Three Essays on Aesthetic Experience

Ewa Bigaj

Dissertation Advisor: Richard Moran

Three Essays on Aesthetic Experience

Abstract

In this dissertation, I connect aesthetics to two unusual areas: mathematics and meditation. In the first two essays, I argue that beauty and aesthetic experience make a difference to mathematical practice. A mathematical proof is very different from more familiar beauties (say, sunsets), but I argue that what mathematicians call "beauty" really deserves the name. "No Mathematics Without Beauty" argues that mathematicians would find it harder to understand or create certain proofs without the capacity for aesthetic experience, and answers the objection that only things perceived through the senses can be beautiful. "The Allure of Elegance" is a detailed case study which traces the role of aesthetic factors in the development of three related proofs of the Quadratic Reciprocity Theorem. In "Where Aesthetics Meets Meditation," I argue that meditation as well as aesthetic appreciation involves adopting an attitude of "accepting attention." This attitude transforms the character of pain (making it less "painful" but more intense), increases the number and intensity of aesthetic experiences, and is a special, "warm," kind of detachment. The last fact sheds new light on old aesthetic notions such as "distance" and "disinterest."

Contents

List of Figures

List of Tables

Acknowledgments

The "Acknowledgments" section of dissertations used to fill me with dread. Not only did I doubt that I had enough ideas to fill such an ominous manuscript—I didn't even have enough friends to fill the first page!

The initial ideas gathered momentum. Like snowflakes in a snowball, thoughts stuck to new thoughts, until the biggest challenge was getting the damned thing to stop rolling. To my great surprise, the number of friends I met along the way increased proportionally to the number of new ideas. Here's a partial list of these helpers.

My dad gave me my love of mathematics and philosophy, my mom—of literature and beauty. This dissertation wouldn't be possible without either.

At Harvard, Zeynep Soysal encouraged the first glimmers of my interest in mathematical beauty. My adviser Dick Moran helped me keep the spark alive. I am especially grateful for the breadth of his conception of philosophy; my ideas would have shriveled in a smaller plot.

The first essay benefited from correspondence with Viktor Blåsjö. I'm grateful for the generous honesty which he calls his "Dutch bluntness." Alison Simmons, Brad Skow, Sam Zimmerman, Anya Yermakova, Zoe Jenkin, Patricia Marechal, and members of the Harvard Metaphysics and Epistemology Workshop provided advice and words of encouragement.

Gabe Goldberg (via Peter Koellner) pointed me to the topic of the second essay. Barry Mazur, Peter Koellner, and Warren Goldfarb looked over the more mathematical sections. All errors are, of course, my own.

I spent January 2018-March 2019 in Senegal. As Edward Bullough would put it, the physical distance to Harvard enabled the insertion of much-needed psychical distance. Without the year away, my writing would have contained twice as much jargon and half the conviction. Thanks to Wave Mobile Money for funding my stay and to the Philosophy Department for permitting it.

When I returned to Harvard, I was delighted to meet Samantha Matherne—a wonderful addition to my committee. She shared her expertise in contemporary aesthetics and helped me locate the deepest sources of my disagreement with philosophical orthodoxy.

The final essay wouldn't be possible without the generosity of the Mid-Atlantic Vipassana Association and the Cambridge Insight Meditation Society, which provide meditation retreats free of charge. Conversations with Parimal Patil and Zachary Irving helped me uncover the linear structure beneath a web of interconnected ideas.

If "Acknowledgments" are painful for the lonely, the last paragraph—the one about the partner—is a real twist of the knife. So I'll spare you all the adorable ways in which Ben Kuhn contributed to this dissertation. Suffice it to say: without him, I would still be busy with the all-consuming task of not writing.

Introduction

In this dissertation, I connect aesthetics to two unusual areas: mathematics and meditation. In the first two essays, I argue that beauty and aesthetic experience make a difference to mathematical practice. If mathematicians had a different aesthetic sense (if they found different pieces of mathematics beautiful) or if mathematical aesthetic experience had a different phenomenal quality than it does, mathematics would be different than it is. In the third essay, I argue that meditation and aesthetic experience both involve the adoption of an attitude of what I call "accepting attention." The pairing of aesthetics and meditation sheds new light on the vexed notions of aesthetic "distance" and "disinterest." It also offers a solution to the so-called "paradox of tragedy:" our puzzling tendency to seek out prima facie painful art.

Aesthetic experience is the thread which connects the two otherwise disparate parts of this dissertation. In time, I will defend my choice of subject, arguing that aesthetic experiences are the locus of aesthetic value; properties such as beauty only matter because aesthetic experiences do. For now, in lieu of arguments, let me offer the eloquence of the art teacher

Robert Henri. These two statements have guided my thinking throughout the writing process.

> No *thing* is beautiful. But all things await the sensitive and imaginative mind that may be aroused to pleasurable emotion at sight of them. This is beauty.
>
> The object of painting a picture is not to make a picture... The object ... is the attainment of a state of being, a state of high functioning, a more than ordinary moment of existence.[1]

That "state of high functioning" is also the object of learning and creating a mathematical proof or of the meditative practice of paying attention to one's breathing—as well as the subject of this dissertation.

1 Aesthetics Meets Mathematics

The topic of mathematical beauty lies at the intersection of aesthetics and philosophy of mathematics. Aestheticians and philosophers of mathematics alike may balk at the very existence of such an intersection. The aesthetician might protest that "from the outset aesthetics has been devoted to the study of things perceived"[2] through the senses. Whatever mathematical "beauty" is, it can't *really* be beauty! For her part, the philosopher of mathematics might exclaim that beauty, which is "very much in the eye of the beholder,"[3] can play no real epistemic role in mathematics—the rational pursuit *par excellence*. Beauty is no more than a nice bonus a mathematician chances upon in her serious pursuit of truth, and anyway the beautiful

[1] [Henri, 2007]
[2] [Binkley, 1977]
[3] [Lange, 2016]

pieces of mathematics are little more than "clever insights of quite limited generality."[4]

Against the backdrop of these theoretical preconceptions, the claims mathematicians themselves make about their discipline can sound shocking indeed. Here's number theorist G.H. Hardy.

> Beauty is the first test: there is no permanent place in the world for ugly mathematics.[5]

In the first two essays of this dissertation, I hope to shift the borders of aesthetics and philosophy of mathematics to accommodate voices such as Hardy's. More precisely, I argue that beauty and aesthetic experience make a difference to mathematical practice. This claim has two components.

> [*Real Beauty*] Mathematicians chose the word "beautiful" for a reason. Aesthetic experience in math has something in common with paradigm cases of beauty.
>
> [*Real Difference*] If mathematicians didn't have aesthetic experiences, mathematical practice would change noticeably.

I linger over *Real Beauty* because the examples from which we learned the concept "beauty" in the first place are all perceived through the senses: flowers, sunsets, and the like. The abstract truths of mathematics may seem very distant from these familiar beauties. Is mathematics really beautiful? I'll argue that it is.

[4] [John Dawson, 2006]
[5] [Hardy, 1992] *A Mathematician's Apology.*

How will I do that? Well, here's how I *won't* do it. I won't reduce the property of beauty to some set of properties of beautiful objects (such as: the beautiful objects are those which are either green or contain the sound of the violin). There is no natural property a beautiful cat has in common with a sunflower (or with Beethoven's Pastoral Symphony!) that makes both beautiful. That is my starting point. So I won't try to find such a property for mathematical objects either.

So instead of providing a general account of beauty, in the first essay I will show you three simple but—I believe—beautiful proofs of the Pythagorean Theorem, presenting them in a way which highlights their beauty. Then I will ask you to apply two very simple tests for whether these pieces of mathematics really are beautiful. The first is whether or not you—you personally!—are compelled to call them beautiful. Insofar as I manage to get you to exclaim, when reading one of my explanations of the proofs I'll present here, "How beautiful!" I will have made progress. The second test is whether or not the thing in question gives you so-called disinterested pleasure—pleasure that is independent from any advantage the thing might bestow upon you (as it would if it were, for instance, a check in your name for a large sum). If, when experiencing the proofs I will present in what follows, you feel pleased, without gaining anything but familiarity with the proof itself, that will be reason to believe that the proofs are beautiful.

In other words, I'll give you a sort of guided tour of a little corner of the garden of mathematical delights, highlighting features which please *me*—in the hopes that you will discover a similar pleasure within your bosom. Guiding your interlocutors' attention in this way is, I think, the thing to do in any domain in which you aesthetically appreciate something. You come to believe that Monticelli's "Still Life with White Pitcher" is beautiful not by consulting your list of beauty-conducive properties, not by a deductive proof, and not even by learning that many before you, including Van Gogh, have found the painting beautiful. You learn it by coming into contact with the painting, and it's something about your inner experience, perhaps about your feelings or the quality of your pleasure, that tells you that you're in the presence of beauty. If I want to get you to share in my appreciation of the painting, the most I can do is to highlight the features which attract me. The most I can do is help you stand roughly where I'm standing.

Getting you to "stand" someplace which gives you a clear "view" of the beauty of various pieces of mathematics is part of what I'm aiming to accomplish in this dissertation. But in my capacity as a tour guide, I will be more of a mathematician than a philosopher. The philosopher's job will begin once we have collected a store of simple specimens of mathematical beauty. Then I will ask you to pay attention to your experience of these pieces of mathematics, thus entering the realm of phenomenology. In the first essay, I will argue that the experience of one variety of mathematical beauty involves a process of "compressing" and "decompressing," of shifting perspective from individual pieces of the proof to the proof as a unified whole.[6] Furthermore, analogous processes occur in our experience of paradigmatically beautiful objects, such as paintings and musical compositions. While beautiful *objects* don't share any natural set of properties, then, it is my

[6]This view is indebted to Henri Poincaré's essay "Mathematical Creation" [Poincaré, 1910].

view that our experiences of these objects when we find them beautiful *do* share some core features.

The first essay is also concerned with answers to the *Role* question. Following Poincaré, I argue that mathematicians wouldn't be able to understand or create certain proofs without the capacity for aesthetic experience. The second essay, "The Allure of Elegance," tackles *Role* in more detail. In this essay, I trace the development of three related proofs of the Quadratic Reciprocity in number theory. I argue that the pursuit of beauty motivated Carl Friedrich Gauss to search for multiple proofs of the theorem and Gotthold Eisenstein to improve Gauss's proof. Interestingly, Eisenstein was misled by his pursuit of beauty into the creation of a needlessly complicated proof. This observation is the starting point for a discussion of the relationship between the epistemic virtue of explanatoriness and the aesthetic virtue of beauty.

2 Aesthetics Meets Meditation

Last November, I went on a silent 10-day meditation retreat. Sitting motionless and paying attention to my breath for hours at a time had some striking effects. First, I began to take intense aesthetic pleasure in the grounds of the meditation center (a parking lot with a border of grass). The boring became beautiful: a patch of dried grass was strewn with warm-colored, richly corrugated leaves and twisted all in one direction, as if someone had carefully combed through a head of stiff, golden curls. Even the repulsive turned gorgeous: a caterpillar's enormous mandible morphed from disgusting to sublime before my eyes. Second, my experience of pain changed. During a meditation session, I paid careful attention to a backache, only to find a tingling sensation: staggeringly powerful, but without a trace of suffering. The only other times I had experienced such "sublime" pain had been in the context of art: with the late string quartets of Shostakovich or the black paintings of Rothko.

Adding to those experiences the fact that aestheticians and meditation teachers use strikingly similar language, recommending attitudes of "detachment," "distance," and "disinterest," I became intrigued. The third essay, "Where Aesthetics Meets Meditation" grew out of that curiosity. I argue that meditation as well as aesthetic appreciation involves adopting an attitude of what I call "accepting attention." This attitude transforms the character of pain (in the words of Robert Wright, such pain has "less 'youch!' than usual and more 'whoa!' than usual") and increases the number and intensity of aesthetic experiences. It is also a special, "warm" kind of detachment.

My experiences may seem "out there," but I urge you to keep an open mind. My view is explanatorily powerful: it explains how emotions experienced in the context of art do and don't differ from ordinary emotions, why we seek out tragic art, and what traditional aesthetic views which use terms such as "distance" do and don't get right.

Chapter 1

No Mathematics Without Beauty

Isn't Figure 1.1 beautiful?

No? Are you sure? Well, think of the Pythagorean Theorem. In case it's been a while since your last geometry class, let me refresh your memory. This is the theorem that the lengths a, b, and c of the sides of a right-angled triangle (where c is the length of the hypotenuse) obey the relationship

$$a^2 + b^2 = c^2$$

Figure 1.1 provides a wordless proof of the theorem. (Figure 1.2 helps explain the proof for those who prefer using some symbols. You will find a fuller explanation in Section 4.1.)

Now do you find Figure 1.1 beautiful? If so, you've just had an aesthetic experience centered around a piece of mathematics. That experience is the topic of this essay. (If you *don't* find the figure beautiful, have no fear—there

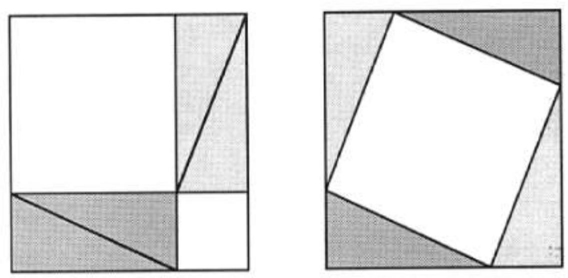


Figure 1.1: Proof #1 of Pythagorean Theorem. (Reproduced from [Nelsen, 1993].)

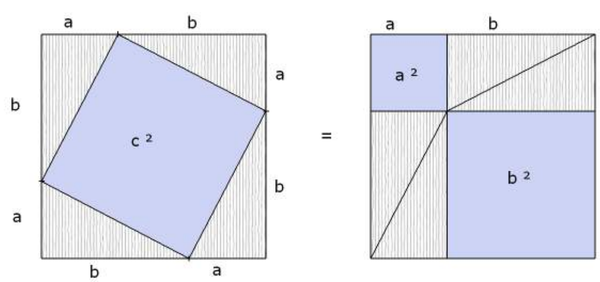


Figure 1.2: Explanation of Proof #1 of Pythagorean Theorem

will be more opportunities for aesthetic experience in the pages ahead.) In particular, I'll try to answer the following questions.

> [*Meaning*] What *is* aesthetic experience in mathematics?[1] What does it mean to find a piece of mathematics beautiful? What does mathematical beauty have in common with the beauty of sunsets, symphonies, ceramics?
>
> [*Role*] What role(s) do beauty and aesthetic experience play in mathematical practice? (How would mathematics change if mathematicians didn't have aesthetic responses?)

The little experience you just had (or didn't) may seem unremarkable. But ask a pure mathematician about it, and you may be surprised at the importance she attaches to such trifles. Thus the number theorist G.H. Hardy claimed that

> Beauty is the first test: there is no permanent place in the world for ugly mathematics.[2]

By contrasts, a philosopher is likelier to echo John Dawson, Jr:

> Some new proofs are thus presented purely as such, because they are deemed to be particularly novel, ingenious, or elegant (...) Such proofs frequently involve clever insights of quite limited generality. Nonetheless, they are often prized for their beauty.[3]

[1] Note that I am choosing to focus on aesthetic experience, rather than beauty. I will explain this choice in Section 4.

[2] [Hardy, 1992] Hardy, *A Mathematician's Apology.*

[3] [John Dawson, 2006] Dawson, "Why do Mathematicians Re-Prove Theorems?"

This isn't exactly a glowing endorsement of the pursuit of aesthetic factors in mathematics. There is growing agreement among philosophers of mathematics that the values which drive mathematical practice go beyond knowldge, truth, and justification, and include mathematical depth, fruitfulness, explanatoriness, understanding, purity of method, visualizability, simplicity, etc.[4] As Dawson's statement illustrates, beauty and elegance are the black sheep of this family of mathematical values. Again and again, philosophers interested in one or another of the values on the list defend their choice of object of study by contrasting it with the suspect realm of aesthetic value. To take another representative example, in his book *Beause Without Cause* [Lange, 2016] Marc Lange remarks:

> One reason that mathematical explanation has received relatively scant philosophical attention may be the temptation among philosophers to believe that when mathematicians apparently characterize some proof as explanatory, they are merely gesturing toward an aesthetically attractive quality that the proof possesses (such as elegance or beauty)—a quality that seems to be very much in the eye of the beholder (like a proof's being interesting, understandable, surprising, pleasing, or witty).

Again and again, philosophers of mathematics relegate aesthetics to the sidelines. Again and again, mathematicians bring up the uncomfortable topic.[5]

On Dawson's and Lange's view, if we ripped out the most beautiful plants from the garden of mathematics, the mathematician would go about

[4]For an overview of research in philosophy of mathematical practice, see [Mancosu, 2008].

[5]While Dawson is a professor of mathematics, his essay is published in the philosophical journal *Philosophia Mathematica*, and so I take it to be representative of the philosopher's outlook.

her day much as she had before—a little sad to see a few "clever insights of limited generality" wither and die, perhaps, but eager to keep growing her nourishing but ugly carrot-theorems. For a mathematician like Hardy, on the other hand, to remove the beauty would be to tear down the garden, to transform it into a wasteland unfit for human cultivation. This is a striking disagreement about what I've called the *Role* question.[6]

Who is right—the wary, disdainful philosophers or the gushing aesthete-mathematicians? Are aesthetic qualities a mere bonus a mathematician chances upon in her serious pursuit of truth, or is their pursuit the driving force behind her search?

This essay starts with a presumption in favor of the mathematicians. After all, they're *mathematicians*! No one else is in a better position to know what their discipline is like and why it's worth studying.

Other things equal, then, I will prefer answers to our questions which satisfy the following constraints.

> [*Real Beauty*] Mathematicians chose the word "beautiful" for a reason. Aesthetic experience in math has something in common with paradigm cases of beauty.
>
> [*Real Difference*] If mathematicians didn't have aesthetic experiences, mathematical practice would change noticeably.

In slogan form: *real beauty makes a real difference.* Of course, these constraints aren't the last word. Mathematicians *may* be confused about

[6] The mathematician Viktor Blåsjö draws attention to this disagreement in [Blåsjö, 2018]. This essay is indebted to correspondence with Dr Blåsjö.

the nature of their enterprise. I'm merely suggesting that, all else equal, we should try to give answers to *Meaning* and *Role* that take their aesthetic claims at face value.

I think we *can* give plausible answers to my questions which satisfy these constraints. To do so, I'll enlist the help of another mathematician: Henri Poincaré. I'll contrast Poincaré's view with that of philosopher James McAllister, highlighting the way McAllister's view fails to satisfy *Real beauty makes a real difference*. I will then introduce three visual proofs in a way which I hope highlights their aesthetic virtues, and I'll argue that Poincaré's view applies well to these examples. While ultimately I think there are types of aesthetic experience which are not of Poincaré's preferred "aha" variety, I believe his view is a persuasive account of one important type of mathematical beauty. In the final sections, I will answer two common objections which threaten the truth of *Real Beauty*: that all beauty is perceptual and that pleasure at purportedly beautiful mathematics is egotistical and hence not genuinely aesthetic.

1 Poincaré's Aesthetics

Poincaré's essay "Mathematical Creation"[7] gives aesthetic responses a central role in mathematical practice. He goes so far as to say that

> the feeling of mathematical beauty, of the harmony of numbers and forms, of geometric elegance (...) is a true esthetic feeling that all real mathematicians know, and surely it belongs to emo-

[7] [Poincaré, 1910]. The essay also appears as chapter 3 of section i of [Poincaré, 1914].

tional sensibility.[8]

While for Hardy, there is no mathematics without beauty (at least, not for long), for Poincaré, there are no mathematicians (at least, not "real" ones) without an aesthetic response.

For Poincaré, this "true aesthetic feeling" plays at least three important roles: it allows mathematicians to create, to understand, and to unify. In this essay, I will focus on the first two: creation and understanding.

1.1 Role 1: Creation

Poincaré claims that the process of solving a mathematical problem (such as coming up with a proof) comes in three stages: two conscious ones and one unconscious. During the first conscious stage, the mathematician sits at her desk with pencil and paper, scribbles furiously, and actively attempts to solve the problem at hand. Then, if the problem is hard enough, she gets stuck. It's time to go for a walk, collect pebbles on the beach, and let the second stage—unconscious work—begin. Poincaré hypothesizes that during this stage the unconscious runs through many combinations of ideas related to ones which the mathematician had considered in the previous stage. If all goes well, this stage ends with an "aha!" moment: the idea of a solution comes to the mathematician in a flash, accompanied by a sense of conviction. In the final, conscious, stage this solution is developed, verified, and sometimes written up.

What role does the aesthetic feeling play during this process? Poincaré

[8]Unless otherwise specified, this and all further quotes are from [Poincaré, 1910].

thinks that there must be a mechanism by which the subconscious stage comes to an end—a mechanism by which some ideas, but not others, come to conscious attention. He claims that the aesthetic feeling provides this mechanism. Speaking metaphorically, the unconscious waves an attractive idea in front of the conscious to get its attention. Mathematicians run through many ideas unconsciously—but only ones which are sufficiently beautiful come to conscious attention.

If this picture is correct, a mathematician without an aesthetic sense wouldn't be able to get out of the unconscious stage of mathematical thought; in Poincaré's words, she wouldn't be able to create. Furthermore, the feature of mathematical beauty which allows it to play this role is one it shares with garden varieties of beauty. Beauty attracts us, compels our attention. The flashily beautiful things like sunsets, from which many of us first learned about beauty, attract us by their literal vividness—but even subdued patches of mist possess a (metaphorically) vivid glow insofar as we find them beautiful.[9]

According to Poincaré, then, if mathematicians didn't have the thing he calls "the feeling of mathematical beauty," they wouldn't be able to solve any mathematical problems difficult enough to require unconscious work. Furthermore, to do its job, this feeling has to possess a certain phenomenal vividness or attractive force. It's not just that mathematical creation

[9]In fact, beautiful images (at least, erotically beautiful ones) "waved" in front of the subconscious do appear to guide attention in a way strikingly similar to what Poincaré posits: in [Jiang et al., 2006], subjects were subliminally presented images of nudes in one half of their visual field. Those who were attracted to the appropriate sex tended to direct more of their *conscious* attention to that half of the visual field. Thanks to Zoe Jenkin for this reference.

requires a feeling which mathematicians happen to call "aesthetic"—it requires a feeling that is similar to the feeling of the beauty of a sunset in this striking respect.

1.2 Role 2: Understanding

To introduce the second role, Poincaré raises a naive-sounding puzzle: if mathematics consists of series of logical inferences starting from self-evident truths, how is it possible that so many people aren't any good at it? After all, the laws of logic are supposed to be so self-evident that one can't even think without being subject to them. It makes some sense that not everyone can *create* mathematics. But how can it be that some people can't even *understand* proofs? Failures of memory are an obvious culprit—during a long series of inferences, the mind gets lost and is apt to modify the premises slightly, which results in errors. But Poincaré claims that many mathematicians (including himself) in fact have quite bad memories. They need some special aptitude to make up for this lack. According to Poincaré, this aptitude consists in a feeling for the structure of pieces of reasoning.

> A mathematical demonstration is not a simple juxtaposition of syllogisms, it is syllogisms placed in a certain order, and the order in which these elements are placed is much more important than the elements themselves. If I have the feeling, the intuition, so to speak, of this order, so as to perceive at a glance the reasoning as a whole, I need no longer fear lest I forget one of the elements, for each of them will take its allotted place in the array, and that without any effort of memory on my part.

The context of this quote suggests that this feeling is the "true aesthetic feeling" that Poincaré mentions earlier. But why think that such a feeling for a piece of reasoning as a whole has anything to do with beauty and aesthetics?

Well, imagine hearing a musical piece in a new genre. At first, you only hear a tonal jumble. You have no sense of overarching structure, don't feel the flow of the music or even recognize repeated phrases. But if you listen to the piece a few times, you may come to hear it differently. Your attention glides smoothly between the notes. You're fully in the moment, but at the same time you have a sense of the piece as a whole, of the order in which the notes are placed. It's apt to call the ability you've just gained—the ability to hear the piece as a smoothly flowing whole—an aesthetic sense. On Poincaré's view, something strikingly similar is at play in mathematics.

If the second part of Poincaré's view is correct, then, a would-be mathematician without an aesthetic sense could only perceive proofs as chains of inferences, without discerning the overall logical structure of the reasoning. She would be less a mathematician than a calculator. Furthermore, a comparison to music suggests that the capacity to unify is continuous with the capacities we ordinarily term "aesthetic."

The two parts of Poincaré's view are related. The unconscious must come up with ideas that are not only vivid and attractive enough to get the attention of the conscious, but compressable enough to be experienced in an "aha!" moment that the mathematician can then unpack in the final stage of creation. In other words, what the unconscious "waves" in front of the conscious is a unified, compressed version of a proof. And when a student follows a piece of mathematical reasoning, she must create something very much like this "aha" idea—she must hold in her head a unified version of the piece of reasoning as a whole.

Poincaré corroborates this interpretation.

> It seems to me then, in repeating a reasoning learned, that I could have invented it. This is often only an illusion; but even then, even if I am not so gifted as to create it by myself, I myself re-invent it in so far as I repeat it.

In other words, understanding is creation in reverse. Given someone else's proof—the product of the third stage of mathematical creation—I compress it into a single "intuition" very much like the one my unconscious would have given me, had I come up with the proof myself. Creation and re-creation share an "aha!" moment.

2 The View's Virtues

We're now in a position to infer Poincaré's answers to our questions: *Meaning* and *Role*.

Appreciating the type of proof Poincaré is interested in involves mental unification. In other words, a mathematician who finds this type of proof beautiful can see the structure of the reasoning "at a glance"—or nearly so.[10] At the very least, she has a sense of the "guiding thread" of the reasoning, a sense which allows her attention to glide smoothly between the inferences.

That's *still* not quite right. Yawning while unifying is hardly an aesthetic experience. What's missing? Plausibly, pleasure. We *like* looking at beau-

[10] It may be that more complicated pieces of mathematics are aesthetically appreciated in ways that don't involve mental unification.

tiful paintings, listening to beautiful music, hiking in beautiful mountains. To find these things beautiful is not to dispassionately judge them to be so, but (in part) to take pleasure in them. So too with beautiful mathematics. To find this type of proof beautiful, then, is to take pleasure in one's mental unification.[11]

This modification is in line with Poincaré's pronouncements; in passages such as the following he explicitly ties beauty with pleasure or delight.

> The scientist does not study nature because it is useful; he studies it because he delights in it, and he delights in it because it is beautiful.[12]

What, then, does Poincaré's mathematician mean when she says that she finds a proof beautiful? I think she means what we all mean in such cases: that she takes pleasure in mere contemplation of the proof. But the pleasure we've been discussing here is a *subtype* of the genus "aesthetic pleasure:" namely, pleasure in mental unification.

For Poincaré, then, a central type of (mathematical) aesthetic experience is pleasant mental unification. Does this answer satisfy *Real beauty*?

We've seen some reason to think so: pleasant mental unification is a type of aesthetic experience at least in music. Similarly, a painting may look like a jumble of brushstrokes until—suddenly, beautifully—it coheres into a whole, allowing one's attention to flow smoothly between the parts. Many critics and philosophers have thought that unity amid variety is an

[11] Or, perhaps, to take pleasure in the proof's capacity to cause such mental unification.
[12] [Poincaré, 1914] Poincaré, *Science and Method*, Ch. 1 ("The Choice of Facts.")

aesthetic value.[13] I believe that this is precisely because pleasant mental unification is a form of aesthetic experience.

Without aesthetic experiences, mathematicians wouldn't be able to create, unify disparate mathematical domains, or understand proofs in more than the minimal sense of merely stringing together inferences. That, in broad strokes, is Poincaré's answer to *Role*. Using a finer brush: if mathematicians didn't find some ideas *attractive*—that is, if they didn't take pleasure in them—they wouldn't be able to create. And if they couldn't mentally unify proofs, they could neither create nor even understand. Furthermore, while a mathematician who never felt aesthetic *pleasure* could theoretically *understand* proofs as long as she could still mentally unify them, she would be unlikely to stay motivated to do mathematics for long. Together, these roles amount to a strong case for *Real Difference*.

