J., and Janusz, R., editors, *Church's Thesis after 70 Years*, pages 420–455. De Gruyter, 2006.

Sigler, J. The New Rhetoric's Concept of Universal Audience, Misconceived. *Argumentation*, 29(3):325 – 349, 2015.

Steiner, M. *Mathematical Knowledge*. Cornell University Press, 1975.

Steiner, M. Mathematical Explanation. *Philosophical Studies*, 34:135–151, 1978.

Tanswell, F. A Problem with the Dependence of Informal Proofs on Formal Proofs. *Philosophia Mathematica*, 23:295–310, 2015.

Tanswell, F. Conceptual Engineering for Mathematical Concepts. *Inquiry*, 61(8):881–913, 2018.

Tanswell, F. S. *Proof, Rigour and Informality: A Virtue Account of Mathematical Knowledge*. PhD thesis, University of St. Andrews, 2017.

Tanswell, F. S. and Kidd, I. J. Mathematical Practice and Epistemic Virtues and Vices. *Synthese*, pages 1–20, 2020.

Tanswell, F. S. and Rittberg, C. J. Epistemic Injustice in Mathematics Education. *ZDM*, 52 (6):1199–1210, 2020.

Tatton-Brown, O. Rigour and Intuition. *Erkenntnis*, 86:1757–1781, 2019.

Thurston, W. P. On Proof and Progress in Mathematics. *Bulletin of the American Mathematical Society*, 30(2):161–177, 1994.

Tindale, C. W. *The Philosophy of Argument and Audience Reception*. Cambridge University Press, Cambridge, 2015.

Troelstra, A. Choice Sequences and Informal Rigour. *Synthese*, 62:217–227, 1985.

Tymoczko, T. The Four-Color Problem and its Philosophical Significance. *Journal of Philosophy*, 76:57–83, 1979.

van der Waerden, B. L. How the Proof of Baudet's Conjecture was Found. In Mirsky, L., editor, *Studies in Pure Mathematics: Papers Presented to Richard Rado on the Occassion of his 65th Birthday*, pages 251–260. Academic Press, London, 1971.

Villani, C. *Birth of a Theorem: A Mathematical Adventure*. Farrar, Straus and Giroux, New York, 2015.

Wagner, R. *Making and Breaking Mathematical Sense: Histories and Philosophies of Mathematical Practice*. Princeton University Press, Princeton, NJ, 2017.

Wagner, R. Mathematical Marriages: Intercourse Between Mathematics and Semiotic Choice. *Social Studies of Science*, 39(2):289–308, 2009. ISSN 03063127.

Waismann, F. Verifiability. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 19:119–150, 1945.

Waszek, D. Rigor and the Context-Dependence of Diagrams: The Case of Euler Diagrams. In Chapman, P., Stapleton, G., Moktefi, A., Perez-Kriz, S., and Bellucci, F., editors, *Diagrammatic representation and inference (Diagrams 2018)*. Springer, 2018.

Waterhouse, W. C. A Counterexample for Germain. *The American Mathematical Monthly*, 101(2):140–150, 1994.

Wiedijk, F. Formal Proof Sketches. In Berardi, S., Coppo, M., and Damiani, F., editors, *Types for Proofs and Programs: International Workshop, TYPES 2003, Torino, Italy, April 30 - May 4, 2003, Revised Selected Papers*, volume 3085 of *LNCS*, pages 378–393, Berlin, 2004. Springer.

Wynn, J. and Reyes, G. M. *Arguing with Numbers: The Intersections of Rhetoric and Mathematics*. Penn State University Press, 2021.