According to Poincaré, then, the experience of mathematical beauty has a rich phenomenology which is importantly similar to the experience of artistic and natural beauty, and which makes a genuine difference to mathematical practice. If the ideas mathematicians call beautiful weren't *attractive* and phenomenally vivid to them, or if they weren't *unified*, mathematicians would be unable to create new proofs or to have more than a superficial understanding of old ones.

The virtues of Poincaré's view come into especially sharp relief when we compare it with accounts taken from the philosophical literature. As an illustrative example, let's consider James McAllister's account of scientific

[13]The classic source is [Hutcheson, 2008] Hutcheson's *An Inquiry Into the Original of Our Ideas of Beauty and Virtue.* John Dewey [Dewey, 1934] lists "unity" as one mark of an aesthetic experience.

beauty.[14] (Note that none of what follows is a direct criticism of McAllister's theory, since McAllister is defending a view of *scientific*, not mathematical, beauty. I am merely using the view as a backdrop against which Poincaré's account can be seen to better advantage.)

Which set of qualities is found beautiful by a scientist varies from epoch to epoch. McAllister hypothesizes a process—called the aesthetic induction—which accounts for such changes of standards. On his view, the qualities scientists find beautiful at any given time are ones which have correlated with empirical adequacy in the recent past; as theories rise and fall, so do aesthetic standards. And since scientists are more likely to propose, study, and test the theories they find beautiful, McAllister infers that beauty makes a real difference to scientific practice.

This inference is too quick. The thing scientists *call* "beauty" makes a difference to science—but McAllister gives us no reason to think that such "beauty" is the genuine article. To see this, consider his answer to *Meaning*: that the experience of beauty is pleasure caused by qualities previously found in empirically adequate theories. What does this have in common with ordinary aesthetic experience? Well, pleasure—but beyond that, the view is ominously silent. For all McAllister says, scientists could have called theories "nice" or even "likely to be empirically adequate" ("promising" for short) instead of "beautiful."

To be sure, McAllister's view is *consistent* with satisfactory answers to *Meaning*. But note that his answer to *Role* is independent of such extensions. For McAllister, scientists' "aesthetic" rankings of theories matter—but their

[14] [McAllister, 1996] *Beauty and Revolution in Science.*

aesthetic *experiences* do not. If scientists ranked theories not by beauty, but by niceness or promisingness, science would go on exactly as before. If they didn't have aesthetic experiences at all, and merely felt pleasure in response to some theories but not others—or even dispassionately preferred some theories to others, perhaps by explicitly reasoning that these are likelier to be empirically adequate—science would go on exactly as it does now. This is in marked contrast to Poincaré's view, on which it is the mental act of pleasing unification that makes a difference to mathematics.

In the next section, I'll argue that this is a damning objection.[15] Real beauty makes a real difference only if it's the *experience* of beauty that makes the difference.

3 Experience First

Philosophical aesthetics is (among other things, perhaps) concerned with qualities or properties such as beauty, sublimity, cuteness, ugliness; with aesthetic judgments or evaluations such as "the Grand Canyon is sublime" or "Kerouac's *On the Road* is painfully boring;" and with aesthetic experiences or moments of appreciation such as looking at Van Gogh's 1888 *Wheat Field* and feeling a sense of profound, almost spiritual calm. My view is that aesthetic experience is vastly more valuable than the other two members of this trio (quality and judgment). I would barter the accuracy of my aesthetic judgments for more aesthetic experiences in the blink of an eye. In this section, I'll try to convince you that you should too.

[15] Again, all this is under the (false) hypothesis that McAllister's view is meant to apply to mathematics instead of science.

Imagine coming across a genie, hunched over a heavy tome. He turns your way and explains: "This is the *Book of Beauty.* Take it, and you'll have perfect aesthetic knowledge: answers to the questions 'Is this beautiful?' and 'Which of these is more beautiful?' for any thing or pair of things for which there is an answer at all." You reach out your hand, and something glints in his eye. "There's a modest price, of course. In return, I only ask for your aesthetic experiences. Take this book, and you won't gasp in front of patches of sunlight on your gerberas anymore. You'll know which things have the highest beauty—but you won't feel any different in their presence."

Would you take such a deal? I would retract my hand in horror. To take the book would be to know that a sunset is more beautiful than any I have ever seen... and shrug my shoulders at it. Evaluation is worthless without appreciation.

Maybe you're more altruistic than me. Would the entries in the *Book* produce valuable aesthetic experiences in many *other* people? If so, you might take the deal for their sake, becoming a blind oracle so that they might see. "Come outside!" you'd entice your neighbors. "There's never been a sunset as beautiful as tonight's!" Watching their faces light up with delight, you'd feel happy for them—truly happy. With time, your joy would be ringed by barely a tremor of regret.

Whether or not you'd take the deal, though, the genie would have taught you a valuable lesson: *someone* must feel beauty for it to matter.

Now suppose the genie offered a *Book of Scientific Beauty* to McAllister's scientist.[16] Would she take it? Well, she'd be a little sad to lose her aesthetic

[16]Perhaps there is a new edition of this book for each generation of scientists. The

pleasure—but think of the payoff! Perfect knowledge of which theories are similar to past empirically adequate ones! When you're a scientist, that's nothing to scoff at.

What if Poincaré's mathematician were offered a *Book of Mathematical Beauty*? At first, he'd shudder at the thought. To lose aesthetic experiences is to lose understanding and the capacity to create. It is to be a mathematician no more. Then, if he were altruistic enough, he might ask: is true beauty correlated with more profound understanding for *other* mathematicians? If so, he might choose to become a blind oracle so that they might see.

Notice how closely this response parallels our reaction to the first genie. I think that is good news for Poincaré. If the thing which makes a difference to mathematics is real beauty, then mathematicians should value it for the same reasons that we, ordinary aesthetes, value garden-variety beauty. Therefore they should be willing to take roughly the same deal with the genie as us. Furthermore, our thought experiment shows that what we value about beauty is primarily the *experience*, and only secondary the ranking. The judgments we make using the word "beauty" are, above all, tools for sharing our experiences with others, for getting them to stand where we're standing. To say "this meadow is beautiful" or "this line is graceful" is to do nothing more or less than to gesture at an aesthetic experience that is to be had in front of the meadow or line. What really matters to lovers of beauty is the experience, not the property. That is also what matters to Poincaré's

original *Book of Beauty* may also come in editions, perhaps relative to cultures as well as epochs. This wouldn't affect my reactions to the thought experiment.

mathematician—and for that reason she (unlike McAllister's scientist) is a true aesthete: real beauty makes a real difference to her.

4 Proofs Without Words

So far, our Poincaré-inspired picture of mathematical beauty and creation is vague and metaphorical. What does it mean for the unconscious to wave ideas in front of the conscious? And what even *are* ideas in this context? Here's a working definition: a mathematical idea is a mental representation which is a compressed version of a proof or a step in a proof. In one way or another, an idea encodes information about a proof—but the encoding can take a bewildering number of forms. An idea might present itself as a belief or hypothesis, which may or may not be explicitly articulated in propositional form. One of the moments of inspiration Poincaré describes is like this: he was suddenly struck with the idea "that the transformations [he] had used to define the Fuchsian functions were identical with those of non-Euclidean geometry." But the phenomenal mark of having formed an idea needn't be an articulated belief—it may be as rich as a mental "video" or as impoverished as a bodily sensation, a sense of conviction, a feeling that one knows how to go on.

This makes mathematical ideas a particularly difficult subject of study. Nonetheless, there is at least one family of cases in which it's possible to say much about these ideas.

This is the case of so-called "proofs without words." These "proofs" are images which distill the key move or moves of a proof. In the words of the editor of a volume of such such proofs:

> Proofs without words are pictures or diagrams that help the observer see *why* a particular statement may be true, and also to see *how* one might begin to go about proving it true.[17]

According to this characterization, then, proofs without words serve two functions: they at least partially explain why their theorem is true and they provide a blueprint for the creation of a more formal proof.

It's my contention in this section that proofs without words are publicly shareable counterparts of mathematical ideas—they are compressed versions of proofs which are reliable blueprints for the creation of the full proofs. Furthermore, such proofs often possess unity and phenomenal vividness (Poincaré's two marks of beauty), in forms which help clarify these two properties.

An image can store much more information than a sentence. For visual beings like us, such information is also easily accessible. So it would make sense for many of the ideas the unconscious presents us with to take a visual form. Indeed, the language Poincaré uses in describing mathematical ideas is strikingly visual.

> If I have the feeling, the intuition, so to speak, of this order, so as to *perceive at a glance the reasoning as a whole*, I need no longer fear lest I forget one of the elements, for each of them will take its allotted place in the array, and that without any effort of memory on my part. (Emphasis mine.)

As we'll see later, it's possible to metaphorically "perceive at a glance

[17] [Nelsen, 1993] *Proofs Without Words.*

the reasoning as a whole" even in cases in which there is no visualization of a proof. Nonetheless, the literal case is a natural starting point.

There's reason to think, then, that our Poincaré-inspired view illuminates the realm of proofs without words particularly well. This is some evidence for the view, since this realm is often thought to be particularly rich in beauty—the proportion of strikingly beautiful proofs among proofs without words seems to be larger than the corresponding proportion among algebraic proofs. Why would that be? The answer would be obvious if beauty were always a perceptual property, but as we'll see in the final sections of this essay, there's little reason to think that's the case. But the Poincaré-inspired picture provides a better explanation: if compression or decompression is a necessary component of one type of aesthetic experience of mathematics, and the existence of a proof without words is a guarantee for the existence of a compressed version of a proof, then proofs without words are to that extent more likely to be beautiful. (Note, however, that compression is a gradable property—a visual proof is more or less compressed depending on the difference in simplicity between the diagram and corresponding proof. When this difference is very small, we probably won't find the proof very beautiful.)

This explanation accords well with the language fans of proofs without words use to describe the beauty of such proofs. In his column exhibiting visual proofs, Martin Gardner remarks:

> In many cases a dull proof can be supplemented by a geometric analogue so simple and beautiful that the truth of a theorem is

almost seen at a glance.[18]

Seeing at a glance is presented as a consequence of the beauty (and simplicity) of visual proofs. Here, then, Gardner and Poincaré are in striking agreement.

Since proofs without words are often held to be strikingly beautiful, then, success in fitting them into our picture would be evidence in favor of this picture. In the following sections I will do just that to three proofs of the Pythagorean Theorem (henceforth "PT.")

4.1 Proof # 1

I'll begin with the proof from the introduction (Figure 1.1). By moving right-angled triangles around inside a square of side $a + b$, this proof allows us to alternately see the white area as a square of side c and as two squares of sides a and b. We see (almost) at a glance that the theorem holds; the proof amounts to an "aha" moment.

Let me spell this out. Given a formal counterpart to Figure 1.1 (that is, a sequence of logical inferences corresponding to the steps carried out in the image), I would struggle to follow the proof unless I created something very much like the figure (whether on a sheet of paper or in my mind). Otherwise I would just be mindlessly stringing together inferences.[19] But as soon as I create something like this figure, my mind will be able to run smoothly through the inferences. I will have a sense for the reasoning as a whole.

[18] In [Gardner, 1973], his column in *Scientific American*, as cited in [Nelsen, 1993] Nelsen, *Proofs without Words*.

[19] Note, however, that there can be a geometric component even to seemingly "algebraic" reasoning. See [Giaquinto, 2005].

And if I find the proof beautiful, it will be precisely at this moment—the moment of "seeing at a glance."

Conversely, if a mathematician's subconscious "waved" Figure 1.1 in front of her in an "aha" moment, the mathematician would see the essence of the proof at a glance and would be able to quite easily convert the image into a more formal proof.

In the following two sections, I'll show in more detail how our experience of the proof fits into the Poincaré picture. This demonstration will have two parts. First, I'll show that contemplating the proof without words itself, without accompanying "decompressed" version, can give rise to a pleasure which is strikingly akin to our pleasure at some art. I will suggest that the capacity for affording such pleasure is one form of the phenomenal vividness which may cause the mathematician to pay attention when the subconscious "waves" such a proof before her. Second (this is the core of the argument), I'll show how to decompress the proof into a more formal presentation. This will show, first, that the visual proof can indeed be thought of as a compressed version of a more formal proof, and, second, that moving between the compressed and decompressed versions gives rise to a further pleasure. I will argue that this pleasure is rightly characterized as aesthetic—a form of appreciating "unity amid variety."

4.1.1 The proof without words

Let's look at the proof very slowly. First, forget about the Pythagorean Theorem and just look at Figure 1.1. What do you see? Two big congruent squares containing four congruent triangles each, arranged in different ways.

Imagine shifting from the left-hand figure to the right-hand one by moving the triangles into the corners. Now shift your focus. In addition to four triangles, the big square contains some white areas, which we have so far treated as negative space. Look at those white areas—two squares in the left-hand picture and one square in the right-hand picture. In both pictures, they're what's left of the big square once you remove the four triangles. That is, the white part of the left-hand image takes up as much space as the white part of the right-hand image.

What did you think of this little exercise? I find it quite nice. It's nothing spectacular, but by making us notice relationships between shapes, move pieces around, and shift from background to foreground, it transforms an unremarkable image into something interesting.

Let's try a similar exercise with the image in Figure 1.3. Notice the four white rectangles, and focus on the ones which clearly aren't square. Notice how one side of the left-most white rectangle is the same length as the side of the big red square. Now imagine sliding the white "vertical" rectangle down next to the yellow square. The sides seem to be about the same length. Finally, take the right-most white rectangle and shift it next to the white square. The lengths also seem to be similar. Now look at the whole image again, paying attention to the way each white rectangle has a companion square. Doesn't it please you more now?

Figure 1.3 is a Mondrian-inspired piece of digital art. I've guided you through this experience of the work to show that we often approach visual art, especially abstract art, in a way which is strikingly similar to our experience of Figure 1.1. We notice relationships between figures, imagine

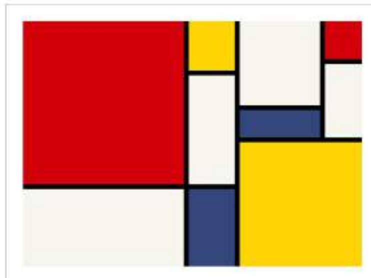


Figure 1.3: Mondrian Style

moving parts of the figure around, shift from focusing on small pieces to taking in the whole image, switch from perceiving an element as foreground to perceiving it as background, etc. (This whole process might be an aspect of what Kant calls "the free play of the imagination and understanding."[20]) In both cases, this enhances our appreciation of the image and compels us towards further exploration. There are further things to notice: in Figure 1.1 you might try to imagine the shape of the paths traced by the three triangles as they move to their corners, in Figure 1.3 you might note the way the white shapes cluster around the diagonal of the image.

All this is aesthetic appreciation of the visual sort, and it explains one aspect of the appeal of Figure 1.1. Nonetheless, if this were *all* this appeal amounted to, mathematical beauty would be rather impoverished. After all, figure 1.3, which is hardly a paragon of visual beauty, affords more opportunities for such visual play than Figure 1.1.

[20] [Kant, 2000] *The Critique of Judgent.* I'm also indebted to Angela Breitenbach's [Breitenbach, 2015] Kantian view of mathematical aesthetics.

Figure 1.4: Seurat, *A Sunday Afternoon on the Island of La Grande Jatte*, 1884

Thankfully (for the fan of mathematical beauty), there is more to the appeal of Figure 1.1 than this. So far, we've abstracted away from the figure being a proof. Take this aspect into account, note that the equality of the white areas in the left and right-hand images *is* the Pythagorean Theorem, and the experience causes a whole new level of pleasure. Now we're looking at the image not just as a piece of abstract art, but as a *representation*. This is akin to noticing the symbolic function of a spatial relationship in a painting.

For instance, notice the way the figures in Seurat's *A Sunday Afternoon on the Island of La Grande Jatte* (Figure 1.4) get smaller not just as we move up parallel to the vertical edge of the canvas (i.e. deeper in the represented three-dimensional space), but also as we move from right to left parallel to the long horizontal edge of the painting. For instance, the lady in the orange dress looking out over the river on the left is much smaller than the lady holding the hand of the little girl in white, despite being in roughly the same plane as her.

Now this feature can be appreciated as just a formal property of the

Figure 1.5: Seurat, *Bathers at Asnières*, 1884

painting—we can enjoy following the descending sequence of figures from right to left.[21] But it's also possible to find interpretations for this feature which enhance our appreciation. Here's one such interpretation.

A Sunday Afternoon on the Island of La Grande Jatte has a companion painting—*Bathers at Asnières* (Figure 1.5), which represents the left bank of the Seine, frequented by working-class people. A boy in the water seems to be calling out to the other side.

If we place the two images side by side—Asnières on the left, Grande Jatte on the right—the descending sequence of figures towards the left amounts to the existence of a second vanishing point on the left, Asnières-bank of the Seine. It's as if Seurat were posing the question: where are these people going? Deeper "inside" the painting, parallel to the Seine, within their own class—or does the future lie on the other bank? The formal feature of the painting now plays a representational role.

In a similar way, the equality of the white areas in Figure 1.1 is pleasant to behold, but our pleasure becomes much deeper once we realize the

[21]Of course, it's also possible to be *displeased* by this feature and see it as a failure of perspective.

import of this equality. Of course, there are many differences between our experiences of the Grande Jatte and of Figure 1.1. For one thing, the white squares in the latter don't just *represent* the Pythagorean Theorem—they also demonstrate its truth. By contrast, if the Grande Jatte gives any evidence at all for the claim it makes (that the future of France might lie with the working class), the evidence is much more meager than a geometric demonstration. But I didn't introduce this example to make an exhaustive list of the similarities and differences between mathematics and visual art. Rather, I wanted to point to a specific structural similarity: that visual proofs as well as art can lead to a visual play which is further enhanced by noticing the representational role of the visual properties.

Hopefully, I've caused you to feel some pleasure by mentally playing around with the pieces of Figure 1.1, and then noticing that it provides a demonstration of the Pythagorean Theorem. Perhaps you were also compelled to exclaim "how beautiful!" If so, the family resemblance between your experience and that in front of a Mondrian or Seurat painting provides some reason to take your exclamation at face value: the proof *is* beautiful. Now let's see how it fits in Poincaré's picture.

4.1.2 The proof with words

Imagining the triangles in our proof sliding into their corners does double duty: it is not only a visually pleasing thing to do, it also allows us to see why the Pythagorean Theorem holds. In fact, the features of Figure 1.1 play *triple* duty: in addition to explaining why the theorem holds, they play the function of a blueprint for a more formal proof. This is the function to

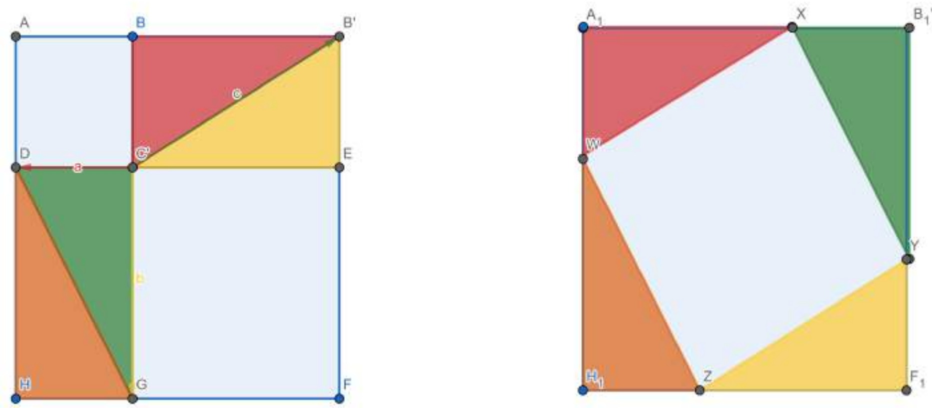


Figure 1.6: Pythagorean Theorem #1, illustration of formal proof.

which we now turn.

Below, I'll provide one way of translating Figure 1.1 into a formal proof. You might want to try your hand at writing out the proof yourself first, to test my claim that the proof without words really is decompressible into a formal proof with little effort.

Proof. Let a, b, and c be the lengths of edges of a right-angled triangle, as in the Pythagorean Theorem. Draw a square $ABCD$ of side a and square $CEFG$ of side b, touching at corner C, with respective edges parallel. Then extend the edges to form a square $AB'FH$ of side $a+b$. (Figure 1.6.)

Now draw auxiliary lines DG and $B'C$. Translate triangle DCG along vector $\overrightarrow{CB'}$, triangle $B'EC$ along vector $\overrightarrow{CG}$, and triangle $B'CB$ along vector $\overrightarrow{CD}$.

Let points D and B' be mapped to point D' and B'' by the mappings defined by vectors CB' and CD, respectively. $D'B'$ is the image of DC

under a translation, so it has length a. But B'' is the image of B' under a translation of a vector of length a parallel to BB'', so it has length a too. So $D' = B''$. Similarly, points G and B' get mapped onto the same point on line $B'F$. Therefore the translations define a quadrilateral $WXYZ$ (the second image in Figure 1.6).

Now the areas of squares $ABCD$ and $CEFG$ add up to the area of figure $WXYZ$ (since the latter is equal to the area of square $AB'FD'$ minus the area of four congruent triangles.)

Furthermore, the edges of figure $WXYZ$ all have length c (since they are the hypotenuses of right-angled triangles of sides a and b. Finally, the angles of $WXYZ$ are all right angles, since (for instance), $\angle DZW + \angle WZY + \angle FZY = 180°$ and $\angle DZW$ and $\angle FZY$, being the two non-right angles of a right-angled triangle, add up to $90°$.

Therefore $WXYZ$ is a square of side c, and its area is equal to the sum of the area of the given squares of sides, respectively, a and b. □

Given the "proof without words," writing out this more formal proof was almost entirely straightforward (though, as is often the case with spelling out geometric proofs, somewhat finicky). I had to shuffle pieces around a bit to fit the ideas implicit in Figure 1.1 into proof-writing conventions. For instance, while in Figure 1.1 we're given the big square divided up into two squares and four triangles as a single perceptual item, with no indication of an order of construction, in a proof of the Pythagorean theorem, where the areas of squares with sides a and b are given, it's customary to start the construction with these squares.

Then I needed to figure out how to formally characterize the notion of "sliding" the triangles into their corners—but the choice was an entirely natural one since "translation" just is the mathematically formal counterpart of "sliding". Finally, writing out the proof makes salient the necessity of confirming that these translations really do define a square ($WXYZ$), which might have gone unnoticed when we first accepted Figure 1.1 as evidence for the truth of the Pythagorean Theorem. Before writing out the formal proof, we might have simply assumed that this was the case, either based on perceptual evidence (the figure *looks* square in the image), or based on inductive inference (a natural construction like the one in Figure 1.1 will tend to produce symmetric results like squares.)

Paying attention to the process of decompression, then, gives us reason to believe that our formal proof *is* a decompressed version of the proof without words—with the one caveat that the argument that figure $WXYZ$ really is a square (or even a quadrilateral!) is a genuinely new—though elementary—addition to the proof without words.

If mathematicians' "aha" moments take forms akin to Figure 1.1, the above discussion explains how these moments of illumination can be strong evidence for the truth of the theorem and the existence of a proof corresponding to the "aha" moment while falling short of the certainty of a deductive proof. It really is possible to see a lot "at a glance" from Figure 1.1: that the four triangles will slide into their corners, that the left-hand and right-hand figures have the same areas, etc. At the same time, there's room for error: the belief that figure $WXYZ$ is a square might be formed based solely on perceptual evidence. When she puts pen to paper, the mathematician might find herself unable to formally demonstrate that this is the case.

We've seen reason to think that the visual proof is an aesthetic object in its own right, independently of its relationship to the formal, decompressed proof. Nonetheless, I believe that full aesthetic appreciation of mathematics involves having both the compressed *and* decompressed versions of the proofs—and shifting between the two perspectives.

To see how the decompressed version of the proof can aid appreciation, consider the way we've replaced the informal notion of "sliding" with the formal notion of "translation" when moving to the more formal proof. We've translated the triangles by vectors $\overrightarrow{CD}$, $\overrightarrow{CB'}$, and $\overrightarrow{CG}$ in Figure 1.6. But these are vectors of length a, b, and c—the construction is elegantly defined out of the most basic ingredients of the Pythagorean Theorem! Noticing this makes me look at the proof without words with an appreciative eye.

In general, shifting perspectives in this way—between the detailed and linear formal proof and the (almost) instantaneous informal proof—is a particularly pleasant way of engaging with the proof. It's akin to looking at a large painting and shifting between admiring the individual parts—the figures, the colors, even the texture of patches of paint—and stepping back to take in the structure of the composition as a whole.

We can shift between perspectives in this way whenever a formal proof has a visualization—but just like in painting, only particular combinations of local detail and holistic structure give rise to aesthetic pleasure. A visual proof of an obvious statement, like the proof in Figure 1.7 of the claim that every square can be divided into two rectangles isn't really a *compressed* version of its formal counterpart—there is no work at all to be done in

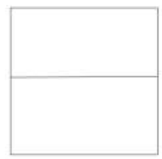


Figure 1.7: An overly simple proof

decompressing the proof. At the other extreme lie visual proofs which are too detailed to be taken in at a glance.

The special power of Figure 1.1 arises in part from the fact that it demonstrates equality of area not by cutting up figures into smaller pieces to reassemble, but by embedding them inside of a larger figure. This makes the proof particularly "instantaneous" and "holistic"—even among the class of visual proofs.

Let me make an important clarification, though. I am not trying to propose a criterion separating the beautiful from the non-beautiful proofs. I am not claiming that the beautiful visual proofs are those which are maximally compressed, that hold a particularly large amount of information relative to their apparent simplicity. Rather, I am claiming that compression and decompression is one type of aesthetic *experience*. Experiences of using a simple mental representation to hold a large amount of information *in your mind* are one kind of aesthetic experience. Particularly simple images which get written up as particularly complex formal proofs are, for many people, likely to cause such experiences; visual proofs are likely to become mathematical ideas. But the real test of beauty is the experience itself: whether contemplating the proof causes you to pleasantly shift between the compressed and decompressed version.

To sum up, aesthetic appreciation of the proof in Figure 1.1 is a complex experience involving appreciating the visual proof itself, decompressing it into a more formal proof, and shifting attention between the formal and informal proofs. The beauty of the informal proof is in part due to its capacity to afford visual play, and in part due to the way these pleasing visual properties simultaneously allow us to see at a glance why PT holds. Decompressing the proof increases our aesthetic appreciation in part by helping us to notice new features of the visual proof (e.g. that the translations involved in it are by vectors of length a, b, and c) and in part by allowing us to shift between taking in the details of the proof and the structure as a whole (as one might in front of a painting).

In conclusion, proofs which have sufficiently compressed "wordless" counterparts fit particularly nicely into Poincaré's picture. A proof without words has all the right features to play the role of an idea "waved" by the subconscious during mathematical discovery: it can be taken in at a glance as well as decompressed into a formal proof. And it can play Poincaré's second role: that of a unified mental representation and aid to memory of a mathematician coming to understand a proof. Faced with the "decompressed" version of our proof above, unless I (mentally) create something very much like Figure 1.1, I will only be mechanically plodding along between the steps of the proof, without full understanding.

Beautiful visual proofs come in pairs: decompressed and compressed, a richly detailed and logically precise outer layer together with an intuitive, unified core—or, perhaps, soul. There is aesthetic pleasure to be found in

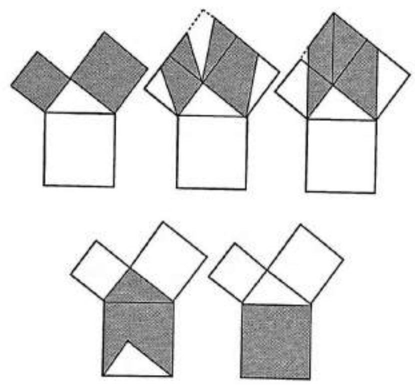


Figure 1.8: Proof #2 of Pythagorean Theorem. (Reproduced from [Nelsen, 1993].)

contemplating each part—and, especially, in shifting between them.

4.2 Proof #2

So far, I have (rather painstakingly) shown how one visual proof fits into Poincaré's model. But how well does the model fare with other test cases? In the next two sections, I'll introduce two further proofs of PT. (Focusing on proofs of a fixed theorem makes comparison easier). I'll show how our model illuminates the differences in aesthetic experience and judgment in these cases.

Our second visual proof of PT, then, is given in Figure 1.8. I recommend taking a moment to try to understand it and convince yourself that it works. Had enough? Then let me tell you how the proof works... or rather, how I *thought* it worked when I first saw it.

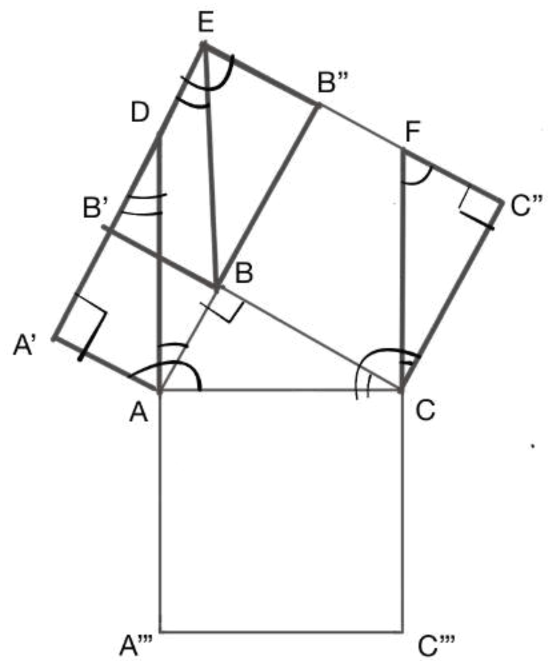


Figure 1.9: Explanation of Proof #2, part 1

The general structure is as follows: we cut up the small squares (of sides a and b) into pieces and rearrange those so that they form the big square (of side c), thus showing equality of area. To do that, we extend the edges of the squares as in Figure 1.9, getting points of intersection D, E, and F. Now the triangles with bolded outlines ($AA'D$, $BB'E$, BEB'', $CC''F$) are all congruent to our original triangle ABC. Why? Well, the corresponding angles are all equal (as marked in the figure) by repeated use of the fact that the two non-right angles in a right-angled triangle add up to 90°. And it's easy to show that each of the triangles has one side equal in length to the corresponding side in triangle ABC—e.g. BB'' has the same length as BC, because $BCC''B''$ is a square.

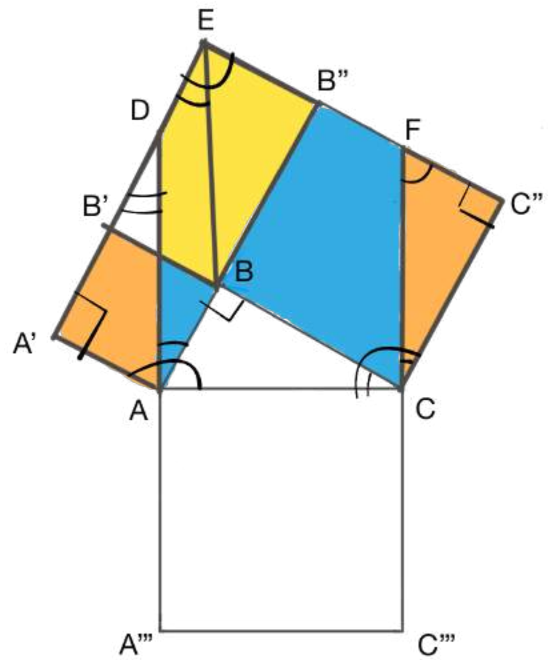


Figure 1.10: Explanation of Proof #2, part 2

This means that the area marked in orange in Figure 1.10 is equal to the area marked in yellow. Why? Clearly, area $EB''B$ is equal to area $CC''F$—as we've seen, these are similar triangles. And area $DEBG$ is the area of triangle EBB' (congruent to $AA'D$), minus the area of BGD, which is an area of overlap of the two congruent triangles. Therefore the area of the two small squares is equal to the sum of blue and yellow areas in the figure.

Finally, triangle DEF is another triangle congruent to ABC. (Why? DF must be parallel to AC because it connects two points at a distance c from AC, and so angle EDF and $A'DA$ add up to $90°$, giving a triangle with corresponding angles equal and hypotenuse length c.) Therefore (translating triangle DEF down to triangle ABC) the blue and yellow areas add up to the area of square $ADFC$, which is congruent to square $ACC'''A'''$, as desired.

I wrote all this, and was about to provide an aesthetic analysis, when I realized that the proof I just gave you isn't quite the one intended in Figure 1.8 (which I took from the book [Nelsen, 1993] *Proofs Without Words*). I had completely ignored the second image in the sequence of 5. That was because I couldn't make sense of its role in the proof—it seemed to represent cutting off and translating arbitrary triangles from the squares of sides a and b. Why not just immediately translate the correct triangles—$AA'D$ and $CC'F$ in Figure 1.9?

It was only when I turned back to Figure 1.8 after writing out my proof that I realized that the significance of the second image is that the greyed out areas are *parallelograms* (and not just triangles affixed to quadrilaterals). The grey areas in the first image are equal to the grey areas in the second because the area of a parallelogram is its height times its base—i.e., in this case, $a \cdot a$ and $b \cdot b$. And now we can note that the greyed out areas in the third image are also parallelograms, thus preserving the areas.

It's true that the second image is still, strictly speaking, superfluous. But it's a useful cognitive aid—a hint that in the third image we should be seeing the grey areas as two parallelograms. Plus, images 1-3 together suggest a continuous transformation of the two small squares into a series of parallelograms of equal areas.

I thus got the structure of the proof in Figure 1.8 (a little) wrong. It doesn't work by *cutting up* the small squares and rearranging the pieces. Rather, the small squares are transformed in several area-preserving ways: continuously transformed into parallelograms (images 1-3), translated by a vector of length c (image 4) and *then* cut up and rearranged (image 5).

Once I understood all this, I found myself in a strange affective state. On the one hand, I was disappointed. I thought the work I had put into writing out a mistaken explanation of the proof had been wasted, and I kicked myself for being so slow to follow what now looked like a perfectly clear blueprint. On the other hand, I felt *pleased*. I suddenly found myself replaying the various transformations in Figure 1.8 in my mind, and I somehow felt grateful that this proof—which had so inconveniently tripped me up—existed.

And then I knew—the proof was beautiful. It caused me to feel pleasure even though its existence was (or at least seemed to be) an inconvenience. As we'll see in more detail in the final section of this essay, such disinterested or non-egotistical pleasure—pleasure at an object independent from anything we might gain from it—is a mark of beauty and the aesthetic.

I find the proof in Figure 1.8 beautiful, then. And note that I found it to be so precisely when I understood it, when it became a full blueprint for the proof. In my imagination, I can now seamlessly run through the transformations it presents, pausing if I wish at various points and assuring myself that I know why these transformations are area-preserving. A visual proof, then, is only fully beautiful for someone who learns to see it as a blueprint. Until I understood the role of image #2, the figure hadn't created an *idea* in my mind.

We now have not one, but two (very closely related) new proofs of PT—a "parallelogram" one and a "triangle" one. How do they compare? Well, note that we still need some of the triangle congruences from the triangle version of the proof to complete the parallelogram version—we need to show that (using the labels from Figure 1.9) EB is parallel to FC, and then that triangle DEF is congruent to triangle ABC. So the proofs aren't that different.

Still, my experience of the proofs differs. In the first proof, I enjoy the fact that wherever I turn, there's a triangle congruent to the one we started with. This is a type of mental unification—the various steps of the proof all involve congruent triangles. It's also a form of so-called purity of method: a fact about right-angled triangles is shown while manipulating pretty much only right-angled triangles.[22] By contrast, the second proof uses the formula for the area of parallelograms, which may seem to have nothing to do with the Pythagorean Theorem. In this sense, the experience of the first proof is

[22]For the notion of "purity of method," see [Detlefsen and Arana, 2011].

more unified than that of the second.

On the other hand, because the second proof employs continuous transformations where the first proof cuts figures up into pieces, its easier to hold all the parts of the proof in one's mind at once. Once I understood it, Figure 1.8 really became a blueprint for me. Though it's made up of several images, it can take on the shape of a single perceptual item, akin to a short video. By contrast, Figure 1.9 is more of a mnemonic than a blueprint. It allows me to remember that the proof involves extending the edges of the square of side c and cutting up figures into triangles, which is enough for me to recreate the proof. But I can't quite hold all the moves in my head at once.

Where does that leave us? It points to the richness of aesthetic experience in mathematics. Two versions of essentially the same proof can lead to different flavor of mental unification and different types of blueprint. I have aesthetic experiences centered around both proofs—but they are different experiences.

Which of the proofs is more beautiful? There's room for disagreement. Someone who highly values purity of method, unity of means to end, may prefer the first version. Someone who wants to be able to hold all the pieces of a proof in their mind at once may prefer the second. But we shouldn't assume the question is even answerable. Which one is more beautiful—Monet's "Haystack in Sunlight" (Figure 1.11) or "Haystack (Sunset)" (Figure 1.12)? As far as I'm concerned, both variations on a theme are beautiful in their own unique ways. Beauty isn't a yardstick allowing us to compare any two objects. Rather, it's a gesture towards an experience. Once we clarify the experience (as I have tried to do in this section), there is nothing more to

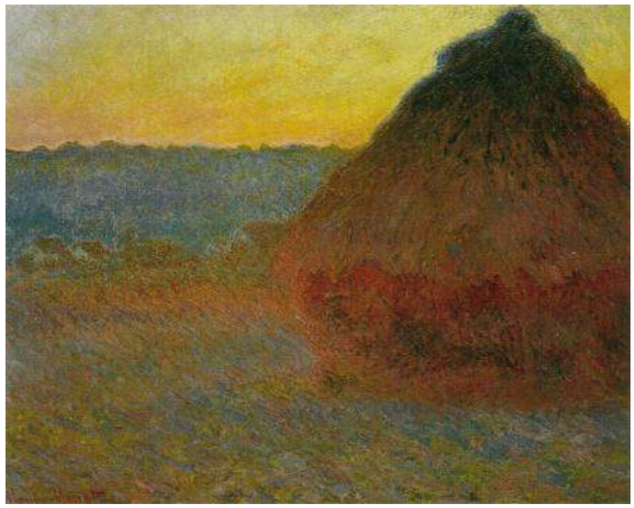


Figure 1.11: Claude Monet, *Haystack in Sunlight*, 1891.

say.

Nonetheless, sometimes a comparison of beauty is possible and fruitful. I find the proof from the previous sections (the one with the triangles sliding into corners) more beautiful than the one(s) from this section. Why? Well, given the first proof, the second seems needlessly complicated. Indeed, proof #2 can be seen as a failed attempt to come up with proof #1. Note that in Figure 1.9, triangles $AA'D$ and $CC'F$ are slid into corner E, forming triangles $EB'B$ and $EB''B$. And in the last step of the proof, triangle DEF is slid into triangle ABC. So along the way to proof #2, we have carried out basically all the steps of proof #1. Noticing this decreases my appreciation for proof #2, because it shows that in this case a higher degree of mental unification is possible.

In the next section, I'll present a proof which has a similarly strong "aha!" moment to proof #1, but where the character of this moment is

Figure 1.12: Claude Monet, *Haystack (Sunset)*, 1890-91.

entirely different.

4.3 Proof #3

For all their beauty, the above proofs are pieces of recreational mathematics, ultimately "clever insights of limited generality." Thankfully (for the fan of mathematical beauty), there are more substantive examples. Such is the case of the following proof.

Here's a fact about areas: they are proportional to the square of lengths. In other words, given two similar figures in the Euclidean plane, if the ratio of any two corresponding line segments is $a : b$, then the ratio of the figures' areas is $a^2 : b^2$. For instance, if the ratio of camel leg heights in Figure 1.13 is $a : b$, the ratio of camel areas is $a^2 : b^2$. This means that the Pythagorean Theorem holds not just for squares, but for *any* two-dimensional shapes. That is, if you base *any* three similar figures (for instance, camels) on the sides of a right-angled triangle, the areas of the two smaller figures will add up to the area of the larger one. In fact, for any such three figures, the Pythagorean Theorem is *equivalent* to the claim that their areas stand in this relationship. (Since $a^2 + b^2 = c^2$ is equivalent to $ka^2 + kb^2 = kc^2$ for any non-zero k.) This means that to prove the Pythagorean Theorem we just need to prove the corresponding claim about areas for *some* three figures.

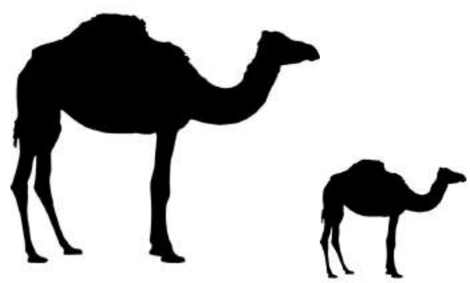


Figure 1.13: A fact about areas

Figure 1.14 presents a fact about right-angled triangles. The altitude AH

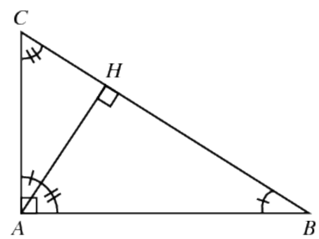


Figure 1.14: A fact about triangles

of a right-angled triangle ABC divides it into two smaller triangles (ACH and BAH) similar to triangle ABC. Why? Because each of the smaller triangles shares one angle with the big triangle and has a 90° angle for its second angle, so that the third angle must also be equal to the corresponding angle in the big triangle.

Clearly, then, the area of triangle ABC is equal to the sum of the areas of triangles AHC and BHC. But now if triangle ABC has side lengths a, b, c (as in the Pythagorean Theorem), this means that we have three similar right-angled triangles of hypotenuse a, b, and c, respectively, such that the areas of the first two triangles add up to the area of the third. In other words (given the fact about areas), Figure 1.14 is a proof of the Pythagorean Theorem!

As far as proofs of this theorem go, I think this one takes the cake. What makes it so beautiful? Above, I suggested that proof #1 is so beautiful in part because because it proves PT by embedding squares of sides a and b in a larger area, rather than cutting them up into smaller pieces. This makes it easier to see the truth of PT "at a glance."

Proof #3 does something similar, but in a less literal sense. Figure 1.14 is also embedded in something that allows it to show us at a glance that PT holds. But that "something," rather than a geometric figure, is *background knowledge*: knowledge of the fact about areas.

Visual proofs do literally what others do metaphorically. They allow us to literally see at a glance why a theorem holds. They show us how the pieces literally fit together.[23] Proof #3, a borderline case of a visual proof, dazzles

[23]This is why I'm sceptical of the tendency in contemporary philosophy of science to

us by working on both a literal and a metaphorical level: we see how the two small triangles fit together to create the big one, and by embedding this image in background knowledge, we (metaphorically) see that this implies the truth of PT.

For someone who has internalized the fact about areas, Figure 1.14 *is* the whole proof. And it has reached the limits of simplicity—as a user of the Mathematics Stack Exchange Forum points out, it's a "one-line proof," where the "one line" is line AH.[24] In a small way, this proof is emblematic of a key trend in pure mathematics: the continued search for frameworks which allow "one-line" proofs of familiar facts, by unifying them under more general theorems (such as, in this case, the extended Pythagorean Theorem about *all* triples of similar figures of ratio $a : b : c$).

How does this proof fit into Poincaré's picture? Well, just like our first proof, it allows you to see at a glance why the theorem holds. It's a compressed version of the whole proof which can serve as an "aha" idea. But that's not all. By drawing attention to the fact that the Pythagorean Theorem holds not just for squares, but for all triples of similar figures, the proof unifies a host of heretofore unrelated facts into one theorem. Appreciating this fact is a second act of mental unification which adds to the depth of the experience of this proof.

Dawson asks: aren't beautiful proofs just clever insights of limited generality? For Poincaré, the answer is a resounding "no." In fact, generality

assume that explanations take the form of propositions and to exclude the possibility of visual explanations. The way I see it, an explanation allows us to see why something is true. The literal case of seeing should be the least controversial, not the most.

[24] [sta, 2010] "What is the most elegant proof of the Pythagorean theorem?"

correlates with one type of beauty: the more general a proof or theorem, the more capacity for mental unification, and hence aesthetic experience. The aesthetic appreciation of a sophisticated piece of mathematics has a rich structure: the mental unification of the various parts of an individual proof can coexist with the sort of unification which allows us to see the place of the proof in a broader theoretical framework. Furthermore, while our first two proofs indeed have "limited generality," aesthetic *experiences* centered around them and similar proofs are nonetheless valuable. The better able you are to hold a proof in your mind as a single entity, the better you can use it, for instance by coming up with similar proofs in other domains. Even elementary beauty makes a real difference.

Proof #3 is a borderline case of a visual proof. While there is some temptation to think of the beauty of proofs 1 and 2 as something which arises out of their purely visual properties, the fact about areas plays an central role in the aesthetic experience of proof #3. Without it, Figure 1.14 is at most a demonstration of an unremarkable fact about triangles. We have therefore reached the borders of purely visual aesthetic value in mathematics. Is there anything past the border? Can there be such a thing as non-visual aesthetic value? This is the question I try to answer in the next section.

5 Aesthetics and the Senses

5.1 Aesthetics and Common Sense

Etymologically, "aesthetic" means "related to the senses." Definitionally, it means "related to art and beauty." The following principle is often thought to link definition and etymology.

> [Aesthetic means perceptual] Only things perceived through the senses can be beautiful or have other aesthetic qualities.

If this claim is true, only the most visual of proofs can count as beautiful. In this section, I will argue that "aesthetic" does *not* mean "perceptual."

I will do this by, first, providing some examples of things many of us would call beautiful which we don't seem to perceive through the senses. I will thus begin locating possible neighbors for mathematics in the realm of aesthetics, in the hopes of finding a home for it there. Second, I will sketch a brief history of Aesthetic Means Perceptual, arguing that common commitment to this thesis is merely a historical accident.

Kant opens his *Critique of Judgment* with a section entitled "The Judgment of Taste is Aesthetical," where "aesthetical" has the etymological meaning. Kant's notion of beauty was peculiarly narrow, a matter of formal visual properties, of the arrangement of shapes (and not even colors).

Hardly anyone is a Kantian formalist anymore, but the conviction that aesthetic means perceptual persists today, even among thinkers who are wholeheartedly committed to re-expanding the notion of the aesthetic beyond Kantian strictures. Thus in his article "Anything Viewed," Paul Ziff urges that an alligator basking in the sun on a mud bank, or even a mound of dried dung, are appropriate objects of aesthetic appreciation. Anything viewed can be aesthetically appreciated.[25] Anything viewed—but not anything conceptualized. Ziff takes "aesthetic means perceptual" as his unexamined starting point, expanding the notion of the aesthetic only in arguing that "perceptual means (potentially) aesthetic."

Yet much of our ordinary discourse belies the assumption that aesthetic means perceptual. At least in our unguarded, unphilosophical moments, we speak of beautiful actions, beautiful souls or characters, beautiful ideas, metaphors, stories, mechanisms, philosophical arguments. At least prima facie, appreciating any of these requires faculties that go beyond our senses—and appreciating some doesn't involve the senses at all.

Traditional aesthetics, which proclaims that beauty is perceptual, also concerns itself with the literary arts. In fact, Baumgarten, who is responsible for introducing the term "aesthetics," was concerned primarily with drama and literature, and was happy to speak of "beautiful thought."[26] But this domain provides particularly striking examples of – at least seemingly—non-perceptual aesthetic qualities. While the sounds and rhythms of poetry are perceptual beauties, literature's beauty outstrips that of its sounds. "Breathing of statues," Rilke's metaphorical description of music (which evokes the way music literally affects our breathing, the way it expresses emotions where one might expect to hear a mere sequence of sounds, and much more) is beautiful whether it's expressed in German or in English,

[25] [Ziff, 1979] "Anything Viewed."
[26] [Baumgarten, 2007], as cited in [Shelley, 2017].

whether it's in Stephen Mitchell's much more rhythmical and sensually beautiful translation of "To Music," which has "breathing of statues," or in Scott Horton's inferior rendering, with "the breathing of statues." The metaphor itself, not just its sound, is beautiful.

While some aestheticians (like Nick Zangwill) insist that, strictly speaking, only the perceptual properties of a literary text are beautiful,[27] and others (like Elaine Scarry [Scarry, 2001]) suggest that literary beauty is ultimately a matter of the beauty of the visual imaginings prompted by literary works, the view that literature abounds in non-perceptual beauties has common sense on its side.

The friend of *Aesthetic Means Perceptual* will likely respond that in such unguarded moments we are speaking loosely, perhaps metaphorically. This view has a credibility that extends beyond mere appeal to tradition. We do sometimes use the word "beautiful" loosely, as a generic form of approval. To borrow an example from dictionary.com, when the chef offers you a beautiful cut of steak, she is unlikely to be making a genuinely aesthetic judgment.[28] She might equally well have called the steak "nice," "good," "wonderful," "a fine specimen."

The sceptic about non-perceptual beauty, then, will claim that when I say that "breathing of statues" is a beautiful metaphor for music, I mean nothing more than that it's nice, good, wonderful, or a fine specimen of metaphor-kind. But it at least *seems* to me that I mean something more.

[27]See [Zangwill, 2001], *The Metaphysics of Beauty*, ch. 8. Zangwill argues that the property of beauty necessarily depends on perceptual properties.

[28]I don't mean to be denying that there can be aesthetic experiences of food. I'm just claiming that our chef hasn't committed herself to their existence.

We may have reason to be wholesale sceptics about beauty. But I see no special reason to be a sceptic about perceptual beauty without being a sceptic about beauty simpliciter. Why can't I mean what I think I mean when praising the beauty of Rilke's metaphor, when I can mean exactly what I think I mean when praising the sound of his poems?

Aesthetic Means Perceptual might be held up as a definitional truth. I don't want to engage in a terminological debate. Merriam-Webster's online dictionary places intellectual and sensible beauty in a single sense.[29] Oxford Dictionaries.com has two subsenses for intellectual and perceptual beauty.[30] It's just a matter of how you slice things up, not a deep question. Maybe the philosophical community by and large sides with Oxford Dictionaries; maybe its concept of beauty really entails perception. But even if that's the case, there is a related, broader, concept which entails no such thing, one which accords well with non-philosophers' discourse. Call it quasi-beauty if you prefer; I'll use the convenient term "beauty." Furthermore, this is a concept that was common currency during British empiricism, and which philosophers have abandoned through misreadings more often than through argument. Or, at any rate, that is the reconstructed history I will sketch in the next section.

[29] "Beauty : the quality or aggregate of qualities in a person or thing that gives pleasure to the senses or pleasurably exalts the mind or spirit."

[30] "1. A combination of qualities, such as shape, colour, or form, that pleases the aesthetic senses, especially the sight. 1.1 A combination of qualities that pleases the intellect."

5.2 Historical Interlude

"From the outset aesthetics has been devoted to the study of 'things perceived'," claims Timothy Binkley in his [Binkley, 1977] "Piece: Contra Aesthetics." Aesthetic Means Perceptual, then, seems to have tradition on its side, even if it departs from ordinary usage.

Binkley continues.

> The commitment to perceptual experience was deepened with the invention of the Faculty of Taste by eighteenth-century philosophers anxious to account for the human response to Beauty and to other aesthetic qualities.

In fact, the tradition Binkley is alluding to is much more recent and less venerable than he makes out. If we take just a step back in time beyond the British empiricists, we'll see a world in which literature and even mathematics are taken to be paradigm examples of beauty, and the presence of beauty itself is thought to be inferred on the basis of principles as precise as those of mathematics. British empiricist aesthetics, I'll argue, is only properly understood against this rationalist backdrop.[31] Eighteenth-century British aestheticians were responsible for shifting the aesthetic paradigm from the non-perceptual to the perceptual, and for denying the claim that the presence of beauty is inferred. But for all that, they were happy to acknowledge the existence of intellectual beauty.

[31] My argument draws on [Shelley, 2017].

5.2.1 Beauty of Principles Without Principles of Beauty

The primary rationalist doctrine opposed by the British empiricists was the belief that judgments about beauty were based on reason, i.e. inference or application of concepts.[32] It was thought that there could be principles about beauty as precise as those of geometry, and that we learned, or at least could learn, about things' beauty by inference from such principles. That is, I could deductively prove that a certain crocus, say, was beautiful, starting from premises which listed some of the crocus's non-aesthetic properties (such us its height and color or its genus) together with general aesthetic principles.

Against this earlier view, the British empiricists posited the existence of a "faculty of taste" fitted for "perceiving" beauty. Typically, this "perception" was thought to amount to an experience of pleasure or displeasure in response to certain objects (or, in the parlance of the day, certain combinations of simple ideas). The use of the term "taste" was selected to highlight similarities between this faculty and literal, bodily taste. One cannot reason one's way into a verdict about the deliciousness of a dish. Instead, one tastes the dish and feels delight, indifference, or disgust. The British empiricists thought that aesthetic taste was more similar to this than to judgments made based on inference.

British empiricism was the beginning of the assumption that beauty "went with" the senses. But while this assumption eventually came to be synonymous with Aesthetic Means Perceptual, initially it only meant that

[32] Cf. [Shelley, 2017].

the faculty of taste was sense-like, rather than intellect-like.

The empiricist thesis that the faculty of taste is sense-like can be divided into three main parts.

1. *No Need for Inference.* At least in the typical case, we don't infer the beauty of objects from principles.

2. *Acquaintance.* To know that something is beautiful, one must come into contact with it.

3. *Pleasure.* To find something beautiful, one must feel pleasure.

No Need for Inference is simply the negation of rationalist aesthetics: at least in paradigm cases, we do not come to find things beautiful based on a series of inferences akin to those made in mathematics. The minimal British empiricist view is consistent with the *existence* of necessary and sufficient conditions on beauty (or on my finding a given item beautiful), so long as it remains the case that we typically do not—and aren't rationally required to—make aesthetic judgments on the basis of such conditions.

Acquaintance is self-explanatory; *Pleasure* provides an explanation for why it might hold: we have to come into contact with the beautiful item precisely so that we can be pleased by it. Pleasure plays an ineliminable role in our aesthetic judgments and experiences. Furthermore, on most British empiricist views, it is a special *kind* of pleasure that signals aesthetic appreciation: a non-egotistical pleasure caused by the operation of the faculty of taste.

The existence of mathematical beauty is consistent with all three empiricist assumptions. *No Need for Inference* is consistent with mathematical beauty because to claim that there are principles of beauty is to claim that aesthetics is itself a mathematical science, that there is a list of axioms from which one can deduce claims of the form "A is beautiful," or perhaps of the form "if x has features P, Q, R, x is beautiful." But whether or not there are such principles—and whether or not we apply such principles in practice when we judge something to be beautiful—claims of the form "A is beautiful," where A is a piece of mathematics, such as a principle, might still be true. If there are no principles of beauty, no such A will make claims about beauty—but since mathematics as currently practiced in fact treats of sets, groups, functions, and the like, rather than the property of being beautiful, this doesn't threaten the beauty of *any* mathematical principle.

Compare other domains which contain beautiful items. There is no temptation to say that if there is to be literary or musical beauty, aesthetics must itself be a branch of literature or music. While the inference might be more tempting in the case of mathematics, it's no less fallacious there.

Acquaintance may appear to pose a threat to mathematical beauty. For, as Paul Benacerraf has forcefully argued [Benacerraf, 1973], we are not acquainted with mathematical objects in anything like the way in which we are acquainted with physical objects. There is an access problem for mathematics.

In response, let me first note that the type of mathematical beauty I am primarily concerned with here is *not* the beauty of mathematical objects such as sets or circles. Rather, I am interested in the beauty of proofs, theories, definitions. And we are acquainted with *these* objects in much the same way in which we are acquainted with novels or musical compositions. Just like you can't find *Anna Karenina* beautiful without having read it, you can't find Galois theory beautiful without having read a treatise or attended a lecture course on it—or invented the theory yourself. While it's true that the identity conditions for mathematical theories are less demanding than those for novels—that is, many different presentations correspond to the same theory—to a lesser degree multiple modes of access to novels and musical compositions are also available. For instance, I can find *Anna Karenina* beautiful whether I read it in the original Russian or in any English translation, whether I read it myself or listen to an audiobook.[33] The important point is that in mathematics as well as in literature, there is a clear distinction between access based on testimony and a more direct acquaintance. If someone merely gives me a list of some of the astonishing theorems of Galois Theory to me, but I have no understanding of how the theorems relate via proofs and only a faint understanding of the meaning of the theorems themselves, I can't know that the theory is beautiful.

It may be objected that in this case I *do* know that theory is beautiful, I merely don't *find* it beautiful—and, perhaps, wouldn't *say* that it's beautiful, because that would conversationally imply that I find it beautiful. That may be right—but there is no special disanalogy between mathematics and other beautiful things here. If *Acquaintance* is false (that is, only true for *finding* something beautiful, and not for *knowing* that it is), it's false in

[33] In my essay on the Quadratic Reciprocity Theorem, I develop some of these themes further, discussing the analogy between mathematics and musics.

other domains than mathematics too.

Furthermore, as I argued in Section 3, what we *care* about is *finding* things beautiful, not (if it even exists) the platonic property of beauty. What we value about beauty is its experiential aspect. And here Poincaré's view is in precise agreement with the British empiricists: to find a piece of mathematics beautiful is to have a certain experience centered around it, an experience you can't have without becoming acquainted with the piece of mathematics.

Finally, we have already seen that on Poincaré's view to find something beautiful is to take *pleasure* in one's mental unification. Therefore *Pleasure* holds for at least one type of mathematical beauty, and certainly *can* hold for other types too.

Two more comments on beauty and pleasure are in order. The term "beauty" has a narrow and a broad sense. In the narrow sense, beauty is just one aesthetic property among many. A rose is beautiful but the *Guernica* is not (at least, for most people). Massenet's idyllic *Meditation from Thais* is beautiful but Shostakovich's harrowing *Piano Trio No. 2* isn't.[34] Anything too dark or too heavy is disqualified from counting as beautiful. It's possible to be aesthetically good without being beautiful in this sense. In the broad sense, by contrast, beauty is the most general aesthetic value, simply equivalent to (all-things considered) aesthetic goodness.

If mathematical beauty exists, it isn't particularly dark. It's beauty even in the narrow sense.

A similar thing may be said about pleasure. In the narrow sense, Aubrey

[34]For a view like this, see Levinson's [Levinson, 2012].

Levinthal's profoundly depressing painting, *Breakfast at 13th Street* (2018), doesn't give me pleasure. In the broad sense, pleasure is the most general (intrinsic) value an experience can provide. I find the experience of looking at Levinthal's painting intrinsically valuable. I seek it out. I can't look away. Therefore, it gives me pleasure.

As with beauty, so with pleasure: the experiences of mathematics are pleasurable even in the narrow sense. Therefore the distinction between the two types of beauty makes little difference to the question of mathematical beauty—but sharing my views on it will make my framework clearer and provide some support for a broadly empiricist view.

I tend to favor the broad sense of "beautiful" as well as "pleasurable," and I think that you can't find something beautiful without being pleased by it. However, it's important to my concept of beauty that the narrow sense is the central one. Shostakovich's music is beautiful, but in a way that gives me pause, makes me hesitate and want to hedge. It pleases me—but in a puzzling way, almost despite myself. Furthermore, while my concepts of beauty and pleasure are broad, they aren't generic. I am not pleased by everything I value—only by all experiences I value for their own sake. Pleasure is necessarily experiential; there's something it feels like to be pleased even by Shostakovich. If I merely dispassionately approve of an artwork, that isn't pleasure and it isn't aesthetic experience.

5.2.2 The Empiricist Paradigm Shift

The British empiricist claim that we do not infer the presence of beauty from principles, but rather that pleasure plays an ineliminable role in our judgment that a given item is beautiful, then, is perfectly consistent with the existence of mathematical beauty. However, there may be another way in which the British empiricist "faculty of taste" was a precursor to Aesthetic Means Perceptual. This is a way which may be conceptualized as a form of paradigm shift.

Some uses of the "faculty of taste" are much more similar to literal taste than others. One's pleasure at a natural beauty such as a crocus is very much like one's pleasure at a mango smoothie. Typically, both pleasures occur as soon as one comes into (perceptual) contact with the item—one sees the flower and tastes the smoothie and feels the pleasure at the same time, or very close to the same time. Indeed, we call the seeing and the tasting themselves pleasant. There is little temptation to think that your judgment about the crocus is inferred from principles. The arts are further from bodily taste than nature, and literature is furthest of all. Here, there can be a temporal gap between our coming into contact with a beautiful item and aesthetic pleasure and judgment. (And what "coming into contact" amounts to is a less straightforward matter than in simple perceptual cases.) We also often give *reasons* for our judgments, a feature which may be thought to imply that we do, after all, base our judgments in such cases on inferences.

The eighteenth century's guiding metaphor, then, inverted the ordering of beauties relative to the rationalist ordering. Natural beauties—hardest to account for on the rationalist picture, on which literature was the paradigm—found their way to the center of the realm of the beautiful, while literary beauties were relegated to the fringes. But for all that, the British empiricists never once thought that these "fringe" cases weren't genuine examples of beauty. Rather, they were test cases for their views, challenges to be faced. Thus, for instance, Hume's "Of the Standard of Taste" (in [Hume, 1757]) was (among other things) an attempt to make sense of reason-giving in the arts, especially literature and drama, under the "taste" paradigm.

Mathematics belongs with literature, on the fringes of the taste paradigm. We don't come into direct contact with mathematical entities the way we do with crocuses, and we come to find proofs or theorems beautiful only after prolonged thought. But for all that, the crucial empiricist insight remains plausible: we don't infer the presence of mathematical beauty from principles, and first-hand experience plays an ineliminable role in our judgments.

We have inherited aspects of the "taste" paradigm from the empiricists. But that is no reason to deny the existence of mathematical beauty. Instead, it's a reason to investigate further. As in science, so in philosophy: the edge of a paradigm is an exciting place.[35]

5.3 Gaining Knowledge Without Egotism

The assumption that the faculty of taste is sense-like, then, doesn't threaten the existence of non-perceptual beauty. But there is one more British empiricist assumption which may be thought to do so. This is the claim that aesthetic pleasure is non-egotistical. In other words, when we are pleased by a beautiful item, our pleasure is independent of any advantage we might expect to gain from the item. We are pleased by the thing itself; we approach

[35]In section 2.6.4 of the final essay I return the question of how mathematical beauty fits into the empiricist paradigm, drawing on the contemporary notion of an "attentional system."

it on its own terms. The beautiful item may be an inconvenience—the work of a competitor artist, for instance—but that fact is irrelevant to our pleasure insofar as the pleasure is aesthetic. In this way, the pleasure we feel at beholding a snowcapped mountain—a pleasure which might arise even if the mountain is a sign that we are lost or that our journey will be more arduous than we had hoped—differs from the pleasure of beholding a check inscribed with our name and a large sum.

Like a check, mathematical beauty can always be cashed—in the currency of epistemic goods such as knowledge. On one view (which may have been Kant's), this is enough to make the pleasure egotistical. The reasoning goes roughly as follows.

1. So-called aesthetic pleasure at mathematics always co-occurs with the acquisition of epistemic goods (such as knowledge or understanding).
2. Therefore, such pleasure is always pleasure *at* gaining epistemic goods.
3. Therefore, such pleasure is always egotistical.

I will respond to the argument in two steps, dealing first with the version of the argument which speaks of knowledge and then with a stronger version which relies on the notion of understanding.

5.3.1 The Knowledge Version of the Argument

Each step in the argument can be questioned. (1) is undermined by the fact that we can experience aesthetic pleasure from pieces of mathematics we are already familiar with, without gaining any new knowledge—but perhaps in those cases our pleasure is like that of the millionaire taking out her cash to count it. The move from (1) to (2) needs further argument: so far, we have a correlation of pleasure and the gain of knowledge, with little reason to think that the pleasure is so much as caused by, let alone directed at, the epistemic gain. Now *something* can be said to support this inference. An art collector whose pleasure at the paintings he owns is highly correlated with their market price may *think* his pleasure is genuinely aesthetic, but we have good reason to suspect otherwise. The contention here is that pleasure at mathematics, and intellectual beauty more broadly, is equally compromised by its correlation with epistemic value.

But there are important differences between the art collector and the collector of beautiful mathematics. Epistemic and material goods are importantly different, in a way which undermines (3) as well as (2). The move from (2) to (3) presupposes that pleasure at gaining any good at all must be egotistical. But while the pursuit of some *practical* knowledge may be egotistical, there's little reason to think that the pursuit of *theoretical* knowledge is.

To appreciate the weakness of our argument, consider a more paradigm case of aesthetic appreciation than that of mathematics. When I am struck by the beauty of a newly found painting, there are many goods I expect to receive from it. I anticipate a valuable experience, a pleasant fifteen minutes. I expect to gain knowledge too: at the very least, knowledge of what the painting looks like, but sometimes also knowledge of historical events, a deeper understanding of human psychology, etc.[36]

[36]As I discuss in the final essay, such expectations can and do get in the way of aes-

However strong we want to make the disinterestedness requirement, it's absurd to make it so strong that even the gain of knowledge of what a painting looks like is forbidden. And the knowledge that a theorem in pure mathematics brings is at most a tiny step away from the knowledge of what the mathematics is like—knowledge which should be allowed to be part of aesthetic appreciation on any reasonable account. Indeed, it's hard to imagine how one could possibly come to fully know what a mathematical proof is like without at the same time coming to know the theorem it demonstrates. The knowledge mathematics brings isn't detachable from the mathematics in the way in which monetary value is detachable from a painting. To claim that this is a problem for the disinterestedness of my pleasure in a beautiful piece of mathematics is to make a mistake akin to claiming that I can't like my wise friends for their own sake because wisdom is a desirable characteristic to have in a friend. To appreciate my wise friend for her own sake *is*, in part, to appreciate her wisdom.

We're not tempted to think that the gain of knowledge of what a painting looks like brings the threat of interest to our appreciation in part because often we wouldn't *want* to know what the painting looks like if it weren't beautiful. I'd like to suggest that some cases of mathematical beauty are like this too: we wouldn't bother to attain the knowledge a piece of pure mathematics brings if the mathematics weren't beautiful.

The argument that the knowledge we receive from mathematics makes our pleasure egotistical is no more plausible than the following cousin argu-

thetic experience if we cling to them too tightly—but that isn't a special problem for mathematics.

ment: the old idea that even your seemingly selfless acts are in fact selfish, since the mere fact of *desiring* something means that you find it to be good *for you*. Both arguments slide from an intuitive notion of egotism to one based entirely on superficial grammatical features like the use of the words "want" or "gain."

5.3.2 The Understanding Version of the Argument

The argument from understanding is more threatening. After all, as we've seen above, the very role one type of beauty (or "beauty") plays in mathematical practice is enabling understanding. Doesn't this mean that mathematicians value this type of aesthetic experience for the understanding it provides, rather than for its own sake?

Three points in response. First, fallacious proofs can be beautiful.[37] Such proofs fail to provide understanding while still providing a pleasurable experience of mental unification. And while the pleasure at a fallacious proof is *diminished* when one spots the flaw, it doesn't entirely disappear. So there seems to be at least a kernel of aesthetic pleasure in such cases. Second, our pleasure at the beauty of artifacts can be similarly diminished when we find out that they don't work, but this doesn't necessarily mean that the pleasure wasn't aesthetic. There is such a thing as functional beauty. It's just that a certain type of experience, an experience of perceiving the item as a well-functioning mechanism, is no longer possible when we find that the item doesn't work.[38] Third, the type of understanding which beautiful

[37] Here, I am in disagreement with [Zangwill, 2001] Nick Zangwill, who claims that "there cannot be proofs... which are dysfunctional yet beautiful or elegant."

[38] I discuss such a case in more detail in the essay on Quadratic Reciprocity.

proofs provide is thoroughly experiential, and valuing such understanding is a form of valuing the proof itself. To say that we value the proof for the sake of the understanding is therefore as misguided as saying that we value an artwork for the sake of understanding it.

Together, these points make it plausible that pleasure at mathematical beauty can play a substantial role in mathematical practice without being egotistic.

6 Conclusion

If there is no mathematics without beauty, how come so many philosophers have thought otherwise? The quote from Marc Lange with which we began this essay sheds some light on this question.

> One reason that mathematical explanation has received relatively scant philosophical attention may be the temptation among philosophers to believe that when mathematicians apparently characterize some proof as explanatory, they are merely gesturing toward an aesthetically attractive quality that the proof possesses (such as elegance or beauty)—a quality that seems to be very much in the eye of the beholder.[39]

Notice the assumption implicit here: if beauty is "in the eye of the beholder," it *shouldn't* deserve philosophical attention.[40] Subjective experience is the land where no philosopher dare go.

[39] [Lange, 2016] Marc Lange, *Because Without Cause.*

[40] There are two senses in which beauty might be said to be "in the eye of the beholder:" (1) there is no *Book of Beauty*, no fact of the matter about which pieces of mathematics are and aren't beautiful independently of the observer, (2) what goes on inside the minds of appreciators matters more than any putatively objective facts about beauty. In this essay, I have argued only for (2).

If you're not allowed to go somewhere, you might forget the place even exists. This may be the reason philosophers like Dawson have tended to underestimate the importance of mathematical beauty: they have become blind to the things they think they can't study.

This essay has been an excursion into the forbidden land of subjective experience. If what I have argued here is correct, this land is *both* amenable to philosophical inquiry *and* extremely important. I hope I have begun uncovering some of the riches to be found there.

When it comes to mathematical beauty, the eye of the beholder is where it's at.

Chapter 2

The Allure of Elegance: Aesthetic Factors in the Development of Three Proofs of Quadratic Reciprocity

Carl Friedrich Gauss begins his 1808 article [Gauss, 1808] presenting his third proof of the Quadratic Reciprocity Theorem in number theory as follows.

> The questions of higher arithmetic often present a remarkable characteristic which seldom appears in more general analysis, and increases the allure of the former subject. (...) [I]n arithmetic the most elegant theorems frequently arise experimentally as the result of a more or less unexpected stroke of good fortune, while their proofs lie so deeply embedded in the darkness that

> they elude all attempts and defeat the sharpest inquiries.[1]

In Gauss's view, then, a core feature which attracts mathematicians like himself to number theory—a feature which makes number theory so *attractive* to them—is the existence of easily discoverable elegant theorems whose proofs are much "deeper" and harder to find. Indeed, it took Gauss a whole year to discover a proof of the Quadratic Reciprocity theorem, a year during which the theorem "tormented [him] and absorbed [his] greatest efforts." Given that the only adjective with which Gauss describes the theorem is "elegant," one may surmise that it is precisely the elegance of the theorem that inspired such devoted work.

Gauss then comes across three other proofs of the theorem, but he remains dissatisfied. The last three proofs "proceed from sources much too remote" (today, we might say that these proofs lacked the virtue of "purity of method"); the first "proceeds with laborious arguments and is overloaded with extended operations." Finally, he finds a proof which *might* qualify as "natural."[2] Gauss appears to be applying the following heuristic: elegant theorems ought to have natural proofs.

Gauss concludes:

> I leave it to the authorities to judge whether the following proof which I have recently been fortunate enough to discover deserves this description ["natural"].

[1]I follow the translation in [Smith, 1929], which has "beauty" where I write "allure." Gauss's Latin has "illecebras" ("lures" or "enticements.") Since my aim is to analyze mathematicians' use of aesthetic language, I don't want to put aesthetic-sounding words in their mouths, and so I have chosen the more neutral "allure."

[2]This is precisely the "third" proof published in [Smith, 1929]. It's the third *published* proof but the fifth discovered by Gauss.

The "authorities" who came after may have judged against him. In 1844, Gotthold Eisenstein published a proof[3] that might be called a cousin of Gauss's demonstration. The proof is widely considered to be superior.[4]

In this paper, I will use Gauss's and Eisenstein's proof, as well as a hybrid Gauss-Eisenstein proof in modern presentation, taken from [Goldman, 1997], to shed some light on the role of aesthetic notions such as beauty and elegance in mathematical practice. I'm interested in the two families of questions which we've seen in the first essay:

> [*Meaning*] What *is* aesthetic experience in mathematics? What does it mean to find a piece of mathematics beautiful? What does mathematical beauty have in common with the beauty of sunsets, symphonies, ceramics?
>
> [*Role*] What role(s) does beauty and aesthetic experience play in mathematical practice? (How would mathematics change if mathematicians had different aesthetic responses?)

A careful examination and comparison of the proofs will reveal a wealth of distinct aesthetic judgments: on the small and large scale, in more and less presentation-bound forms, and in ways that do and don't take the functions of proof (to prove and to provide understanding) into account. In particular, I'll argue that, depending on one's aesthetic preferences, it's equally justified to find Eisenstein's proof more beautiful than the modern one as it is to find it *less* beautiful. I will compare these judgments to ones we make in the realm of sport, functional design, and interior decoration to help answer the *Meaning* question and bolster the claim that these judgments are genuinely

[3] [Eisenstein, 1844]

[4] [Reinhard C. Laubenbacher, 1994]

aesthetic—and to show that such indeterminacy is to be expected in any aesthetic domain.

As to *Role*, I will argue that the pursuit of beauty motivated Gauss as well as Eisenstein, and led to the discovery of their proofs, thus making a difference to the history of mathematics. However, the effect of this pursuit of beauty wasn't entirely positive: I believe that Eisenstein was misled by his pursuit of beauty into the creation of a needlessly complicated proof. A core aim of the practice of mathematics is mathematical understanding: understanding, as deeply as possible, how mathematical concepts and truths are interrelated and why mathematical theorems hold. Beauty is a good guide in the pursuit of such understanding, but it is not an infallible guide. Nonetheless, Eisenstein's proof provides more understanding than Gauss's, and was an important step on the way to the hybrid proof. In this way, aesthetics played a genuine role in the development of number theory.

In Section 1, I'll provide a toy example that will ease us into the sort of aesthetic comparison we'll be carrying out in further sections. Sections 2 and 3 introduce the Quadratic Reciprocity Theorem and its three proofs. The last proof (Gauss's) may be hard to follow, so it's okay to skim that section—but I encourage even the less mathematically inclined reader to examine the outline of the proofs provided at the beginning of Section 3 and tackle at least the first proof. At the end of Section 3, you'll find a table which summarizes the proofs. I hope it will provide a useful cognitive aid to Section 4—the comparison and aesthetic analysis of the proofs. Section 5 compares the aesthetic qualities of the proofs to the those of functional items such as bikes, and Eisenstein's project to the borderline aesthetic project of cleaning one's room. This highlights the many similarity between the aesthetic features of proofs and other objects.

1 A Toy Example

In what follows, I will provide a detailed comparison of three pieces of reasoning. In one sense, these pieces of reasoning are all versions of the same proof. In another, they are three different proofs, with different aesthetic evaluations. Before we jump into the details of this mathematically and historically rich example, I'd like to introduce a simpler, toy case of the same sort: three proofs which can be seen, alternately, as importantly different and essentially the same. I will also develop an analogy to music which will help show that *both* ways of looking at the proofs (as same or as different) can be aesthetically relevant.

You probably know the following theorem.

Theorem 1 (Euclid)**.** *There are infinitely many primes.*

Imagine the following conversation between three geeky friends, Asia, Basia, and Kasia.[5]

> Asia: Don't you just love Euclid's proof of the infinity of the primes? It's so beautiful!
>
> Basia: I know, Euclid's proof was my first ever favorite proof!
>
> Kasia: Right?! The way you can just see in a flash that you've constructed a prime that couldn't have been on your original list!

[5]The following section is a version of a talk I presented at the New Ontologies of Art conference in Warsaw, Poland. In deference to the location, I have given my characters Polish names, pronounced "AH-sha," "BAH-sha," and "KAH-sha."

This is a commonplace sort of conversation: an item which the friends are all acquainted with in some fashion is judged beautiful. This is how mathematicians exchange aesthetic views. But a puzzle appears if we consider the sort of acquaintance the girls might have with the proof in question. If the friends aren't talking past each other, "Euclid's proof" ought to refer to the same thing in their respective utterances. But suppose Asia, Basia, and Kasia each went to different schools, where they learned, respectively, the following three proofs.

Asia's Proof. Let n be some natural number greater than 1. Then $N = n!+1$ has a prime factor p. But p can't be any of the numbers $\{1, ..., n\}$: otherwise p would be a divisor of N and of the product $n!$, and thus also of the difference $N - n! = 1$, which is impossible. So $p > n$. So for any n, there is a prime number $p > n$, hence there are infinitely many primes. □

Basia's Proof. Suppose there were only finitely many primes, $\{p_1, ..., p_r\}$. Then consider the number $n = p_1p_2...p_r + 1$. This n has a prime divisor p. But p is not one of the p_i: otherwise p would be a divisor of n and of the product $p_1p_2...p_r$, and thus also of the difference $n - p_1p_2...p_r = 1$, which is impossible. So we found a prime number p greater than any of the p_i's, contradicting our assumption. So there are infinitely many primes. □

Kasia's Proof. For any finite set $\{p_1, ..., p_r\}$ of primes, consider the number $n = p_1p_2...p_r + 1$. This n has a prime divisor p. But p is not one of the p_i: otherwise p would be a divisor of n and of the product $p_1p_2...p_r$, and thus also of the difference $n - p_1p_2...p_r = 1$, which is impossible. So any finite set

$\{p_1, ..., p_r\}$ of primes can be extended to a larger set of primes—hence there are infinitely many primes.[6] □

I'd like to take some time to compare and contrast these three proofs. (Though, as we'll see, in some sense they are all one and the same proof, I will continue to refer to them as "proofs" in the plural.)

First, the differences. To begin with extreme pedantry: the proofs aren't literally the same, since they consist of different strings of symbols. (This is pedantic, but for certain purposes, important. The mathematical discipline of proof theory requires such a coarse-grained notion of proof.) A slightly (but only slightly) less pedantic difference: the strategy (in one sense of the word) is different in each proof. Asia's proof gives a construction which multiplies together all numbers $1, ..., n$, while Basia's and Kasia's multiply together only prime numbers. And Basia's proof is a *reductio*, while Kasia's proof is a constructive proof. This gives us reason to think that all three proofs are distinct.

If you're writing a program searching for prime numbers, you'll want to stress a particular difference between Asia's proof and Basia's or Kasia's—Asia's proof gives an algorithm which "narrows down" the search space for prime numbers into hopelessly large chunks. (Basia's proof is, in turn, somewhat inferior to Kasia's, in that the latter provides an algorithm allowing one to construct a prime out of any set of primes, not just the set of all primes below a given one.)

A cursory glance reveals a similarity: it suggests that Basia's and Kasia's

[6]The final proof is adapted from [Aigner and Zieglier, 2010] *Proofs from the Book.*

proofs are mere presentational variants of each other. Basia's proof poses as a *reductio*, but comparison with Kasia's proof shows that this structure isn't doing any work; it's just a constructive proof in disguise.[7]

But this disguise obscures something else. Given our assumption that $\{p_1, ..., p_n\}$ included *all* the primes up to p_n, we were entitled to conclude that we'd found a prime number *greater* than p_n. So picking a special set of primes gives us additional information about our new prime number. On the other hand, Kasia's proof shows how to construct a new prime out of *any* finite set of primes, but we're not entitled to assume that the new prime will be larger than the primes we were given so far—e.g. the set $\{3, 5\}$ gives us $n = 3 \cdot 5 + 1 = 16$, with new prime factor 2. Kasia's proof is more general—we can start with any set of primes whatsoever, not just the set of the first n primes—and for this reason I find it more satisfying, perhaps even elegant.

So much for the differences. If you're trying to understand why there are infinitely many primes, or to commit enough of the proof to memory to be able to retrieve a related valid proof later, you're likelier to see the similarities. Proofs 1-3 give you extremely similar understandings, which come down to the "trick" of multiplying together the numbers in a set and adding one, thus building a large number forced to be relatively prime with the numbers you start with.

The upshot of this discussion is that the use to which we put proofs constrains what differences between them we see as significant and which

[7]Georg Kreisel's "unwinding" program in proof theory is concerned precisely with finding the constructive content of prima facie non-constructive proofs. For an overview of the program, see [Feferman, 1996].

ones we judge irrelevant. There's a *core idea* common to the three proofs which, for many purposes (such as gaining understanding or retrieving a related valid proof later), makes it natural to say that they are "essentially the same proof." But for various other purposes it makes sense to distinguish them. For computational purposes, Asia's proof is different from Basia's. (If Asia says "Euclid's proof is computationally inefficient" and Basia denies this, they are talking past each other.) For the purpose of gleaning a broader understanding of why there are infinitely many primes, they are essentially the same. (If Asia says, "Euclid's proof gives you a good understanding of why there are infinitely many primes," and Basia denies this, they may be genuinely disagreeing.)[8]

Identity conditions of proofs, then, are relative to purpose. What it takes to be an instance (or presentation) of Euclid's proof will vary from context to context. What if our purpose is aesthetic evaluation? If Asia says "Euclid's proof is beautiful," and Basia denies this, are the two girls talking past each other? The fact that exchanges like the one with which we began this section take place all the time between mathematicians who have seen variant presentations of "the" proof in question suggests that the requirements on being (essentially) the same proof for purposes of aesthetic evaluation are relatively undemanding. The three proofs above will get to count as "the same proof"—arguably because they have the same core idea.

And yet above I have made an aesthetic distinction between the "three" proofs, claiming that Kasia's proof is more elegant than Basia's. I seem to

[8]However, as we'll see vividly later in this chapter, some versions of a proof obscure the core idea so thoroughly that it takes much work to uncover it.

be contradicting myself.

To clarify these contradictory intuitions, let's compare Asia's, Basia's, and Kasia's conversation above to one from a more paradigmatically aesthetic domain: a discussion of the song "Mr Tambourine Man." Suppose the girls are waxing poetic about the song's beauty, and suppose further that Asia has heard the Dylan song, Basia heard The Byrds' cover, and Kasia read the lyrics and a score for an orchestral-choral version.

Are the girls talking past each other in this case? It depends. There's much they might be agreeing about: the rhythm (at least in a coarse-grained sense), the melody, the lyrics. There's a core idea shared by the song and its cover which allows us to call them essentially the same, versions of the same thing. And if Kasia is particularly musically gifted, there's no reason for her not to access this core idea through the score, without hearing any performances of the music. So if the girls' assessment of the beauty of the song is drawing on these common features, the girls are, in an important sense, not talking past each other. (If you have doubts about this particular case, consider classical music: surely Asia and Basia can agree about the beauty of Beethoven's Pastoral symphony while having heard different performances.) On the other hand, if Basia continues "I love the song for the airy vocals!" she *is* talking past Asia; (needless to say) there are no airy vocals in the Dylan version.

The broad takeaway from this discussion is the familiar point that aesthetically evaluating a given song or musical piece can come apart from aesthetically evaluating its particular performances or interpretations. Asia can like "Mr Tambourine Man" (and not just in the sense of liking the

Dylan version), but think the airy vocals of the Byrds' version are a fatal flaw. What it means for a given performance to be of a given song, and what a song is over and above its performances are vexed questions; for our purposes we need only note that the question closely parallels that of the relationship between a proof and its various presentations.[9]

Let me spell out some analogies between music and mathematics which are beginning to emerge here. Two performances of the same song or piece share a "core idea;" two presentations of the same proof share a core idea too. Evaluating core ideas can come apart from evaluating their particular performances or presentations—but evaluating particular performances or presentations is often guided by our sense of the core idea. Some presentations are more adequate, do better justice to the core idea: hence the judgments that Basia's proof is a "constructive proof in disguise" or that the Byrds' airy vocals fail to do justice to the mood of the song. Proofs and performances are normative kinds: there are better and worse instances, and often goodness is measured at least in part by faithfulness to the core idea.

However, the core idea isn't some distant ideal, floating free and unmoored from all the presentations of the corresponding proof or perfor-

[9]One might argue that in music, the intentions of the original songwriter constrain the core idea more than the intentions of the discoverer of a new proof. I think that's not actually the case: there are covers which are better than the originals and reveal something that was there in the initial song, but only vaguely sensed by its creator. Bob Dylan considered Jimi Hedrix's version of "All Along the Watchtower" to be a case of this sort, remarking [Wikipedia contributors, 2020] that Hendrix "could find things inside a song and vigorously develop them. He found things that other people wouldn't think of finding in there. He probably improved upon it by the spaces he was using. I took license with the song from his version, actually, and continue to do it to this day." Thanks to Peter Koellner for the example.

mances of the corresponding piece. We only have access to proofs through their various presentations; we only have access to musical songs and pieces through their performances (or scores or recordings). It takes work to extract a core idea, and often this work can only be done by comparison with variant presentations. That Euclid's proof is essentially a constructive proof, that there is a weariness to the mood of Dylan's song, is best discovered by comparing alternative presentations.[10] Still, there is room, in mathematics and in music, for a sense that no presentation has ever fully done justice to the "core idea." Gian-Carlo Rota, in [Rota, 1997] "The Phenomenology of Mathematical Beauty," appears to agree:

> Some beautiful theories may never be given a presentation which matches their beauty.

Whether or not two people aesthetically evaluating a given proof or song (as opposed to a given performance or presentation) are talking past each other is discovered through examining the reasons they give for their judgments. If Basia says "I love the airy vocals" in the musical case and "I love how it's a proof by contradiction" in the mathematical one, she is talking past her friends. (Once she says this, the friends may exchange the versions of the proof or song they are familiar with, and a productive discussion may develop: the girls can ask whether the feature in question is good-making, whether it's essential to the identity of the proof, and whether it furthers or detracts from the core idea, etc.) But in many cases there is enough of a core idea shared between various presentations of a proof or

[10]See [Tappenden, 2008] for an argument that the correct definition of "prime number" could only be developed after significant development in abstract number theory.

performances of a song for our interlocutors to be talking essentially about the same thing.

There is a sense of content—namely, core idea —which is common to music and mathematics, and in music, too, we make aesthetic judgments about the content in a way somewhat severed from our judgments about the form. Conversely, in mathematics, too, we care not only about the core idea, but about particular presentations of a given proof—we judge them to be better or worse, more or less beautiful, more or less adequate to the core idea of the proof. And in mathematics as in the arts, any small feature of a presentation of a given proof may turn out to be aesthetically relevant: that we multiplied together the first n numbers rather than the first n primes, say.

All of this will be even starker in our case study. I have explored this toy example to prepare you for the richness of that study: the way in which what is essentially a single proof may have multiple presentations which vary on multiple aesthetic and epistemic dimensions. It's time to dive into this richness.

2 Quadratic Reciprocity

If you're a mathematician, you may want to skip the following subsection, which is a crash-course in some number-theoretic notions.

2.1 A Gentle Introduction

In the United States, after the hours on a clock reach 12, they return to 1. Other countries use a 24-hour system: after 12 comes 13—but after 24 (or 0) o'clock, it is 1 o'clock once again. We all know how to translate between these two systems: 13 o'clock is 1, 14 is 2, etc. We can use the symbol $\equiv$ to summarize these translations:

$$13 \equiv 1 \bmod 12.$$

The "mod" (pronounced "modulo") here just means that it's a 12-hour clock, which resets to 1 after 12 (or, equivalently, to 0 after 11). We could imagine using clocks which reset after any number of hours. Then if our clock resets to 1 after c, we could write $a \equiv b \bmod c$ whenever a is a whole number of rotations from b (backwards or forwards) on the clock. We can translate back from this modular notation to the familiar = sign: $a \equiv b \bmod c$ just means that $a = b + kc$ for some integer k. This is equivalent to saying that c divides $a - b$. (The standard mathematical notation for this claim is $c|a - b$. I will occasionally use this in what follows.)

Number theorists have made good use of this notion of clock—or "modular"—arithmetic. Though we don't usually multiply hours together, we can extend the notation to cover multiplication, with $13 \cdot 14$ equivalent to $1 \cdot 2 \bmod 12$, etc. This extended notion works particularly nicely for a "clock" which resets to 1 after a prime number p of hours: for instance, every hour a (as long as a isn't divisible by p) has an inverse b such that $ab \equiv 1 \bmod p$.

A convenient piece of modern notation is to write Z_p for the set of numbers $\{1, 2, ...p - 1\}$, with operations of addition and multiplication defined as above—by clock arithmetic with a p-hour clock. Then instead of writing "$a \equiv b \bmod p$" we write "$a = b$ in Z_p." In mathematical jargon, the "nice" properties alluded to above (e.g. the existence of inverses) are summarized as the claim that Z_p is a field.

In many ways, $\equiv$ parallels our old notion of $=$, and pushing the parallels between these notions has been, historically, a fruitful avenue of mathematical research. For instance, we're all familiar with the notion of a *square root*: a solution to the equation $x^2 = a$, for some number a. In a revelation which we're all used to, but which shocked the Pythagoreans, not every such solution is a rational number. For instance, $\sqrt{2}$ is not rational. So the quadratic equation $x^2 = a$ has rational solutions for some a (e.g. 4), but not all (e.g. 2). We can ask a similar question about the solubility of quadratic equations in the field Z_p. That is, we can ask when $x^2 \equiv a \bmod p$ has a solution. This question is important enough that it has acquired its own terminology, which I introduce in the following subsection.

2.2 The Theorem

Definition 1. *A number a is a square mod p iff there is a number x such that $x^2 \equiv a \bmod p$ (that is, if there is a number $y \in \mathbb{Z}$ such that $x^2 = a + py$).*

The Quadratic Reciprocity Theorem relates the conditions under which an odd prime number p is a square mod (odd, prime) q to those under which q is a square mod p. Namely, in most cases p is a square mod q if and only if q is a square mod p. The only exception is if p and q are both of the form $4n + 3$ for some integer n (that is, when p and q are both equivalent to 3 mod 4). In that case, p is a square mod q iff q is *not* a square mod p.

This is a striking and surprising theorem. There is no *prima facie* reason to think that the existence of a solution to $x^2 = p$ in Z_q would have anything to do with the existence of a solution to $x^2 = q$ in Z_p. After all, Z_p and Z_q are completely unrelated fields! This simple and surprising relationship is part of what Gauss meant when he called this theorem "elegant."

The theorem is usually stated in slightly different and more perspicuous form, using the Legendre symbol, $\left(\frac{x}{p}\right)$.

Definition 2. *For prime number p and integer x with p not dividing x,*

$$\left(\frac{x}{p}\right) = \begin{cases} 1 & \textit{if } x \textit{ is a square mod } p \\ -1 & \textit{otherwise} \end{cases}$$

Note that the exponent $\frac{p-1}{2}\frac{q-1}{2}$ is odd precisely when p and q are both equivalent to 3 mod 4. This means that the Quadratic Reciprocity theorem is equivalent to the following.

Theorem 2 (Quadratic Reciprocity)**.** *For p, q–distinct odd primes,*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\frac{q-1}{2}}.$$

The Legendre symbol has some nice properties which justify its introduction. For instance, it is multiplicative. That is, for any a, b not divisible by p,

$$\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$$

This property implies that the question “is a a square mod p?” reduces to questions about whether a’s prime factors $p_1, ...p_n$. are squares mod p. This is because

$$\left(\frac{p_1 p_2 ... p_n}{p}\right) = \left(\frac{p_1}{p}\right)\left(\frac{p_2}{p}\right) ... \left(\frac{p_n}{p}\right) \tag{2.1}$$

Another nice property is periodicity in the top argument: if $a \equiv b$ mod p,

$$\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right) \tag{2.2}$$

Finally, *en route* to proving the Quadratic Reciprocity, one can prove that[11]

$$\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}} \tag{2.3}$$

Together, these properties allow us to use the Quadratic Reciprocity Theorem to answer the question “Is a a square mod p?” for *any* a. We just use 2.1 to reduce the question to a series of questions of the form “is p_i a square mod p?”, for prime p_i. Then we use the Quadratic Reciprocity Theorem, together with 2.2, to reduce the question to the calculation of Legendre symbols with progressively flipped “numerators” and “denominators” and progressively decreased numerators. Eventually, we’ll end with something with either 1 or 2 in the numerator, which can be easily calculated by 2.3.

The formulation of the theorem which uses the Legendre symbol, then, highlights the great utility of the theorem. Since it involves an algebraic

[11] I will omit the proof for the sake of brevity.

expression, it is also easier to prove, and so this is the formulation of the theorem which we'll show in what follows.

Just a few preliminaries remain before the proofs. They all draw on the following fact.

Theorem 3 (Euler's Criterion). *For any p prime and x not divisible by p,*

$$\left(\frac{x}{p}\right) \equiv x^{\frac{p-1}{2}} \ mod\ p.$$

In what follows, we'll also make use of the following notion.

Definition 3. *A number r is the* residue *of x mod p iff $r \equiv x$ mod p and $0 \leq r < p$.*

r is the least absolute residue *of x iff $r \equiv x$ mod p and $-\frac{p}{2} < r < \frac{p}{2}$.*

Recall also that $\lfloor x \rfloor$ is the *floor* of x, i.e. the greatest whole number $a \leq x$.

3 The Proofs

Unless otherwise specified, q and p are throughout as in the statement of the Quadratic Reciprocity Theorem.

3.1 The Modern Proof

The general structure of the proof may be summarized as follows.

1. In the end, the exponent $\frac{p-1}{2} \cdot \frac{q-1}{2}$ is interpreted as the area of a $\frac{p-1}{2} \times \frac{q-1}{2}$ rectangle formed as the union of two congruent right-angled triangles.

2. Since $\left(\frac{q}{p}\right) \equiv q^{\frac{p-1}{2}} \bmod p$ (Euler's criterion), multiplying together $\frac{p-1}{2}$ appropriately selected multiples of q mod p gives $\left(\frac{q}{p}\right)$ times a certain manageable quantity.

3. We show that the manageable quantity is equal to ± 1, where the sign is determined by the parity (evenness or oddness) of the number of points in one of the right-angled triangles from (1), so:

4. Switching the roles of p and q and putting things together, we get the theorem.[12]

Going back in time, to Eisenstein and then to Gauss, leads us to proofs which ultimately share this structure, but in progressively more obscure forms. Keeping the structure in mind will therefore help us understand the proofs, but we should be careful to remember that this is an anachronistic imposition.

Here, then, is the modern proof.

Lemma 3.1.1 (Gauss's Lemma)**.** *Let p, q—odd primes. Take the list*

$$q,\ 2q,\ 3q, \ldots\ \tfrac{p-1}{2}q,$$

and replace each number by its least absolute residue mod p, forming the list

$$b_1, b_2, \ldots, b_{\frac{p-1}{2}}$$

Let α be the number of negative numbers on the list.

Then

[12] I owe this formulation to Prof. Barry Mazur.

$$\left(\frac{q}{p}\right) = (-1)^{\alpha}.$$

Proof. We show that $\{|b_i| : i < \frac{p}{2}\} = \{1, 2, ..., \frac{p-1}{2}\}$. All equivalences are mod p.

$\{|b_i| : i < \frac{p}{2}\} \subseteq \{1, 2, ..., \frac{p-1}{2}\}$, so it suffices to show that all the $|b_i|$ are distinct. Suppose not, i.e. $|b_i| = |b_j|$ for some $i \neq j$. Then there are two cases.

(1) If $b_i = b_j$, then $qi \equiv qj$, so (since q, p coprime), $i = j$.

(2) if $b_i = -b_j$, $qi \equiv -qj$, so that $q(i + j) \equiv 0$. Since q, p coprime, this entails $p|(i + j)$. But this can't be the case, since $i, j < \frac{p}{2}$.

So indeed $\{|b_i| : i < \frac{p}{2}\} = \{1, 2, ..., \frac{p-1}{2}\}$.

Now, mod p,

$\Pi_{i=0}^{i=\frac{p-1}{2}}(qi) \equiv \Pi b_i \equiv (-1)^{\alpha}\Pi|b_i|$.

The LHS is equal to $q^{\frac{p-1}{2}}(\frac{p-1}{2}!)$, and by what we have just shown, the RHS is $(-1)^{\alpha}(\frac{p-1}{2}!)$. Therefore (dividing by $\frac{p-1}{2}!$), $q^{\frac{p-1}{2}} \equiv (-1)^{\alpha}$.

By Euler's criterion, it follows that

$\left(\frac{q}{p}\right) \equiv (-1)^{\alpha}$, and since the LHS and RHS are both equal to ± 1, we have equality.

□

Lemma 3.1.2. $\alpha \equiv \Sigma_{0<a\leq\frac{p-1}{2}}\lfloor\frac{qa}{p}\rfloor$ *mod* 2.

Proof. For each $0 < a \leq \frac{p-1}{2}$, we have

$qa = \lfloor\frac{qa}{p}\rfloor p + l_a$, some $0 < l_a < p$. Now if qa's least absolute residue is negative, write $-s_a$ for this residue, and if it's positive, write r_a for it. Then if qa's least absolute residue is negative,

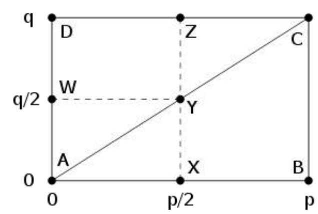


Figure 2.1: $y = \frac{q}{p}x$

$qa = \lfloor \frac{qa}{p} \rfloor p + p - s_a$, so that

$\Sigma_{0<a\leq \frac{p-1}{2}} qa = \Sigma \lfloor \frac{qa}{p} \rfloor p + p\alpha - \ \Sigma s_a + \Sigma r_a$

Now by the proof of Gauss's Lemma, $\Sigma s_a + \Sigma r_a = \Sigma a$, so that mod 2

$q\Sigma a \equiv \Sigma \lfloor \frac{qa}{p} \rfloor p + p\alpha - \Sigma a + 2\Sigma r_a$.

Since p and q are odd, this gives

$2\Sigma a \equiv \Sigma \lfloor \frac{qa}{p} \rfloor + \alpha + 2\Sigma r_a$, i.e.

$\alpha \equiv \Sigma \lfloor \frac{qa}{p} \rfloor$.

□

Lemma 3.1.3. $\Sigma_{0\leq a\leq \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor + \Sigma_{0\leq a\leq \frac{q-1}{2}} \lfloor \frac{pa}{q} \rfloor \equiv \frac{p-1}{2}\frac{q-1}{2}$ *mod* 2.

Proof. $\Sigma_{0\leq a\leq \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor$ is the number of points with integer coordinates (so-called "lattice points") inside triangle AXY in figure 2.1. Since $y = \frac{q}{p}x$ is equivalent to $x = \frac{p}{q}y$, $\Sigma_{0\leq a\leq \frac{q-1}{2}} \lfloor \frac{pa}{q} \rfloor$ is the number of points inside triangle AWY. Since there are no points on the diagonal, the required sum is the number of points in rectangle $AXYW$, i.e. $\frac{p-1}{2}\frac{q-1}{2}$. □

Lemmas 3.1.1 and 3.1.2 hold for p and q swapped, so that

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\Sigma_{0\leq a\leq \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor + \Sigma_{0\leq a\leq \frac{q-1}{2}} \lfloor \frac{pa}{q} \rfloor},$$

and so Lemma 3.1.3 gives the Quadratic Reciprocity Theorem.

3.2 Eisenstein's Proof

For Gauss's and Eisenstein's proofs, I try to stay as close as possible to their presentation, while taking the liberty of presenting the proofs in a more modern style and regrouping parts into lemmas which correspond to the modern proof. Here, then, is Eisenstein's proof.

Lemma 3.2.1 (Eisenstein's Lemma)**.** *Let r range over least residues mod p of products qa with a—even, $0 < a < p$. Then*

$$\left(\frac{q}{p}\right) = (-1)^{\Sigma r}.$$

Proof. We show that for r ranging over the residues in the theorem, the residues of $(-1)^r r$ range over the set of positive even numbers, without repetition.

Clearly, the residue of $(-1)^r r$ is even (as it's equal to r if r is even and equal to $p - r$ if r is odd—a difference of two odd numbers.)

There is no duplication among the list of residues of $(-1)^r r$, since otherwise for some even $a \neq b$ with $0 < a, b < p$, mod p

$$qa(-1)^{qa} \equiv qb(-1)^{qb},$$

which (since q and p are coprime and $a \neq b$) would imply that $a \equiv -b$. But then $a = p - b$ and since p is odd and a, b—even with $a, b < p$ this is impossible.

Hence, mod p,

$$\Pi r \equiv \Pi_{a \text{ even}}(qa) \equiv q^{\frac{p-1}{2}} \Pi_{a \text{ even}} a$$

and

$$\Pi_{a \text{ even}} a \equiv \Pi(-1)^r r,$$

hence (by Euler's criterion) $1 \equiv q^{\frac{p-1}{2}}(\Pi(-1)^r) \equiv \left(\frac{q}{p}\right)(-1)^{\Sigma r}$.

Since $\left(\frac{q}{p}\right)$ and $(-1)^{\Sigma r}$ are both equal to ± 1, it follows that

$\left(\frac{q}{p}\right) = (-1)^{\Sigma r}$. □

Lemma 3.2.2 (Eisenstein)**.** $\Sigma r \equiv \Sigma_{even\ a} \lfloor \frac{qa}{p} \rfloor$ *mod* 2*.*

Proof. We have, by the definition of the r's:

$\Sigma qa = p\Sigma \lfloor \frac{qa}{p} \rfloor + \Sigma r.$

Now since all the a's are even, the LHS is even too, and since p is odd, the parity of $\Sigma \lfloor \frac{qa}{p} \rfloor$ must be the same as the parity of Σr. □

Lemma 3.2.3 (Eisenstein)**.** $\Sigma_{even\ a} \lfloor \frac{qa}{p} \rfloor$ *is equivalent, mod* 2*, to the number of lattice points (i.e. points with integer coordinates) in triangle AXY in Figure 2.1.*

Proof. Note that $\Sigma_{\text{even } a} \lfloor \frac{qa}{p} \rfloor$ is the number of lattice points with even x-coordinates in the triangle below $y = \frac{q}{p}x$, with $x < p$ (triangle ABC in Figure 2.1).

Take $a > \frac{p}{2}$. There are $q - 1$ lattice points at a inside the rectangle $XBCZ$. These divide into the number of points below $y = \frac{q}{p}x$ plus the number of points above (since no integer with $x < p$ is of the form $\frac{q}{p}x$).

Now $q - 1$ is even, so the parity of the number of points below $y = \frac{q}{p}x$ must be the same as the parity of the points above. Therefore replacing the numbers below $y = \frac{q}{p}x$ with the ones above for $a > \frac{p}{2}$ won't change the parity of the sum.

Now note that the number of points above $y = \frac{q}{p}x$ at a is the same as the number of points below $y = \frac{q}{p}x$ at $p - a$, so that the parity of the number of lattice points with x-coordinate bigger than $\frac{p}{2}$ is the same as the parity of the number of points in triangle AXY with odd x coordinates. This means that $\Sigma_{\text{even } a<p}\lfloor \frac{qa}{p} \rfloor$ has the same parity as the number of lattice points in triangle AXY. □

Eisenstein deduces Quadratic Reciprocity from these three lemmas as follows (paralleling the modern proof).

Proof. (Eisenstein).

Since $y = \frac{q}{p}x$ is equivalent to $x = \frac{p}{q}y$, by a symmetric argument $\left(\frac{p}{q}\right) = (-1)^n$, where n is the number of points in triangle AWY in Figure 2.1.

Hence $\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{n+m}$, where m is the number of points in triangle AXY and n is the number of points in triangle AWY. Since there are no points on the diagonal, $n + m$ is the number of points in rectangle $AXYW$, i.e. $\frac{p-1}{2}\frac{q-1}{2}$, which is the exponent we want. □

3.3 Gauss's Proof

Like the modern proof, Gauss's proof begins with Gauss's lemma. However, Gauss states the lemma in a slightly different form, with least absolute residues instead of residues. The lemmas are obviously equivalent, but working with least absolute residues makes the modern proof ever so slightly more streamlined. I include this variant formulation here to highlight the way in which aesthetic improvements in mathematics range along a continuum, with some improvements, like this one, involving extremely minor presentational differences.

Lemma 3.3.1 (Gauss's Lemma)**.** *Let p, q—odd primes. Take the list*

$$q,\ 2q,\ 3q, \ldots\ \tfrac{p-1}{2}q,$$

and replace each number by its residue mod p, forming the list

$$b_1, b_2, \ldots, b_{\frac{p-1}{2}}.$$

Let α *be the number of items on this list which are greater than* $\frac{p}{2}$.

Then

$$\left(\frac{q}{p}\right) = (-1)^{\alpha}.$$

The proof begins by forming a list $(c_1, c_2, \ldots, c_{\frac{p-1}{2}})$ by replacing b_i with $p-b_i$ whenever $b_i > \frac{p}{2}$—that is, by forming the list of least absolute residues—and continues as in the proof of Lemma 3.1.1.

Lemma 3.3.2 (Gauss)**.** $\alpha = \Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{2qa}{p} \rfloor - 2\Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor$

Proof. For any rational number x, recall that $\{x\}$—the fractional part of x. Then,

$$\{x\} \geq \tfrac{1}{2} \text{ iff } \lfloor 2x \rfloor - 2\lfloor x \rfloor = 1$$

and

$$\{x\} < \tfrac{1}{2} \text{ iff } \lfloor 2x \rfloor - 2\lfloor x \rfloor = 0.$$

(This follows by a straightforward calculation from the definitional $x = \lfloor x \rfloor + \{x\}$.)

Let r be the residue of some qa mod p. Then $r = \{\frac{qa}{p}\}p$. It follows that

$$r \geq \frac{p}{2} \text{ (i.e., since } \frac{p}{2} \notin \mathbb{N},\ r > \frac{p}{2}) \text{ iff } \lfloor \frac{2qa}{p} \rfloor - 2\lfloor \frac{qa}{p} \rfloor = 1$$

and

$$r < \frac{p}{2} \text{ iff } \lfloor \frac{2qa}{p} \rfloor - 2\lfloor \frac{qa}{p} \rfloor = 0.$$

So counting the elements which α counts is exactly the same as adding up the $\lfloor \frac{2qa}{p} \rfloor - 2\lfloor \frac{qa}{p} \rfloor$ with $0 < a < \frac{p-1}{2}$. □

Lemma 3.3.3 (Gauss)**.** *For* $0 < l < p$, $\lfloor \frac{(p-l)q}{p} \rfloor = q - 1 - \lfloor \frac{lq}{p} \rfloor$

Proof. It follows from the definition of $\lfloor x \rfloor$ that for $b \in \mathbb{Z}, x \in \mathbb{Q} \setminus \mathbb{Z}$,

$$\lfloor x \rfloor + \lfloor b - x \rfloor = b - 1.$$

Now for $0 < l < p$,

$$\frac{(p-l)q}{p} = q - \frac{lq}{p},$$

so that

$$\lfloor \frac{(p-l)q}{p} \rfloor = q - 1 - \lfloor \frac{lq}{p} \rfloor,$$

as required.

□

The next lemma features some intimidating calculations, so let me provide an explanatory gloss. In Lemma 3.3.2, we have reduced α to $\Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{2qa}{p} \rfloor - 2\Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor$, something of the form (*)

$$A - 2B.$$

In Lemma 3.3.4, we will show that this formula is in turn equivalent to something of the form (**)

$$(q-1)C - 2D - B.$$

We do this by applying Lemma 3.3.3 to roughly half of the terms in A, then regrouping the terms. More concretely, recall that $A = \lfloor \frac{2q}{p} \rfloor + \lfloor \frac{4q}{p} \rfloor + ... + \lfloor \frac{p-3}{p} q \rfloor + \lfloor \frac{p-1}{p} q \rfloor$. Using Lemma 3.3.3, we reduce the second (roughly) half of this equation into a sum of $(q-1)$s minus a sum of terms of the form $\lfloor \frac{lq}{p} \rfloor$ with odd l. We group the $(q-1)$s together, giving the first term in (**). The remaining terms in A are of the form $\pm \lfloor \frac{lq}{p} \rfloor$, and we combine those with terms from $-2B$,[13] depending on the parity of l. More precisely, what remains of A is a sum of $\lfloor \frac{lq}{p} \rfloor$ with even l, minus a sum of $\lfloor \frac{lq}{p} \rfloor$ with odd l. The former cancels with corresponding terms in $-B$, the latter combines to form $-2D$ in (**). $-B$ remains, giving (**). Finding the actual values of C and D is just a calculation—the only hard part is keeping track of the indices. Here's Gauss's formulation.

Lemma 3.3.4 (Gauss). *If $p \equiv 1 \mod 4$, then*

$$\alpha = \frac{(q-1)(p-1)}{4} - 2\Sigma_{b=1}^{b=\frac{p-1}{4}} \lfloor (2b-1)\frac{q}{p} \rfloor - \Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor$$

If $p \equiv -1 \mod 4$, then
$\alpha = \frac{(q-1)(p+1)}{4} - 2\Sigma_{b=1}^{b=\frac{p+1}{4}} \lfloor (2b-1)\frac{q}{p} \rfloor - \Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor.$

Proof. From Lemma 3.3.3, we have

$$\lfloor \frac{(p-l)q}{p} \rfloor = q - 1 - \lfloor \frac{lq}{p} \rfloor$$

Substituting this into the last $\frac{p \mp 1}{4}$ terms of $\Sigma_{0<a\leq \frac{p-1}{2}} \lfloor \frac{2aq}{p} \rfloor$ in Lemma 3.3.2 (that is, into the terms $\lfloor 2\frac{qa}{p} \rfloor$ with $a \geq \frac{p+1}{2}$, writing all these terms as $\lfloor \frac{(p-l)q}{p} \rfloor$) and regrouping the terms, we get the required equations. $\square$

[13] It's easier to think of this as $-B - B$, with only the first $-B$ combining with terms from A.

Let me make a remark which Gauss leaves until the very end of the proof. It follows from the above lemma (Lemma 3.3.4) that the parity of α is equal to the parity of $\Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor$. Why? Using the notation from (**) above, the formula in the lemma is $(q-1)C - 2D - B$. D is an integer, so clearly $-2D$ is always even. C is also an integer (since it counts the number of $q-1$s we added together in the proof), and $q-1$ is even, so that $C(q-1)$ is also even.

Eventually, we'll argue that Lemmas 3.3.1 - 3.3.4 hold with the roles of p and q reversed, and so we'll end up with a counterpart to Lemma 3.3.4 with the term $\Sigma_{a=1}^{a=\frac{q-1}{2}} \lfloor \frac{ap}{q} \rfloor$. Setting $\gamma = \frac{p}{q}$, these are two sums of terms of the form $\lfloor \gamma a \rfloor$ and $\lfloor \frac{a}{\gamma} \rfloor$, respectively. Since we'll eventually be multiplying $\left(\frac{p}{q}\right)$ and $\left(\frac{q}{p}\right)$, which corresponds to adding the exponents in Lemma 3.3.4 (and ignoring the exponents of obviously even parity), it makes sense to investigate what happens when you add sums of terms of the form $\lfloor \gamma a \rfloor$ to sums of the form $\lfloor \frac{a}{\gamma} \rfloor$. This is what Gauss does in the following lemma.

Lemma 3.3.5 (Gauss)**.** *For $\gamma \in \mathbb{Q}$, $\gamma > 0$, and $n \in \mathbb{N}$ such that there are no integers among γ, 2γ,..., $n\gamma$, set $b = \lfloor n\gamma \rfloor$. Then*

$$\Sigma_{0<a\leq n} \lfloor a\gamma \rfloor + \Sigma_{0<a\leq b} \lfloor \tfrac{a}{\gamma} \rfloor = nb$$

Proof. First, note that for natural a with $0 < a \leq b$, $\frac{a}{\gamma} \notin \mathbb{N}$. (For otherwise $\frac{a}{\gamma} = k \leq n$, i.e. $a = kx \in \mathbb{N}$, with $0 < k \leq n$.)

Let $l \in \mathbb{N}$. Note that for $\frac{l}{\gamma} < a < \frac{l+1}{\gamma}$, $l < a\gamma < l+1$, so that $\lfloor a\gamma \rfloor = l$.

Now by the definition of $\lfloor x \rfloor$ and by the fact that the $\frac{l}{\gamma}$ are non-integral, $\frac{l}{\gamma} < a < \frac{l+1}{\gamma}$ is equivalent to $\lfloor \frac{l}{\gamma} \rfloor < a \leq \lfloor \frac{l+1}{\gamma} \rfloor$.

Let $\Omega = \Sigma_{0<a\leq n} \lfloor a\gamma \rfloor$.

Group the terms in Ω into chunks: the ones up to and including the $\lfloor\frac{1}{\gamma}\rfloor$th, the ones from the $(\lfloor\frac{1}{\gamma}\rfloor+1)$th up to and including the$\lfloor\frac{2}{\gamma}\rfloor$th, from the $(\lfloor\frac{2}{\gamma}\rfloor+1)$th up to and including the $\lfloor\frac{3}{\gamma}\rfloor$th, ..., from the$(\lfloor\frac{b}{\gamma}\rfloor+1)$th up to and including the nth. We've just shown that the terms in Ω from the $(\lfloor\frac{a}{\gamma}\rfloor+1)$th up to and including the$\lfloor\frac{a+1}{\gamma}\rfloor$th all have value a, so that (canceling terms):

$$\Omega = (\lfloor\tfrac{2}{\gamma}\rfloor - \lfloor\tfrac{1}{\gamma}\rfloor) + (\lfloor\tfrac{3}{\gamma}\rfloor - \lfloor\tfrac{2}{\gamma}\rfloor)2 + (\lfloor\tfrac{4}{\gamma}\rfloor - \lfloor\tfrac{3}{\gamma}\rfloor)3 + ... + (n - \lfloor\tfrac{b}{\gamma}\rfloor)b =$$
$$nb - \Sigma_{0<a\leq b}\lfloor\tfrac{a}{\gamma}\rfloor.$$

□

Lemma 3.3.6 (Gauss)**.** $\Sigma_{0\leq a\leq\frac{p-1}{2}}\lfloor a\frac{q}{p}\rfloor + \Sigma_{0\leq b\leq\frac{q-1}{2}}\lfloor b\frac{p}{q}\rfloor = \frac{p-1}{2}\frac{q-1}{2}$

Proof. This lemma follows from the previous one with $\gamma = \frac{q}{p}$, $n = \frac{p-1}{2}$, $b = \frac{q-1}{2}$.

We just need to verify that the conditions of the lemma are satisfied. Without loss of generality, $q < p$ (for otherwise we can swap the roles of n and b).

There are no whole numbers among $\frac{q}{p}, 2\frac{q}{p}, ..., \frac{q}{p}\frac{p-1}{2}$ because q and p are coprime and $q < p$. So we just need to show that

$b = \lfloor n\gamma\rfloor$, i.e. $\frac{q-1}{2} = \lfloor\frac{p-1}{2}\frac{q}{p}\rfloor$.

Now $\frac{p-1}{2}\frac{q}{p} = \frac{q}{2} - \frac{q}{2p} < \frac{q}{2}$.

Also, since $q < p$, $\frac{q}{2p} < \frac{1}{2}$, so $\frac{p-1}{2}\frac{q}{p} = \frac{q}{2} - \frac{q}{2p} > \frac{q-1}{2}$.

So indeed $\frac{q-1}{2} = \lfloor\frac{p-1}{2}\frac{q}{p}\rfloor$. □

Recall that α was the exponent in Gauss's Lemma (Lemma 3.3.1). Gauss completes the proof of Quadratic Reciprocity by noting that if we set $M = \alpha + \Sigma_{0<a\leq\frac{p-1}{2}}\lfloor\frac{aq}{p}\rfloor$, it follows from Lemma 3.3.4 that M is even. (Recall

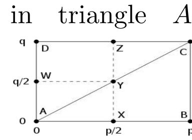
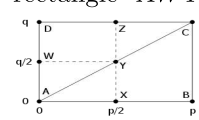
that I spelled out this argument immediately after Lemma 3.3.4.) By a symmetric argument to all of the above, with β playing the role of α in $\lfloor \frac{p}{q} \rfloor$ and N defined correspondingly to M, we get $\alpha + \beta = M + N - \frac{p-1}{2}\frac{q-1}{2} \equiv \frac{p-1}{2}\frac{q-1}{2} \bmod 2$, since M, N—even. Putting all this together,

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\alpha+\beta} = (-1)^{\frac{p-1}{2}\frac{q-1}{2}},$$

which, of course, is the Quadratic Reciprocity Theorem.

It's time for the aesthetic comparison of the proofs. For ease of reference, table 3.3 contains the lemmas from the three proofs.

Table 2.1: Comparison of the Three Proofs

Gauss		Eisenstein		Modern	
Lemma #	Statement	Lemma #	Statement	Lemma #	Statement
3.3.1	$\left(\frac{q}{p}\right) = (-1)^{\alpha}$, where α counts the residues of qa with $a < p$ which are greater than $\frac{p-1}{2}$	3.2.1	$\left(\frac{q}{p}\right) = (-1)^{\Sigma r}$, where r ranges over residues of qa with a—even, $0 < a < p$	3.1.1	$\left(\frac{q}{p}\right) = (-1)^{\alpha}$, where α counts the least absolute residues of qa with $a < p$ which are negative
3.3.2	$\alpha = \Sigma\lfloor \frac{2qa}{p} \rfloor - 2\Sigma\lfloor \frac{qa}{p} \rfloor$ (a ranges over positive numbers $< \frac{p-1}{2}$)	3.2.2	$\Sigma r \equiv \Sigma\lfloor \frac{qa}{p} \rfloor$ mod 2 (a ranges over even numbers $< p$)	3.1.2	$\alpha \equiv \Sigma\lfloor \frac{qa}{p} \rfloor$ mod 2. (a ranges over positive numbers $< \frac{p-1}{2}$)
3.3.4	If $p \equiv 1$ mod 4, then $\alpha = \frac{(q-1)(p-1)}{4} - 2\Sigma_{b=1}^{b=\frac{p-1}{4}} \lfloor (2b-1)\frac{q}{p} \rfloor - \Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor$ If $p \equiv -1$ mod 4, then $\alpha = \frac{(q-1)(p+1)}{4} - 2\Sigma_{b=1}^{b=\frac{p+1}{4}} \lfloor (2b-1)\frac{q}{p} \rfloor - \Sigma_{a=1}^{a=\frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor$	3.2.3	$\Sigma_{\text{even } a}\lfloor \frac{qa}{p} \rfloor$ is equivalent, mod 2, to the number of lattice points in triangle AXY  $y = \frac{q}{p}x$	3.1.3	$\Sigma_{0 \le a \le \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor + \Sigma_{0 \le a \le \frac{q-1}{2}} \lfloor \frac{pa}{q} \rfloor \equiv \frac{p-1}{2}\frac{q-1}{2}$ mod 2 (the number of lattice points in rectangle $AWYX$)  $y = \frac{q}{p}x$
3.3.5	$\Sigma_{0<a\le n}\lfloor ax \rfloor + \Sigma_{0<a\le b}\lfloor \frac{a}{x} \rfloor = nb$ ($x \in \mathbb{Q}^{+}$, $n \in \mathbb{N}$, $b = \lfloor nx \rfloor$, with no integers among x, $2x$,..., nx.)				

4 Comparison

4.1 Gauss vs. Eisenstein

I will begin my comparison of the three proofs by looking only at Gauss's and Eisenstein's proofs and ignoring the hybrid proof, the way a contemporary of Eisenstein's might have done. This is to highlight the ways our aesthetic and other judgments are affected by knowledge of later proofs.

4.1.1 The First Pair of Lemmas

Gauss's and Eisenstein's first lemmas (Lemma 3.3.1 and Lemma 3.2.1) are closely related, and they play the same role in the overall strategy of each proof: the lemmas reduce $\left(\frac{q}{p}\right)$ to a power of -1 so that the rest of the proof can be carried out by manipulating the exponent. To this end, Gauss and Eisenstein both make use of Euler's criterion and reduce $q^{\frac{p-1}{2}}$ to a power of -1 by multiplying together $\frac{p-1}{2}$ multiples qb of q. The main difference is that for Eisenstein the b's run through the even numbers $< p$ and for Gauss they run through $\pm a < \frac{p-1}{2}$, with the sign of a varying depending on whether the remainder of qa is $> \frac{p}{2}$ or not.

Both proofs have a part in which there is shown to be no duplication in a list since otherwise we would have $a \neq b$ with $q(a+b) \equiv 0 \bmod p$, which is shown to be a contradiction, either because $0 < a+b < p$ (Gauss) or because $0 < a+b < 2p$ and a, b—even (Eisenstein).

As the authors of [Reinhard C. Laubenbacher, 1994] point out, Eisenstein's lemma has an algebraic formula for the exponent of -1, which makes it easy to manipulate so that the parity stays the same. By contrast, Gauss's lemma counts the elements of a set, and it's only in the next step of the proof that he reduces this to an algebraic expression. (On the other hand, Gauss's lemma turns out to be useful in other proofs—and, as I'll discuss in the next section, the modern proof utilizes Gauss's lemma while avoiding most of Gauss's calculations.) In general, Eisenstein seems to always keep his end destination in mind, cleverly and quickly manipulating expressions when he knows that only the parity will matter.

The idea to focus on even numbers $< p$ rather than all numbers $< \frac{p}{2}$, and to replace Gauss's lemma with Lemma 3.2.1 was, chronologically, the last one Eisenstein had in coming up with his proof. In a letter to Moritz Stern, Eisenstein writes:

> I did not rest until I freed my geometric proof, which delighted you so much, and which also, incidentally, particularly pleased Jacobi, from the Lemma of Gauss on which it still depended.[14]

At this stage, Eisenstein's choice to focus on even a's can seem like an ad hoc maneuver. By the end of the proof, we'll be able to appreciate the cleverness and foresightedness of this choice. Then, Eisenstein's lemma will take on an elegance akin to that of a well-orchestrated pass of a soccer ball which proves to be a necessary component of a subsequent goal. But this judgment, too, may be short-lived: comparison with the modern proof will reveal that there was, after all, something convoluted in Eisenstein's approach.

Viewed in isolation from the rest of the proof, then, Gauss's and Eisenstein's proof of their first lemmas seem about on a par, aesthetically and mathematically. In the context of the entire proofs... well, we'll see.

[14] The translation is from [Reinhard C. Laubenbacher, 1994], following [Hurwitz and Rudio, 1895].

4.1.2 The Second Pair of Lemmas

Gauss's second lemma (Lemma 3.3.2) involves quite a nice trick that obtains an algebraic expression counting the elements in a set (let's call this set A). Gauss repurposes a straightforward fact about $\lfloor x \rfloor$ to build a binary function which perfectly corresponds to what is called the *characteristic function* of A—a binary function $\chi(x)$ equal to 1 for values $x \in A$, and 0 otherwise. The pleasure I get from this part of the proof is a very modest form of unity amid variety. I see the same object ($\chi(x)$) in two guises at the same time: as the non-constructive characteristic function, and as a simple algebraic formula $\lfloor 2x \rfloor - 2\lfloor x \rfloor$. Furthermore, the move from the nonconstructive to the algebraic version of the function is clearly seen to be another well-orchestrated pass *en route* to proving the theorem.

By contrast, the corresponding part of Eisenstein's proof (Lemma 3.2.2) is ascetic to the point of unremarkableness. But this, too, can be aesthetically appealing—like a minimalist chair with no features other than those needed for sitting.

Note that the two lemmas arrive at almost the same place: Gauss has reduced the exponent of -1 in the formula for $\left(\frac{q}{p}\right)$ to $\Sigma_{a \leq \frac{p-1}{2}} \lfloor \frac{2qa}{p} \rfloor - 2\Sigma_{a \leq \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor$. Since $2\Sigma_{a \leq \frac{p-1}{2}} \lfloor \frac{qa}{p} \rfloor$ is even, this has the same parity as $\Sigma_{a \leq \frac{p-1}{2}} \lfloor \frac{2qa}{p} \rfloor$, i.e. $\Sigma_{\text{even } a<p} \lfloor \frac{qa}{p} \rfloor$, which is the sum to which Eisenstein has reduced the exponent.

4.1.3 The Tail Ends of the Proofs

Comparing the two second parts of the proofs—and noting the role they play in the proofs as a whole—substantially alters the judgment we'd make about the lemmas and their proofs in isolation. Gauss's proof of the second lemma may initially strike us as pleasingly clever, whereas Eisenstein's second lemma is boringly obvious—it amounts to a restatement of the definition of the residue r of qa. But in the context of the whole proofs, we can see Gauss's lemma together with the trick in his second lemma as a detour through counting the elements in a set only to reduce this to an algebraic expression, where it was possible to treat things algebraically all along. Eisenstein has no need for such deceptive cleverness—or so it seems right now, before the hybrid proof.

The geometric argument (which I've presented in two parts, in Lemma 3.2.3 together with the unlabeled final part) is the crowning achievement of Eisenstein's proof. In good visual-proof fashion, areas are rearranged and put back together, and symmetries are made use of. As in all of Eisenstein's proof, parities are simplified as soon as possible. The way $q-1$ (that is, the points at a given even coordinate inside rectangle $XZCB$) is divided up into two parts and the new parts fit exactly in the odd slots in the smaller triangle is particularly lovely, as is the way the second half of the proof follows by symmetry. $\frac{p-1}{2}\frac{q-1}{2}$ is given a simple and natural interpretation as the area of a rectangle.

The beauty of this part of the proof is just so obvious that it's hard to know what to point out. Symmetry, areas fitting together, simplicity all contribute.

How does the last part of Gauss's proof compare to Eisenstein's geometric argument? Well, it turns out that it says almost exactly the same thing as Eisenstein's image—just in overcomplicated words. Here's why.

Recall that Lemma 3.3.3 states that $\lfloor \frac{(p-l)q}{p} \rfloor = q - 1 - \lfloor \frac{lq}{p} \rfloor$, i.e. that

$$\lfloor \frac{(p-l)q}{p} \rfloor + \lfloor \frac{lq}{p} \rfloor = q - 1.$$

Given Figure 2.1, we can see that this statement, when applied to odd $l < \frac{p-1}{2}$, is basically Eisenstein's observation that the points inside rectangle $ABCD$ at a given point $(p-l)$ subdivide into the points below and above the line $y = \frac{q}{p}x$. ("Basically," because what the statement really says is that the points inside rectangle $XZBC$ at a given point $(p-l)$ subdivide into those inside figure $XYCB$ at $p-l$ and those inside triangle AYX at l. In other words, this move of Gauss's condenses Eisenstein's two moves: from $XYCB$ to ZYC to AXY.)

Turning to Lemma 3.3.4, let's start with a presentational issue. Gauss ends his proof of the whole theorem by introducing special notation (M) for the even part of the formula in Lemma 3.3.4. He might have spared himself the trouble, by noting right after the proof of Lemma 3.3.4 that, mod 2, α is equivalent to $\Sigma_{0 \leq a \leq \frac{p-1}{2}} \lfloor \frac{aq}{p} \rfloor$. (Better yet, that should have been the content of Lemma 3.3.4.) Gauss does essentially this in his final step, but this is two long calculations after Lemma 3.3.3, so that the reader is left carrying around a cognitive burden: a formula divided into cases with complicated terms which don't end up being used in the rest of the proof.[15]

[15] Gauss's skill at mental arithmetic was notorious. It's entirely possible that he would

Once we note that Gauss could have struck M from the formula in Lemma 3.3.4, we can see that this lemma ends up in the same place as Eisenstein's Lemma 3.2.3; the number of lattice points in triangle AXY is $\Sigma_{0\leq a\leq\frac{p-1}{2}}\lfloor\frac{aq}{p}\rfloor$.

The parallels don't stop there. Gauss's Lemma 3.3.4 is a formidable symbol-salad, but using Eisenstein's geometric representation as a guide, we can simplify things drastically, revealing a unity beneath Gauss's division into cases.

The last $\frac{p\mp1}{4}$ terms of $\Sigma_{0\leq a\leq\frac{p-1}{2}}\lfloor\frac{2aq}{p}\rfloor$ are just the terms with $2a>\frac{p-1}{2}$ (since the number of these varies depending on whether $\frac{p+1}{2}$ is even or not.) Counting how many of these there are (which $\frac{p\mp1}{4}$ does) is a distraction. We can appreciate here how Eisenstein's presentational decision to write $\Sigma_{0\leq a\leq\frac{p-1}{2}}\lfloor\frac{2aq}{p}\rfloor$ as $\Sigma_{\text{even } a\leq p-1}\lfloor\frac{aq}{p}\rfloor$ highlights the crucial information in this part of the proof.

Lemma 3.3.5 has the appearance of seriousness. Gauss deserves some credit for stating it in more general terms than its corollary Lemma 3.3.6; Eisenstein could have proved the more general theorem by having Figure 2.1 represent not just $y=\frac{q}{p}x$, but a more general $y=\gamma x$ satisfying the conditions in Lemma 3.3.5. However, the price for this generality would have been the necessity of additional notation, which would have made the argument slightly harder to follow.

The proof of Lemma 3.3.4, too, seems respectable. The way each sub-

not have perceived carrying M and N around in his mind as a cognitive burden at all, and indeed would immediately see Eisenstein's and his own treatment as equivalent here. This aspect of the superiority of Eisenstein's proof is almost certainly "subjective" in the sense which Lange is worried about.

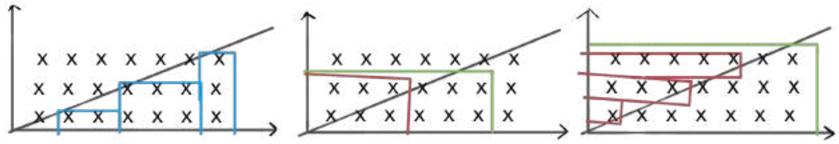


Figure 2.2: $y = \frac{q}{p}x$ (Dividing triangles into rectangles)

sequent term of the form $a(\lfloor\frac{a+1}{\gamma}\rfloor - \lfloor\frac{a}{\gamma}\rfloor)$ cancels out components of the previous term seems like quite a little gem.

I say "seems," because this impression evaporates once we compare Gauss's argument with Eisenstein's. Then we see that where Gauss has this calculation, Eisenstein notes that the two halves of a rectangle add up to form its area.

We can use Eisenstein's geometric represenation to visualize Gauss's proof of the Lemma. To calculate the area of triangle AXY, Gauss divides it into boxes (of side length $(\lfloor\frac{a+1}{\gamma}\rfloor - \lfloor\frac{a}{\gamma}\rfloor)$) whose heights ($a$) increase one unit at a time (the blue boxes in Figure 2.2. He then reconceptualizes each box as the difference of two boxes with corners at the origin (e.g. the difference of the green and red box in the second image in Figure 2.2), and notes that when you keep adding and subtracting the new boxes like this, you're left with the green box minus the red strips in the third image of the figure.

Using this visualization makes Gauss's "gem" look ridiculous: without realizing it, Gauss is calculating the area of a triangle by dividing it up into rectangles, essentially using the definition of an integral for a figure whose area can be calculated by any child.

I've made a series of evaluative comparisons between Gauss's and Eisenstein's proofs. Table 2.2 includes the main evaluative terms I used in making the comparison. The arrows symbolize a temporal sequence of judgments, updated based on engagement with the proof and its alternatives. For example, coming to understand the tail part of Gauss's proof leads me to move from finding it a cognitive burden to thinking it's serious; learning about Eisenstein's proof causes me to update this judgment again and find Gauss's proof ridiculous.

Table 2.2: Gauss and Eisenstein: Evaluative Comparison

	Evaluative terms	
Part	Gauss	Eisenstein
1		ad hoc → clever, well-orchestrated → somewhat ad hoc
2	nice trick, unity amid variety, well-orchestrated → unnecessary detour	ascetic, unremarkable → pared down, minimalist
3	symbol-salad, hard to follow, cognitive burden → serious → general → gem → needlessly complicated → ridiculous	simple, clear, beautiful (because of symmetry, seeing at a glance, areas fitting together)

4.2 Enter the Hybrid Proof

Eisenstein's later proof changes our judgment about Gauss's proof. Impressive tricks—the algebraic expression for α in Lemma 3.3.2, the way terms cancel in Lemma 3.3.4—are reevaluated as unnecessary detours.

The hybrid proof forces us to change our judgment yet again, doing to Eisenstein what we have just done to Gauss. Like Gauss's proof, the hybrid one works with $a < \frac{p}{2}$, rather than even a, throughout. But unlike Gauss's proof, this one is easy to understand. The proof of Lemma 3.1.2 is almost as straightforward as that of Eisenstein's Lemma 3.2.2—it involves the same adding up of definitional equalities, just additionally making use of a part of the proof of Gauss's Lemma. And it lands us right at Eisenstein's Lemma 3.2.3, without the need for the first half of Eisenstein's geometric argument.

Remember what we said about the relationship beteween Gauss's and Eisenstein's lemmas (Lemma 3.3.1 and Lemma 3.2.1)? They seem just about equally simple and easy to understand—and about on a par aesthetically, *until* one sees the role they play in the whole proofs. But now consider the hybrid proof. Intrinsically, its first part (Gauss's lemma) is just as simple as Eisenstein's, and its second part is only very slightly more complicated than Eisenstein's proof of Lemma 3.2.2—and lands us at Eisenstein's Lemma 3.2.3, without the need for Eisenstein's proof of that theorem. So, together, the first two parts of the hybrid proof are at least as simple, insightful, and easy to understand as the first three parts of Eisenstein's proof. And given that the proofs share the final part, that holds true of the entire proofs.

This means that in an important sense Eisenstein didn't have to free his proof from Gauss's lemma. It was possible, after all, to have a non-algebraic exponent and still have an extremely simple and streamlined proof. Eisenstein's decision to look at even numbers $< p$ now looks like a clever trick of limited generality. It's the sort of thing that makes us wonder "how would you come up with this?" and it ultimately turns out to be unnecessary.

Furthermore, our observation above that Gauss's proof of Lemma 3.3.3 just is an overcomplicated version of the first half of Eisenstein's geometric argument can be turned around and used against Eisenstein. Rather than revealing anything essential about why Quadratic Reciprocity holds, the first part of Eisenstein's geometric argument is jut a visual representation of Gauss's. Why did Eisenstein choose to focus on even $a < p$? Only because these make an appearance in Gauss's Lemma 3.3.3. Eisenstein's project in creating his proof was that of cleaning up Gauss's earlier demonstration. But there is only so far you can go in such a project if the proof you're trying to clean up is clunky to begin with.

The hybrid proof, then, is at least as easy to understand as Eisenstein's proof. But does it provide as much understanding? There is reason to think so. Eisenstein's proof of Lemma 3.2.3, while lovely, doesn't seem to have much to do with the reasons for which Quadratic Reciprocity holds. It's just a step getting us back from the detour through even numbers to numbers $< \frac{p}{2}$. By contrast, the second geometric argument, the one which is maintained in the hybrid proof, gives a natural geometric representation of the exponent in Quadratic Reciprocity, and is plausibly part of a good explanation of why the theorem holds.

5 Comparison to other Aesthetic Domains

5.1 Functional Beauty

In terms of simplicity and explanatoriness, then, the proofs' ordering aligns with chronology: Gauss, Eisenstein, hybrid. What about the aesthetic ordering? I find Eisenstein's proof to be more beautiful than either Gauss's or the hybrid proof. I feel disappointed that the more streamlined hybrid proof is missing Eisenstein's first geometric argument.

And yet one can reasonably come to a different aesthetic conclusion. Coming in contact with the modern proof may cause Eisenstein's proof to lose its magic. For me, the beauty of some of the apparent "gems" in Gauss's proof (such as the way terms cancel in the final part of his proof) vanished when they were shown to be unnecessary detours; for you, the same may happen to Eisenstein's geometric argument.

Is Eisenstein's proof *really* more or less beautiful than the modern one? I think there's simply no answer to this question. This isn't a failure of the concept of mathematical beauty, though—rather, it is a fact about beauty in all domains. In particular, whenever an object has a function, our judgments about how well it fulfills this function and how beautiful it is interact in complicated ways, which make multiple incompatible aesthetic experiences and judgments possible and appropriate. To show that this is the case, I'd like to tell you about my bike.

My bike is a happy bright turquoise, it has smooth graceful curves, and it radiates personality. Strangers sometimes stop me on the street to tell me that it's beautiful.

Some people are happy to stop at that, never wondering whether the bike works well, but one lady continued: "it must be quite good!" "It's okay, but the brakes are squeaky, and it's a little heavy," I responded. She seemed disappointed—as if the bike had failed to keep a promise it made with its beauty.

I recently brought my bike to a repair shop (to fix those squeaky brakes), and the staff there remarked that the bike was "cursed with innovation." The parts were connected in ways which made it harder to fix.[16] It pained the staff to look at my bike—they could see that it was badly put together—and though they didn't quite say it was ugly, I bet they thought so. It would be difficult, if not impossible, for them to see the bike as merely a pleasing array of shapes. When they look at it, they *see* how its function is implemented—they see the bike as a mechanism and they judge it to be ugly.

My experience of Eisenstein's proof is like that of the lady who stopped me on the street: I feel deceived, as if the proof's beauty had made and broken a promise of maximum explanatoriness. I'm unwilling to let go of my aesthetic judgment, but I feel uneasy about this unwillingness.

You may be more like the bike repairmen. You may be unable to see Eisenstein's proof as anything but an unnecessary detour. Even in that case, though, you are making an aesthetic judgment: that the proof doesn't

[16]Note that this shows that a bike is a multifunctional artifact, which ought to be not only aerodynamic but easily fixable. Similarly, proofs aim at proving, but also at explanation, generalizability, and more.

please you for its own sake, that contemplating it doesn't give rise to a disinterested pleasure for you. Or you may be making just the opposite judgment: that the proof is beautiful completely independently of how well and transparently it proves. If so, you are like the people who find my bike beautiful whether or not it works well.

5.2 Clean-up, Structure, and Deeper Beauties

All this points towards the richness of aesthetic judgments in mathematics. Some aspects of our judgments, like judgments of the beauty of a bike or a soccer match, have to do with how the pieces of a proof work together, how they achieve the goals of proving and explaining. Other aspects are more self-contained, intrinsic to a piece of mathematics. And among these, there are deeper and shallower beauties: the beauty of the Quadratic Reciprocity theorem itself is deeper than that of Eisenstein's geometric argument.

I'd like to conclude with a few more words about a shallower aesthetic notion: that of clean-up.

Consider the quotidian task of cleaning your room. What are your aims in such a task? You may be looking for something you've misplaced. You may be trying to make the arrangement of furniture in your room more perceptually salient. Or you may be trying to make your room look nicer. Typically, the task of cleaning accomplishes all three goals at once: the structure of the room becomes more transparent, everything is in the right place—and therefore more easily retrieved, and simultaneously more pleasing to behold.

Similarly, mathematical cleanness is is a borderline aesthetic and cognitive value. A clean proof is more pleasing to behold *and* it provides more understanding—at least, if it is simultaneously a deep proof. It's easier to see what the pieces are and how they fit together. Furthermore, these are *both* good and common reasons for wanting to produce clean proofs.

Cleanness is a relatively "shallow" notion—to say that one version of a proof is cleaner than another is to say nothing about the merit of the proof itself, insofar as this can be divorced from a presentation. Even so, the pursuit of clean-up can take a mathematician away from the proof they set out to clean and towards a new, better one, and some forms of cleanup are less about presentation than others.

In his [Leddy, 1995], Thomas Leddy has argued that "clean" is an aesthetic predicate, and we engage in cleaning partly for the aesthetic pleasure clean and organized things afford us. But there is an epistemic aim too: "one of the main functions of neatening and cleaning is revelation of underlying form or structure." Leddy speaks of multiple "ontological layers" which can be cleaned: a painting's surface may be cleaned of impurities, without affecting the underlying structure—but the structure itself, the composition, can also be "cleaned up."

> The features of the composition which are cleaned might still be considered surface features in that they are surface features of the composition. There is something that underlies those features, a more basic form of the composition, which is cleaned up.

Similarly, parts of Eisenstein's proof are cleaned-up versions of Gauss's. Gauss's argument at the end of the proof, with M and N, differs only in surface features from Eisenstein's decision to cut out even parts of the exponent as early as possible. This makes Eisenstein's proof easier to understand—Gauss has a distracting pile of even coefficients added up to his exponent until the very end of the proof, something he could have swept away much earlier—but the difference is cosmetic. Similarly, Gauss's division into cases in Lemma 3.3.3, and his decision to explicitly calculate the number of terms he is adding ($\frac{p\pm1}{4}$) are messy distractions. When Eisenstein decides to add the $\lfloor \frac{qa}{2} \rfloor$ with $a > \frac{p-1}{2}$, he is putting Gauss's two cases into one drawer for ease of cognitive access.

The line between designing and cleaning a room can be blurry, but there is a distinction. As part of the cleaning process, you might buy a new piece of furniture for your room, but if you're replacing the whole set with one in a different style and repainting the walls, that's redesigning. Similarly, Eisenstein's replacement of Gauss's Lemma with "Eisenstein's Lemma" is a borderline case of cleaning and redesign—but the second part of his geometric argument is a real case of redesign. While Eisenstein's visualization may have originated in an attempt to understand Gauss's proof, it ended up being a genuinely novel tool for understanding why the Quadratic Reciprocity Theorem holds.

6 Conclusion

Let's return to the two questions, *Meaning* and *Role*, with which I started this essay.

What does it mean to find a piece of mathematics beautiful? The same thing it means in other domains: to take pleasure in mere contemplation.

I have utilized several comparisons—to the functional beauty of a bike, the cleanness of a room, the elegance of a soccer-ball pass—to argue that, despite entanglements with the cognitive and epistemic, mathematical beauty and elegance are bona fide aesthetic notions. This claim is further bolstered by the richness of aesthetic judgments made in this case, as seen in Table 2.2.

What role does beauty play in mathematical practice? It motivates mathematicians: Gauss wanted a proof whose beauty would match that of the theorem; Eisenstein wanted to free his proof from Gauss's lemma so that he could embed his gem of a visual proof into a complete proof of Quadratic Reciprocity. While beauty can be misleading—it led Eisenstein to a needlessly complicated proof—ultimately it motivates mathematicians' search for understanding.

Chapter 3

Re-Attaching Detachment, Putting the ‘Interest’ Back in Disinterest: Where Aesthetics Meets Meditation

In their simplest forms, meditative practices involve keeping one’s attention fixed on a “meditative object” such as the breath. While such things may seem to have nothing to do with aesthetics, three sets of facts suggest otherwise. First, the experience of pain (or perhaps “pain”) in meditative contexts has much in common with the pain of “negative” emotions in the context of tragic art. In the words of Robert Wright, such pain has

> less ‘youch!’ than usual and more ‘whoa!’ than usual.[1]

[1] [Wright, 2017]

Second, at least anecdotally, meditators have more, and more intense, aesthetic experiences than non-meditators. Third, aestheticians and meditation teachers use strikingly similar language, recommending attitudes of "detachment," "distance," and "disinterest."

In this essay, I argue that these similarities are no coincidence. Meditation as well as aesthetic appreciation involves adopting an attitude of accepting attention. This attitude transforms the character of pain, increases the number and intensity of aesthetic experiences, and is a special, "warm," kind of detachment. It is also an attitude which allows us to see the world and other people more clearly, and so one which ought to be adopted in many contexts on moral and epistemic grounds. A core value of art is that it teaches and motivates us to do so.

The two major parts of the essay investigate the effect of aesthetic and meditative attitudes on pain and pleasure, respectively. I argue that the notion of accepting attention offers a solution to the "paradox of tragedy" in the first case, and clarifies traditional aesthetic notions such as "distance" and "disinterest" in the second.[2]

1 Sublime Pain: Less "Youch!" and More "Whoa!"

1.1 Art-emotions

Artworks in genres ranging from film to music, from literature to painting, appear to arouse our emotions. Indeed, the arousal of emotions is the

[2]The line between the two halves of the essay might also be called the line between the sublime and the beautiful. I choose not to do so simply because I tend to use "beauty" in a wider sense that includes sublimity.

apparent aim of some artistic genres (such as thrillers) and of some aesthetic movements (such as romanticism). But for all their ubiquity, these emotions—if that is indeed the right term—possess many puzzling features.[3] If the artwork in question is a work of fiction, we may feel pity or empathetic joy for people who don't exist, in response to events that never happened. Things get even weirder when we move to other art forms, especially to wordless (so called "pure") music. To listen to Mozart's Symphony No. 40 is to feel distress which is meant for no one and about nothing at all. And music itself appears to be happy or sad, triumphant or wistful, despite, well, not being a conscious agent. It's also not uncommon to feel confused about whether musical emotions are "in" the music or in the listener; whether I am feeling sadness in response to the symphony, or whether I am merely recognizing the sadness in the music.[4]

Perhaps the most puzzling feature of art-emotions is the so-called "paradox of tragedy:" our strange tendency to seek out art which, on the face of it, makes us feel awful. Hume eloquently introduces it in his essay [**?**] "Of Tragedy" [Hume, 1757]:

> It seems an unaccountable pleasure, which the spectators of a well-written tragedy receive from sorrow, terror, anxiety, and other passions, that are in themselves disagreeable and uneasy. The more they are touched and affected, the more are they delighted with the spectacle; and as soon as the uneasy passions

[3] I use the term "art-emotion" for the reactions to art we're pre-theoretically inclined to call "emotions," whether or not that ultimately turns out to be the right word, and the term "life-emotion" for the sort of emotion many philosophers consider paradigmatic (such as fear of a lion). As we'll see below, it's entirely possible to feel art-emotions in daily life, and vice versa.

[4] For the view that the emotions are "in" the listener, see e.g. [Matravers, 1998]; for the "mere recognition" view, see e.g. [Kivy, 1980].

cease to operate, the piece is at an end.

Outside of the realm of art, we tend to avoid the things which make us sad, scared, or angry. So what makes art—high or low, fictional or not—different? Why do we accept the pain it brings, even seek it out? And are we rational to do so?

1.2 Methodology

Emotional responses to art vary wildly between audiences. So when I say that art makes "us" sad, who does this pronoun really refer to? Philosophical discussion has tended to focus on an idealized appreciator with few personal quirks. This makes sense if we are interested in aesthetically "obligatory" responses, especially if we assume that the "real" aesthetic properties are the ones which cause such responses.

This is not my approach. I am interested in a different sort of ideal: that of the most *valuable* art-caused emotions, the responses that might be termed the aesthetically "supererogatory" ones. Aesthetics is not a branch of metaphysics; I see no point in studying the "real" aesthetic properties if they are not tied to valuable experiences. Neither is it a branch of ethics; if there is such a thing as an aesthetic obligation, it's not one anyone would want to discharge by doing the bare minimum. The question I'm interested in, then, is: when art really shakes us up, how are we to characterize *that* emotional response?[5]

[5] The "us" in *this* question is, in the first place, *me*. As I explained in the Introduction, my project is one of getting you to stand where I'm standing, not of prescribing a single "right" way of engaging with art and beautiful objects.

The puzzling features of art-emotions I listed in the previous section are all ways in which they seem "less real" than their real-life counterparts, mere make-believe shadows of the genuine article. These are the features philosophers have tended to highlight—but the sort of art-emotions I'm interested in here are additionally puzzling in the opposite sort of way. Namely, they can appear *realer* than real-life emotions.[6] They can be purer and more intense. At its most powerful, art makes it *easier* to cry at a fictional tragedy than at a real-life one.

The paradox of tragedy acquires a new dimension in such cases. Here, the unpleasantness of negative emotions *contributes* to the value of the experience rather than detracting from it. The experience is valuable *because*, not *in spite* of the negative emotions. In such cases, it's not enough to point to the benefits of art-emotions which outweigh the cost. Such benefits certainly exist, and are numerous (e.g. alleviation of boredom; lessons about history, psychology, morality; purely aesthetic gifts such as beauty), but in the cases I'm interested in below, they are not the whole story.

1.3 Anne of Green Gables

Let me give an example of the sort of valuable art-emotion I'm interested in here. This is my experience of reading an episode of my childhood favorite novel, L. M. Montgomery's *Anne of Green Gables*.

At the death of her beloved stepfather, Matthew, Anne is stunned with shock and unable to cry. This is especially poignant because Anne is a

[6] For the "make-believe shadow" view, see [Walton, 1990]. For a rare example of the opposite sort of view, see [Moran, 1994].

very emotional person, who is easily moved to tears at novels. As I read about this, weeping, I could relate. I, too, found it easier to cry for fictional characters than for real people, easier to mourn their losses than mine.

When I read Anne's story, I had recently experienced a loss too: my grandmother had died. Like Anne, I had been unable to cry. I had hardly known my grandmother—certainly I knew her less well than I knew Anne's Matthew—and what I felt at her death was a mixture of indifference, numb shock, and guilt.[7] And when I cried for Anne's loss, I cried also for myself: for my own inability to cry at real-life tragedies.

Eventually, Anne does start crying for Matthew. So she did love him, even though she couldn't cry? I cry with her. I cry for Matthew too, but my tears are, finally, also for someone else: my grandmother.

This is the sort of experience I am interested in here. Many philosophers would dismiss responses such as mine as "aesthetically irrelevant." After all, if there were such a person as the "ideal reader," she needn't cry for Matthew—and certainly needn't cry for my, or even her, grandmother. That's all true. But it's equally true that fiction is made *for* experiences like mine. If the emotions I experienced in front of tragic art never taught me lessons about my own life, I would stop consuming tragic art. Fiction is a dress rehearsal for life—not the ideal reader's, but any and every reader's life.

There's nothing wrong in principle with focusing on idealized responses to art which track fiction's more "objective" properties. But there is a risk

[7]The guilt was partly caused by the memory of the last thought I had had about my grandmother while she was still alive: whether she might buy me a particular toy.

involved in doing so: the risk of mistaking a theoretical ideal for a practical one. Someone who wants to be an"ideal reader" in the philosophical sense will be encouraged to dismiss, even repress, their own emotional reactions, and focus instead on, say, intellectually appreciating the emotional arc of the plot. If I thought that I ought to be an ideal reader, I might dismiss my sadness for Matthew and my ambivalent feelings about my grandmother as neither here nor there, and focus my attention on, say, the literary qualities of Montgomery's prose. And so I would be robbed of the chance to work through my grief. It is precisely by encouraging acceptance, rather than repression, of our emotions that fiction gives us one of its greatest gifts. We should be wary of a notion of the "aesthetic" so pure that it leads us to scoff at such things.

1.4 Distance

In his 1912 essay, [Bullough, 1912] "Psychical Distance as a Factor in Art and as an Aesthetic Principle," Edward Bullough argues that aesthetic experience requires "inserting" what he calls "psychical distance" between the self and its "affections" (thoughts, feelings, sensations, etc.).

It's not immediately clear what Bullough's scale of "distance" is supposed to measure, but marking a few extremes on the scale should help elucidate the notion. Complete *loss* of distance occurs when we treat our affections as indubitable truth, look "through" them without noticing them, and unquestioningly let them push us around—as in the case of uncontrolled rage. At the other extreme, the person who experiences even his visual perceptions as somehow "unreal" and his emotions as belonging to someone other than himself, who views his experience "with the marveling unconcern of a mere spectator" is overly distanced form his affections.

Bullough's motivating example is that of a fog at sea. Lost in such a fog, we will ordinarily adopt a very close attitude to our affections: the fear of shipwreck is a call to action which blinds us to the particular qualities of the fog, and of the fear itself. We inhabit the emotion, and our entire experience, from the inside, so to speak. But even in such dire circumstances, we can "step outside" our affections and adopt a more distanced attitude, merely observing how things are for us now, seeing the fog more "objectively," not as a threat but as a richly detailed perceptual item, and even our fear as an "uncanny mingling of repose and terror." This, Bullough believes, is when an aesthetic experience is born.

While the title of Bullough's essay is "*Distance* as a Factor in Art (...)," it could equally well have been called "*Closeness* as a Factor in Art." He points out that too much distance can get in the way of aesthetic appreciation. For instance, if we can't relate to the characters of a play at all, we'll have no trouble detaching ourselves from our responses, but we won't appreciate the play. Conversely, too *little* distance can get in the way of appreciation. Here Bullough gives the example of a jealous husband watching *Othello*, for whom the play is too painful for appreciation.[8] Bullough calls this combination of facts—that distance and closeness can *both* get in the way of appreciation—"the antinomy of distance." Bullough's solution to the antinomy is that

[8]This is a somewhat unfortunate example: since the audience knows that Desdemona has *not* been unfaithful to Othello, I'd hope that the play would help Othello's real-life counterpart examine his jealousy more critically. But the theoretical point stands for a real-life Desdemona with an unreasonably jealous husband.

the attitude which leads to aesthetic appreciation is "the utmost decrease of distance without its disappearance"—minimum non-zero distance.

I will eventually criticize various aspects of Bullough's view. But since my own account is an attempt to preserve the virtues of Bullough's picture while correcting the failings, I'd like to begin by pointing out these virtues.

Bullough's view is explanatorily powerful. It provides a single characterization of the aesthetic attitude in art as well as in life. It accounts for several kinds of appreciative failures in a unified way: there can be too little distance or too much, and minimum distance may be hard to reach due to failures of the appreciator as well as of the appreciated object (and, in the case of art, its creator). The metaphor of distance is also apt because literal distance—spatial or temporal—affects aesthetic distance. Books about far-off events tend to succeed only if they bring the story "closer" to us by making it vivid or relevant; depictions of recent tragedies are too close to bear; if they are to be successful, descriptions of familiar everyday life call for a process of defamiliarisation.

Finally, though Bullough's focus on "affections" may initially seem puzzling (we're unaccustomed to thinking of our visual perceptions as "affections" and tend instead to look "through" them to their corresponding objects), it allows him to solve the paradox of tragedy. Here, too, distance comes to the rescue: viewing one's emotions from the outside removes the sting from them, perhaps even causes them to be pleasant. Such detached perception is often enabled by the fictional character of art, but it can also arise spontaneously, as in the case of the person lost at sea who sees his terror as something beautiful.

So much for the virtues of Bullough's view. In the next section, I correct an unsatisfactory aspect of his position.

1.5 Distance vs. Repression

In the case of emotions, Bullough's antinomy is closely related, if not equivalent, to the observation I made above that the most valuable art-emotions can be simultaneously more and less "real" than garden-variety emotions.

While Bullough deserves credit for noticing the antinomy, his own solution is ultimately unsatisfactory. It implies that someone who is just a little dissociative or repressed, who sees her "affections" as almost, but not *quite*, her own, is having an aesthetic experience. The difference between appreciation and repression is qualitative, not quantitative; what Bullough needs is a way of distinguishing between "good" and "bad" distance, and "tiny" is simply not the same as "good" here.

How should we distinguish between aesthetic detachment and repression, then? Well, suppose I am stomping my feet, banging my fists on the table, and yelling at you at the top of my lungs: "I'm not angry!" Clearly, I am mistaken about my inner state. I have separated myself from my feelings so far as to be *unaware* of them. This is a far cry from Bullough's fog-appreciator, who is aware of a rich experiential field which ranges from the way the fog blurs the outlines of things to the terror it causes in him.

At a minimum, then, aesthetic detachment requires awareness of the states one is detaching from. Does it require anything else? Well, consider a slightly less extreme form of repression. Suppose that on some level I'm *aware* of my feeling of anger, but that it's part of my identity that I am a calm and measured person. This causes me to compartmentalize my anger and, as much as possible, ignore it. I look away from it, and notice very few of its properties.

This, too, is strikingly different from the fog at sea. The fog-appreciator turns *towards* her experience, not away from it. She pays attention to it, in all its details, and she accepts all of its aspects, from the milkiness to the terror. If this is right, then the aesthetic attitude has two crucial components which distinguish it from negative forms of detachment such as repression: attention and acceptance.

In the following section, I will carefully consider another valuable experience of art-emotion. Through this example, we'll see more concretely that "good" distance arises out of attention and acceptance. Furthermore, the example will suggest that the antinomy of distance is better thought of adjectivally: aesthetic distance is "close," or affectively warm, distance.

1.6 Love Warrior

In her memoir, [Doyle, 2016] *Love Warrior*, Glennon Doyle describes how listening to the emotionally charged music of the Indigo Girls helped her recover from alcoholism and an eating disorder. If there ever was a valuable experience of art-emotion, this is one.

During the course of the book, Doyle's responses to music undergo a transformation. At first, the music is unbearably painful.

> Her voice makes my whole body ache, like she's holding me down and operating on me without anesthesia.

But the next time she listens to the Indigo Girls, her experience has changed.

> I lie back down and wait for the music to hurt too much. When they start singing, I begin to feel that familiar ache that music always brings. I hold my breath, but I quickly realize that my ache feels different than it used to. Music usually makes me feel left out and yearning, like I'm looking at a photograph of a party I wasn't invited to. But now I feel drawn in, pulled closer, like the music is a bridge between these two women and me. I feel comforted. The Indigo Girls promise me that it's okay to feel too much and know too little. They insist that my sadness is not new, it's ancient. I listen for hours and every song makes me feel less alone and more part of a universal, underground sisterhood. Gradually, I feel something like joy growing inside of me. This joy brings me to my feet and I start to dance.

Note that *both* experiences have aspects of distance and closeness. When the music is like an operation without anaesthesia, it's too close in the sense of being too intensely painful, and in the sense of reminding Doyle too much of her own life. Here, she is like Bullough's real-life Othello. But if we are to trust her later characterization, the first experience also makes her feel "left out and yearning," and it's only during the second one that she feels "drawn in, pulled closer."

Doyle's two experiences show, among other things, that there isn't a single way to feel art-emotions—and that the paradox of tragedy doesn't hold across the board. That is, tragic art *can* feel like an operation without anesthesia, can cause us suffering so profound that we avoid the artwork altogether. The paradox of tragedy, then, is really the question: what makes the difference between Doyle's two experiences? What makes her ache feel "different than it used to," the pain bearable, even joyful—and valuable?

To answer this question, let's look at an experience from a prima facie very different realm: that of meditation. In *Why Buddism is True*, Robert Wright describes how, on the tenth day of a meditation retreat, he began experiencing horrific toothaches whenever he drank anything. As a experiment, he meditated for half an hour, then took a swig of water—maintaining the same attitude towards his pain as towards his breath while meditating. The results were striking.

> I felt a throbbing so powerful that I got absorbed in its waves, but the throbbing didn't consistently feel bad; it was right on the cusp between bitter and sweet and just teetered between the two. At times it was even awesome in the old-fashioned sense of actually inspiring awe—breathtaking in its power and, you might even say, its grandeur and its beauty. Maybe the simplest way to describe the difference between this and my ordinary experience with tooth pain is that there was less "youch!" than usual and more "whoa!" than usual.

Wright's experience isn't isolated. Indeed, meditation-based interventions are increasingly used as a way to manage chronic pain. A recent study [Zeidan et al., 2015] found that such interventions were not only more effective than placebo in decreasing (self-reported) pain intensity and unpleasantness, but also reduced the unpleasantness much more than the intensity. (Placebo reduced both at about the same rate.) In other words, the experiment supports the claim that meditation decreases the "youch!" of pain without proportionally decreasing its intensity.[9]

[9] If we think of Wright's "whoa!" as equivalent to the startlingly pure part of pain, i.e. the *difference* between intensity and painfulness, rather than as intensity itself, then meditation does indeed increase the "whoa!"

What is particularly noteworthy is the parallel to Glennon Doyle's experience. Wright and Doyle both moved from excruciating pain to something still related to pain (an ache that felt different than it used to in Doyle's case, something on the cusp between bitter and sweet for Wright). Both ended in an experience of beauty, or perhaps sublimity. Perhaps most strikingly, "less 'youch!' than usual and more 'whoa!' than usual" is as good a description of the difference between the "pain" of tragic art and garden-variety pain as any I've read. "More 'whoa!' than usual" is part of what I mean when I say that art-emotions are "realer" or more intense than life-ones.

Meditation and art, then, appear to have the power to transform pain in strikingly similar ways. The parallels between art-appreciation and meditation don't stop there. The attitude meditators are supposed to cultivate is strikingly similar to Bullough's "distanced" attitude of the fog appreciator. For instance, they are taught to pay attention to the meditative object (e.g. their breath) in an "objective" and "detached" or "distanced" manner.

What does the metaphor of distance amount to in this context? Let's take the case of distance from one's emotions. We habitually do as our emotions bid—for instance, when we are angry, we believe the feeling and shout accordingly. It's as if we live *inside* the sea of our emotions, and unquestioningly let the waves push us around. The distanced attitude is one of seeing the sea from the *outside*—seeing the forces with a clarity that is impossible from the inside, without necessarily doing as they bid.

So far, we have just a restatement of Bullough's notion of psychical distance. But there is more than this to the instructions of meditation teachers.

The meditative attitude towards one's "affections" is one which permits us to see them more clearly and recognize our habitual reactions. It is not, however, a "cold" attitude: welcoming rather than indifferent, it involves treating our affections with kindness and curiosity. In this sense, it is an attitude of "close" distance.

Another way to put it is that the meditative attitude involves a clarity of vision and a warmth or openness. Indeed, meditative practices appear to teach two core skills: attention and acceptance. On the one hand, by focusing on small and seemingly unimportant sensations such as those accompanying the inhale and the exhale, the meditator learns to increase her focus or attention, and consequently sharpen her ability to notice new aspects of her experience. On the other hand, by treating all sensations and experiences (including the inevitable mind-wandering), pleasant or unpleasant, desirable or not, equally, she learns to cultivate an accepting, open, equanimous attitude.

I believe that this is precisely the attitude Doyle was able to adopt during her second experience of the Indigo Girls. She accepts her feelings; she knows that "it's okay to feel too much and know too little." And she pays careful attention; she's "drawn in, pulled in closer." In the next section, I'll say more about how precisely music enables such a change of attitude.

1.7 Music as a Safe Place

Why did Anne find it easier to cry at fictional deaths than at Matthew's death? It's not too hard to give some answers. Fiction can enhance empathy by making the inner lives of its characters more vivid than those of even our loved ones. It condenses so much action into a small space that we can feel the whole emotional arc reverberating almost at once. Real-life emotions are diluted by imperfect knowledge of others and ourselves, by insufficient time to process the triggering event, by ambivalence arising from the interference of *other* emotions, and by simple distraction.[10]

So fiction "purifies" our emotions by removing counterbalancing feelings, the numbing effect of (too much or too little) time, and ignorance. Pure music continues the process of purification by removing even the propositional content of our emotions—leaving us with their purely bodily aspect (the breathing pattern, the clenched fists, the increased heartbeat, the tension and the release).[11] Since there is no danger of discovering a horrible fact beyond these feelings, we are free to embrace them. Doyle describes the process of coming to embrace or accept her feelings poignantly.

> This becomes my ritual. Instead of drinking, every night I shut my bedroom door and meet with the Indigo Girls. Sometimes I whirl, but usually I just lie in bed and practice feeling my feelings. The music is a safe place to practice being human. In the span of one song I can feel it all, let it all come—joy and hope and terror and rage and love—and then let it pass. The song always ends. I survive every time.

[10]Proust has an eloquent description of such factors on p. 116-117 of *Swann's Way* (Volume 1 of [Proust, 2003]): the narrator's afternoons of reading aroused emotions because they were "crammed with more dramatic events than occur, often, in a whole lifetime" and because "it is only in one small section of the complete idea we have of [a real person] that we are capable of feeling any emotion (...) The novelist's happy discovery was to think of substituting for those opaque sections, impenetrable to the human soul, their equivalent in immaterial sections."

[11]A recent study [Farb et al., 2010] shows that the brains of meditators presented with affectively-laden stimuli show more activation in areas related to bodily awareness. When the stimulus was negative, the meditators scored lower on depression questionnaires than controls, while having comparable levels of self-reported sadness. This is in striking parallel to the experience of emotions when listening to "dark" music: more embodied, less likely to lead to depression, and yet at least as intense.

Feeling it all, letting it all come—that is precisely attention and acceptance, precisely what meditation is about. Art is a "safe place to practice being human" because it's fictional (or non-propositional) and because we know it is finite. You can always close the book, exit the concert hall, pause the music player.

Art teaches acceptance of our emotions—or their lack. As we've seen, *Anne of Green Gables* taught me to accept my ambivalence about my grandmother's death. And one of the core things Doyle learns from the Indigo Girls is that "it's okay to feel too much and know too little." Music, especially pure music, is particularly well-placed to teach this lesson: that it's okay to have feelings which have no propositional content, which aren't "about" anything. In fact, such feelings can be beautiful!

Many philosophers have insisted that such contentless feelings aren't properly called emotions. I don't want to engage in a terminological debate, but I do think that overemphasizing the differences between feelings with and without propositional content can have troubling practical consequences.

In daily life, our conscious access to the content of our feelings varies along a continuum. On the one end is the fear you might feel when chased by a lion. Here, there is no room for doubt: what you feel is fear of the lion. On the other extreme is a perfectly opaque, free-floating depression, which either has no content at all, or no consciously accessible content. In between are cases where affect misattribution is possible: you may wrongly suppose that you feel anxious when in fact you're jittery because of the coffee you just had. But if you introspected just a moment longer, you would have found that it is only jitteriness. So with psychological work, a feeling that seems to belong to one category naturally shifts to the other. To do that work, one must recognize this shift as a conceptual possibility.

We often deal poorly with feelings on the far end of the opacity spectrum, especially if they are negative. Sometimes, we follow them blindly, finding them ever more inappropriate contents, attaching our free-floating anger to anyone who comes in contact with us. Other times, we take the route of repression, turning away from the feeling since we think it doesn't correspond to reality in any way and would only cause us suffering.

Music is emotion externalized, and it can teach us a middle way: a way of simply listening to our feelings. But the conceptual separation of "mere feelings" from emotions can block this lesson. If we think that whatever the music is stirring up in us isn't at all relevant to our daily emotional life—a life which recognizes only the most introspectively transparent emotions—we are taking the road of repression. Too often, that is the philosopher's road.

1.8 Pain Without Suffering

I hope to have made it plausible that the relationship between the pain (or "pain") of art-emotions and life-ones is just the same as the relationship between ordinary pain and pain perceived in a meditative state: less "youch!" and more "whoa!" I have also argued that both differences are due to an increase in attention and acceptance. It's easier to achieve a state of accepting attention in response to art than in daily life because the context of art is a safe and finite one, and because art is rich and interesting enough that we naturally want to attend to it closely.

But why exactly would accepting attention remove the "youch!" from pain? To answer this question, let's look more carefully at the difference between Doyle's first and second experience. Here's a little more context around her first experience.

> Her voice makes my whole body ache, like she's holding me down and operating on me without anesthesia. Her voice and the music are true and deep with longing and both seem directed straight at my heart. This is not a day, or a lifetime, in which I can tolerate remembering my heart.

Doyle feels extreme pain, like an operation without anesthesia, and she can't tolerate remembering her heart. What's the relationship between these two facts? The natural view is that it's the intensity of the pain that causes Doyle to flinch away.

I think the natural view is wrong. During her second experience, (1) Doyle's ache feels more bearable and (2) she accepts the feeling rather than flinching away. If it was (1) which caused (2), we would need a further explanation of (1)—the decrease of "youch." But nothing in Doyle's story explains why the same music could suddenly have felt less painful to her. On the other hand, flinching away is to some extent consciously controllable. While for a given person at a given time, there is probably an upper limit of "youch!" they can withstand without automatically flinching away, for levels below that limit it's up to us whether we turn "towards" the pain or away from it. My diagnosis of Doyle's first experience, then, is the reverse of the natural view: the pain was unbearable in part *because* she couldn't tolerate it.

This is a radical view, but it's supported by experiences like Wright's. Keeping the physical cause of one's pain constant and decreasing one's aversion to the pain appears to decrease the pain's "youch." In the limit, complete lack of aversion leads to a pain with "whoa" instead of "youch" Meditation teacher Shinzen Young summarizes this correlation in a formula:

$$\texttt{suffering} = \texttt{pain} \times \texttt{aversion}.$$[12]

"Suffering" is Young's term for what I have been calling "youch!"—i.e. the (experiential) part of pain that is bad.

This is a radical view, but I have no alternative way of making sense of the transformation that occurs in experiences like Doyle's and Wright's.

1.9 Rilke

In the above sections, I have drawn attention to the ways in which art-emotions are simultaneously more and less "real", more and less fully "ours" than life-ones. I have argued that this transformation is due to the working of accepting attention. In the case of "negative" emotions, acceptance—equivalently, lack of aversion—removes the "youch!" of pain and replaces it with "whoa!" This amounts to an increase in intensity; the "youch!" of pain is a turning away, an empty sensation or perhaps the lack of sensation. This is the sense in which negative art-emotions are "realer" than life ones.

I'd like to conclude this discussion of art-emotions with a poem of Rilke's which captures the special strangeness of music —itself a sonorous piece of word-music. Here's Stephen Mitchell's wonderful translation.

[12] As cited in [Yates, 2015].

To Music

Music: breathing of statues. Perhaps:
silence of paintings. You language where all language
ends. You time
standing vertically on the motion of mortal hearts.

Feelings for whom? O you the transformation
of feelings into what?–: into audible landscape.
You stranger: music. You heart-space
grown out of us. The deepest space in us,
which, rising above us, forces its way out,–
holy departure:
when the innermost point in us stands
outside, as the most practiced distance, as the other
side of the air:
pure,
boundless,
no longer habitable.

Notice the pairs of contrasting terms which suggest that music is simultaneously "inside" and "outside" us. On the one hand, it is a "stranger," a "departure" "rising above us," "the most practiced distance," "the other side of the air." On the other, it's also our "heart-space," "the deepest space" and "the innermost point" in us.

Together, these descriptions capture the double strangeness of musical emotions. The heartache I feel when listening to Shostakovich's Piano Trio No. 2 has a level of intensity I (thankfully) have never experienced in my day-to-day life. The piece is made of distilled despair. The deepest space in us. But in another sense, what I feel with Shostakovich is nothing like what I would feel if I were to, say, lose a loved one. I don't really believe anything tragic has happened when I listen to it, and in a tangled way I'm still enjoying my experience—something unthinkable in the face of a real personal tragedy. The other side of the air.

To listen to Shostakovich is to find, startlingly, that there never *was* a paradox of tragedy. The most intense of "painful" emotions can be felt fully without any suffering at all. Music provides a safe environment in which we can feel more deeply—and fearlessly. And once we experience this, we may find a strange truth: by themselves, our feelings can't hurt us—at least, not in a sense of "hurt" that is any reason at all for avoidance. Once we experience this, we can begin to extend the lesson to real life, by correcting our habit of turning away from our scarier emotions. Though we can't always *enjoy* them, we *can* embrace them. Sometimes such embracing means recognizing that our feelings had been perfectly contentless. Sometimes it entails looking our losses squarely in the eyes, and turning the feelings into forms of honoring.

Perhaps below even the darkest grief there is no suffering—only beauty.

2 Beautiful Pleasure: From Detachment to Accepting Attention

I started this essay by focusing on emotions, especially negative ones, because this is the case in which the difference between the aesthetic and non-aesthetic experience—between art-emotions and life-ones—is the most striking. But Bullough's distanced attitude is supposed to make the difference not just between art- and life-emotions, but also between the beautiful and the merely agreeable. In other words, Bullough thinks that the experience of beauty is a pleasant experience in which we are distanced from our affections, including the pleasure itself. This is the claim I will examine in the remainder of the essay, arguing that the attitude of accepting attention does indeed lead to an experience of beauty.

Just like in the previous section, the experiences I am interested in here are the most *valuable* ones. That is, I'm interested in the kind of beauty that really blows our socks off. (It's perfectly consistent with my view that less than fully accepting attention can lead to aesthetic experience, as long as this is not one of the most valuable experiences. I'm inclined to call these *partial* aesthetic experiences.) This is why I will, once again, lean quite heavily on memoirs as sources of paradigm aesthetic experiences.

2.1 D-theories

Bullough's aesthetics isn't the only one with a d-adjective at the forefront. In fact, philosophical aesthetics abounds in d-adjectives: *disinterested*, *detached*, *distanced*. Variously applied to types of pleasure, judgment, attention, and attitude, these words appear to share a host of unwelcome properties. As the prefixes suggest, they are all negative characterizations—in several senses. On a purely logical level, they tell us only what the thing they describe (pleasure, attitude, etc.) is *not*: interested, attached, close. One negativity breeds another: such characterizations appear to build thick and tall walls between aesthetics and the rest of the world, with little indication of why we should ever cross this barrier. Indeed, the more one focuses on what the aesthetic *isn't*, the smaller one tends to make the area enclosed by these thick walls—till the temple of beauty shrinks to the size of a broom closet. Thus (at least in some moods and on some interpretations), Kant ends up claiming that aesthetic pleasure is incompatible with any desires or with the use of any concepts whatsoever. What sorts of objects can be appreciated in this way? Perhaps nothing more than purely formal patterns: visual and auditory wallpaper. And even if the temple of beauty is a little roomier than this, we may wonder whether it's a place worth visiting. Indeed, d-word aesthetic attitudes can appear downright undesirable. *Dis*interestedness hovers on the brink of *un*interestedness. Detachment has pathological cousins: coldness, dissociation, repression. And those who propose such characterizations tend to uphold unpalatable role models: Kant's disinterested judge is a pretentious fellow who thinks he can command people to like the same things he does; Bullough likens the "distanced" aesthetic attitude to one we experience when "we watch the consummation of some impending catastrophe with the marvelling unconcern of a mere spectator."

My own view is a cousin of d-theories. In the following sections, I'll argue that it doesn't share these theories' (prima facie) flaws. The notion of "accepting attention" introduced above is my attempt at a positive characterization of what Kant, Bullough, and others characterize only negatively.[13] "Acceptance," which may also be called "equanimity," "openness," or even "sympathy," is particularly positive in the sense of being a desirable and warm attitude. As to the shrinking dimensions of the temple of beauty, I'll argue that the temple of beautiful *things* is in fact enormous; what is the size of a broom closet is only the space we must inhabit to get the best

[13]More precisely, Kant's *followers* have only negative characterizations. Kant's own notion of the "free play of the faculties" *is* a positive characterization.

view of this expansive realm. The point of the harsh- and arbitrary-seeming prohibitions—against desires, concepts, and the like—is only to help us get a better view. And I'll argue that we shouldn't think of these injunctions as *prohibitions*, only invitations to relate to our concepts, desires, etc. in a slightly different way. Finally, we've already seen that it's possible to distinguish psychical distance from repression. In the next section, I'll spell this out in more detail, arguing that despite appearances to the contrary, repression doesn't aid appreciation.

Why defend a cousin of d-theories? Because for all their apparent flaws, the d-words have a habit of resurfacing in unexpected places, far removed from philosophical aesthetics. For example, here's how the contemporary artist Catherine Kehoe describes her artistic process.[14]

> I start with a subject that is before my eyes and respond to it in paint. Mostly I remain true to what is there, only to discover how odd and surprising the appearance of things becomes when one puts aside the kind of seeing that helps us navigate in the world. There is another kind of seeing that kicks in when I am painting. It is less about things and more about the surprising relationships between things.

Note the commitment to the existence of two types of seeing—and the claim that even the minimal aim of navigating the world leads to the non-aesthetic, interested kind of seeing. The other kind of seeing, by contrast, isn't even "about things"—presumably this means, in part, doing as Kant bids and abstracting away from one's concepts.

In her memoir, [Milner, 2011] *A Life of One's Own*, Marion Milner de-

[14]In [Seed, 2019].

scribes a striking experience in which she found herself unable to appreciate her sumptuous setting—until, that is, she let go of every one of her desires.

> And once when I was lying, weary and bored with myself, on a cliff looking over the Mediterranean, I had said, 'I want nothing', and immediately the landscape dropped its picture-postcard garishness and shone with a gleam from the first day of creation, even the dusty weeds by the roadside.

Here, the "negative"—and extreme!—attitude of renouncing all desires seamlessly leads to a rich and overwhelmingly positive aesthetic experience. Examples like these suggest that d-views get something right after all.

As we've seen, meditation trains the capacity for accepting attention. Strikingly, meditators appear to have more, and more intense, aesthetic experiences than non-meditators. In his book [Wright, 2017] *Why Buddhism is True*, Robert Wright describes a series of experiences of this sort. In one of them, he was in a melancholy mood during a meditation retreat. He noticed the last pink and purple light of the sunset, and at first, the sight only added to his melancholy. But using his newly acquired meditative skills, he "detached" himself from the feeling, and—behold! –

> the horizon took on a different aspect: it was stunningly beautiful. It had gone from being a reflection of sadness to being a source of delight, even awe.

Wright concludes that there

> seems to be a natural tendency of contemplative practice to strengthen the sense of beauty.[15]

[15] By his own admission, Wright is "flummoxed" by this tendency. Shouldn't we expect

This anecdotal evidence is extremely suggestive: practicing a state of detachment like the one described by Bullough appears to lead to more and richer aesthetic experiences. It may even be that expert meditators find *everything* beautiful. Whether this is really so is ultimately an empirical matter, but in what follows, I will adopt the empirical claim that more meditation leads to more aesthetic experience as a working hypothesis, arguing that it does so by training our capacities for acceptance and attention.

Together, acceptance and attention lead to aesthetic experience. “Lead to” is intentionally ambiguous: it may mean “cause,” or “(partially) constitute.” It may even be that acceptance and attention put us in touch with the platonic property of beauty. I mean to remain neutral between these possibilities. I am also neutral about how precisely attention and acceptance combine—I’ll suggest that they form a special attitude of “accepting attention,” but I’m not wedded to the view. In general, I’m not providing an analysis of the concept of the aesthetic or of beauty, only drawing attention to a cluster of important facts about aesthetic experience in the vicinity of old-fashioned d-theories.

2.2 Revisiting Repression

Above, I argued that aesthetic detachment is distinct from repression. Yet Marion Milner’s experience may suggest otherwise. Recall that her appreciation of the Mediterranean came on the heels of the thought “I want nothing.” Did she really want nothing when she said this? She tells us that

meditative detachment to *lessen* the intensity of our feelings, including the feeling of aesthetic delight? This essay is in part an attempt to flesh out a negative answer to this question.

at the moment of the thought, she was "weary and bored" with herself. It seems likely, then, that she *did* have desires: for entertainment, for a day that wouldn't weary her. Wasn't her "I want nothing," then, a form of repression, of silencing her boredom and weariness, rather than a form of accepting attention?

I can't speak for Milner—but I have had similar experiences in which the answer is invariably "no." In one memorable case, I was on my first spelunking trip. I had been wading through underground rivers for several hours, and I was cold and exhausted. If only I weren't this tired, I'd be able to appreciate the glistening stalactites, the roar of the underground waterfall, the spacious darkness... I was disappointed, in myself and in the caving experience.

Then, a thought appeared in my mind: "There is nothing else." What had I wanted? To wade through an underground river for eight hours with all the thrill of adventure and none of the difficulty? Where could the thrill come from then? To come down here and find that I no longer got in my own way—to come someplace I (with my aches, desires, poor circulation) couldn't reach?

No. The ache in my feet, the cold, the exhaustion, even the disappointment —this *is* the caving experience, this *is* what I came here for.

It only took that thought. In my cold, sore feet, I could feel the miles I'd just walked, the depth of the world I'd reached, the thrill of adventure. Faintly and in the background of that ache, there was the waterfall and the spacious darkness. I didn't mind that they were faint; they were beautiful.

Then, somehow the ache lessened. The desire for a different experience fell away completely. Suddenly the waterfall and the cavernous darkness became the center of my experience, the ache barely registered. What a roar! What spectacular spaciousness!

If Milner's experience was anything like mine, her "I want nothing" should be read as "I accept everything." If my reconstruction is right, these sorts of experiences are dynamic processes, with two crucially different moments. First, one pays attention to and accepts one's current state: aches, desires, boredom, and all. In this first moment, what we may come to think of as the object of the experience (the cave or the view of the Mediterranean) is often only faintly perceived. Rather, there is a holistic, multifaceted and largely positive experience. Then the "negative" aspects of the experience—or rather, those we would have identified as negative prior to our stance of acceptance—tend to fall away, either by lessening in intensity or having less of a grip on our attention. Then, in the second moment, we are freed to focus on the objects around us—which suddenly astound us with their beauty.

2.3 Kantian Disinterest

Milner's experience suggests that letting go of one's desires aids appreciation. This is reminiscent of Kant's view that the pleasure which determines the judgment of taste (i.e. the judgment "this is beautiful") is *disinterested*, i.e. independent of desires. For Kant, the notion of disinterest helps distinguish the beautiful from the merely agreeable on the one hand and the good on the other. Pleasure at the merely agreeable (e.g. at the taste of a dish or at winning an award) is interested because it is pleasure at the satisfaction of a personal desire (e.g. to sate one's hunger or to be admired). Pleasure at the morally good is also interested because we desire the good—indeed, we are rationally required to do so.

If we lose our hunger, food no longer pleases us. By contrast, the lack (or presence) of a desire is irrelevant to whether an object is *beautiful*. Kant illustrates this with the following example: suppose you ask me whether a certain palace is beautiful. I may think that palaces are deplorable wastes of resources. I may even [Kant, 2000]

> easily convince myself that if I found myself on an uninhabited island without the hope of ever again coming among men, and could conjure up just such a splendid building by my mere wish, I should not even give myself the trouble if I had a sufficiently comfortable hut.

But as far as your question about the palace's beauty is concerned, all such considerations are irrelevant. To find it beautiful is to be pleased by its appearance, independently of any preferences for or against its existence.

Kant's view has many critics, but it has had considerable influence on the history of aesthetics. I bring it up here because it is another place where aesthetics meets meditation: the Buddhist tradition from which meditation arose emphasizes letting go of desires, arguing that they lead to suffering. In the first part of this essay, we saw that aesthetic contexts reduce the suffering inherent in "negative" emotions. This fits well with Kant's claim that aesthetic appreciation involves setting aside our desires.

Furthermore, the types of aesthetic experiences which occur on meditation retreats often involve letting go of desires. For instance, Robert Wright describes how a meditation retreat marked the first time he was struck by the beauty of a weed... of a species which he had previously been trying to eradicate in his yard. In a real-life counterpart to Kant's "palace" scenario, it appears that Wright's prejudice against the weed's existence had been getting in the way of his appreciation. (I should reiterate that "letting go" doesn't mean "repressing." Rather, in a two-step process, we first *accept* the desires and then either we end up with a holistic aesthetic experience one of whose objects is our desire, or the desire naturally falls away and we are freed to focus on the beautiful object.)

Wright's desire is mediated by the concept "weed." This lends some plausibility to another of Kant's controversial claims: namely, that the judgment that something is beautiful is independent of concepts. I will revisit this claim in Section 2.6.

2.4 Meditation and Accepting Attention

When I went on a 10-day meditation retreat, after a few days of equanimously attending to my breath and bodily sensations, I began to see dozens of new beauties in the grounds of the meditation center (a border of grass and trees around a small parking lot). Meditation increased the number and intensity of my aesthetic experiences. It also increased my capacity for accepting attention. In this section, I'll give a few examples from my meditation retreat which suggest that the two facts are causally related: it is precisely acceptance and attention that contribute to such aesthetic experiences.

Prima facie, many of my experiences were simply a matter of noticing more. One evening, I saw that the whole lawn was covered in spiderwebs which glinted in the setting sun. The next morning, I admired strings of dewdrops on grasses, spiderwebs, and weeds. After a while, it felt like my eyesight had sharpened: I noticed things I had never seen before, like individual flappings of birds' wings, the shape of raindrops as they hit puddles. All these were things I would have found beautiful without meditation too—had I noticed them. Here, meditation seems to have helped precisely by teaching me to notice more—by sharpening my capacity for attention and perhaps even sharpening my perceptual system.

Even here, though, acceptance was required to *enable* attention. I realized that as a painter, I had been unconsciously filtering out those perceptions I deemed unpaintable. I paid almost no attention to tiny things and to motion. I focused most of my attention on landscape-sized scenes rather than on individual objects, and looked for beautiful color-relationships to the exclusion of all else. Here, my (aesthetic!) desire to make beautiful paintings was, if not precluding aesthetic experience, at least noticeably impoverishing it. It was only when I became open to all my perceptions, whether or not I could extract a painting from them, that I found beauty all around me.

Other experiences involved equanimity more directly. Towards the end of the retreat, I found myself captivated by the pattern of bumps in the asphalt parking lot, strewn with acorns casting long blue shadows. "God help me, I'm appreciating asphalt!" I thought. I worried that this was evidence that, rather than putting me in touch with the world's beauty, meditation worked like a drug which artificially induced rapture. After all, asphalt couldn't *really* be beautiful... It was only when I let go of my embarrassment that I could feel the asphalt's beauty again.

Another day, when examining the dew-strewn grass, I noticed a fuzzy caterpillar glittering with dew. I could see the way it munched on grass with its disproportionately large mandibles, and initially I felt a wave of revulsion. It was only with some effort that I managed to view the caterpillar "objectively," at which point I was flooded with awe. Here, I went through a two-stage process akin to my caving experience: I had to first notice and accept my revulsion, and then I could appreciate the mandible as the powerful source of my reaction.

You may be sceptical that acceptance plays any real role in the examples above, let alone that there is such a thing as "accepting attention." If so, you're in respectable company. George Dickie's 1964 essay, "The Myth of the Aesthetic Attitude" [Dickie, 1964] is a forceful and influential rebuttal of theories such as Bullough's.

Dickie claims that the reason that e.g. the real-life Othello fails to appreciate the play is not that he is watching it in an insufficiently "distanced" way. Instead, he is simply failing to attend to the play (or, at least, to many of its aesthetically relevant features) because he is too preoccupied with his own troubles. Apparent failures of distance are simply failures of (complete) attention.

Dickie further claims that attention is *attention*—that is, two instances of attention can only differ in two respects: what we are attending to, and how closely we are attending. Since we attend to all manner of things during every waking moment of our lives, sometimes extremely closely, the degree of attention can't by itself be a difference-maker between aesthetic and non-aesthetic experiences.

If you're sympathetic to Dickie, you'll probably say the same of my view, substituting "acceptance" for "distance:" acceptance is an idle wheel in the view, and attention can't possibly make the difference between the aesthetic and non-aesthetic.

In the next section, I'll provide some neuroscientific evidence for the existence of various types of attention, including one that may be called "accepting attention." For now, let me say some things which don't depend on such heavy machinery but also call Dickie's assumptions into question.

My starting point in this essay is the claim that meditation increases aesthetic experiences. I think attention and acceptance, the two main capacities trained by meditative practices, have something to do with how it accomplishes this feat, but the precise causal mechanism is open to debate. My examples are consistent with several different causal routes: (1) attention and acceptance might be two independent causes of aesthetic experience, or (2) they might combine into a joint attitude of "accepting attention," or (3) acceptance might cause or enable attention, which in turn causes or constitutes aesthetic experience. Figure 3.1 summarizes these options.

Figure 3.1 leaves a lot of questions unanswered. How and why would acceptance and attention lead to finding more things beautiful? I see several different explanations. On the first, the positive affect associated with aesthetic experience is *built into* the "accepting" part of "accepting attention." That is, the meditator or the aesthete cultivates a "warm" detachment, and the joy of "discovering" that more and more things are beautiful is simply a matter of projecting this warmth onto these things. In other words, the

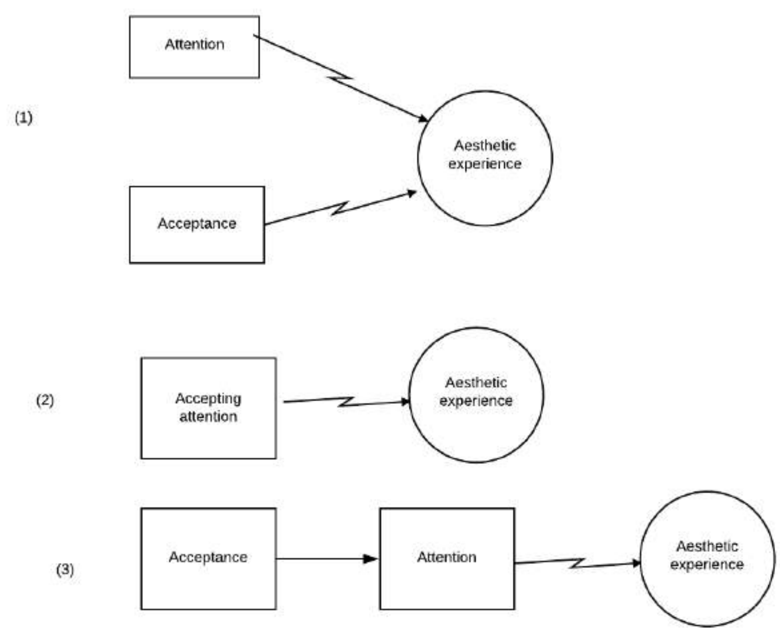


Figure 3.1: Attention, acceptance, aesthetics: causal relationship. (Regular arrows are causation; squiggly arrows are ambiguous between causation and (partial) constitution.)

more we cultivate an open and welcoming attitude towards our experiences, the more we are biased in their favor.

On an alternative, and radically different, view, accepting attention puts us in touch with things' *real* beauty. By cultivating accepting attention, one learns that positive affect is in fact an appropriate response to everything.

Robert Wright describes what may be a middle way when he says that

> Another possibility is that a certain affinity for the universe is a kind of default state of consciousness, a state to which it returns when it's not caught up in the inherently distorting enterprise of operating a self.

On this view, aesthetic experiences *do* put us in touch with a truth, but the truth has at least as much to do with our mind as with the beautiful objects.

First-personally, aesthetic experience *feels* like the second or third option. When I found that caterpillar beautiful at the meditation retreat, it felt like I was *discovering* that it was beautiful, in just the same way in which I was discovering that it had an enormous mandible. At the very least, I was discovering that the caterpillar and my perceptual system where somehow appropriate to each other.

Converting that feeling into philosophical argument lies beyond the scope of this essay. Whatever explains options (1)-(3) in Figure 3.1, I believe that one of them is correct.

Recall Dickie's objection: that acceptance is an idle wheel and that attention can't make the difference between the aesthetic and non-aesthetic. For the sake of the argument, let me assume the first part of the objection:

that acceptance plays no real role in the creation of aesthetic experience. If that's right, option 3 in Figure 3.1 is the only one which could be true. Ultimately, I think option 2, which posits a special kind of accepting attention, is likelier, but I don't want to rule out option 3. In fact, the first few aesthetic experiences from my meditation retreat seem to more closely follow this causal pattern. In these cases, meditation worked primarily by sharpening my ability to notice things, my power of attention, and only minimally, if at all, by working on my acceptance.[16]

Dickie assumes that mere strength of attention can't possibly be the difference-maker between the aesthetic and non-aesthetic. I disagree. It may be that our ordinary attention is much less close than we'd like to believe, and if we just looked more closely, we would find *everything* beautiful: the sunset as well as the dead mouse, the scribble and the mathematical proof. The art teacher Robert Henri believed in this possibility [Henri, 2007]:

> It may be that the enthusiast does not exaggerate and that an excited state is only an evidence of the thrill one has in really seeing.

If, for a fixed object, there is only one dimension on which attention can vary, and if meditation increases aesthetic experience, then Henri's radical claim may well be correct.

A study carried out by Arthur Deikman [Deikman, 1963] suggests as much. Deikman instructed subjects to "meditate" on a blue vase he placed before them. The instructions on what "meditation" meant in this context

[16]These cases are still *consistent* with options (1) and (2). Since I was already inclined to be accepting of perceptions of dewdrops, say, meditation didn't need to sharpen that ability for it to play a causal role in my experience.

were minimal, and required primarily focusing on the vase, to the exclusion of all else. After the experiment, the subjects reported that the vase was "more vivid," "luminous," and "animated" than before. Perhaps they were, for the first time in their adult lives, feeling the thrill of really seeing.

So even if "attention is attention," the difference between an aesthetic experience and its lack can amount to an attentional difference—the difference in *degree* of attention. In the next section, I'll use neuroscientific evidence to argue that attention is in fact *not* just attention. This evidence will also shed some light on the fraught notion that aesthetic appreciation is independent of concepts.

2.5 Attentional Systems

So far, I have been oversimplifying things by treating all types of meditation as if they were the same. In fact, there are two distinct types of meditative practice: "focused" or "concentrative" ones on the one hand, and "open" or "receptive" ones on the other.[17] The former involve focusing on a single object, such as the sensations associated with the breath, almost to the exclusion of all else. The latter cultivate a wide attentional field that takes in the entirety of one's experience: sounds, sights, sensations, emotions, thoughts.

The two types of meditation appear to correspond to two neural pathways: the dorsal attention system and the ventral attention system [Austin, 2009]. The former, corresponding to concentrative practices, is more top-down: that is, it is guided by concepts, purposes, and expectations, which

[17] See e.g. [Austin, 2009].

affect what is perceived. The latter is responsible for what is sometimes called "peripheral awareness" and is more bottom-up: that is, it builds an interpretation out of relatively unfiltered sense-data. However, the correspondence is a matter of degree, and both systems are in fact involved in both types of meditation. To complicate matters further, practitioners of concentrative meditation need to develop their peripheral awareness so that they can catch their wandering mind before they lose track of their object of attention completely. Most, if not all, meditative practices appear to strengthen *both* attentional systems, perhaps by teaching a particular interaction between the two. They also appear to strengthen the bottom-up system more, since even in concentrative meditation there is a progression from seeing one's breath, say, as little more than a series of labels ("inhale," "exhale") to a richly detailed sequence of sensations; a progression from top-down to bottom-up processing.

What is the relationship between "open" meditation and the sort of open, accepting attention I have described above? They are distinct, since one can practice a maximally focused type of meditation such as focusing on breath sensations, and remain open and welcoming to *all* sensations within this narrow set. However, receptive meditation appears to be *maximally* open in the "accepting" sense: it involves openness to *everything* within one's potential field of experience.

The precise relationship between the two attentional systems, meditation, and aesthetic experience has yet to be worked out. Aesthetic experience ranges from the narrowly focused examination of dew-drops to the multifaceted experience of a mountain hike, encompassing types of attention involved in open as well as concentrative types of meditation. But the hypothesis that the ventral (bottom-up) attention system is more involved in aesthetics and meditation than the dorsal (top-down) one, and that the two systems interact in a special way in such experiences, has some philosophically intriguing consequences. Let me spell out a few.

First, as the name suggests, the peripheral attention system is heavily involved in peripheral vision. It may be no coincidence, then, that we sometimes catch beauty out of the corner of our eyes. I had an experience of this sort recently when I visited the Nicholas Roerich Museum in New York City. Roerich was, well, a little mad. He thought his wife was a medium receiving messages from Tibetan lamas, telling her that Roerich was the next Dalai Lama. His paintings are in-your-face spiritual—enormous, vividly colored mountains in front of small, pious people. The other visitors to the museum were, well, not 100% sane either; one complained loudly that the museum stopped displaying Roerich's healing crystals. "I'm not like *those people*," I thought as I looked at the art. "How sentimental," I muttered. Then, while I was staring vacantly at another painting, my peripheral vision caught a shade of purple which took my breath away. I turned towards this work and found mountains hanging like sunset clouds in a space charged with meaning. *Sentimental? What an empty word!*

Second, many beautiful objects are praised for their "unity amid variety." The aesthetic experience of such items involves the interplay of narrow and wide focus, attentional shifts between, say, the brushstrokes and the composition. This suggests that a particular interaction between the dorsal and ventral systems may be at play here. Furthermore, the ventral system is often involved in acts of unification [Yates, 2015]. Bence Nanay's recent claim that one type of aesthetic experience involves attention which is distributed between many of an object's properties [Nanay, 2015] would also be supported by the involvement of the ventral system.

Finally, and most intriguingly, Kant and others' claim that aesthetic judgments are independent of concepts and purposes finds an echo in the bottom-up character of the ventral attention system. This connection is worth lingering over, since it lends support to one of the most controversial aspects of traditional aesthetics. Surely concepts can *aid* appreciation just as much as they can hinder it...

Kant's *Critique of Judgment* is the *locus classicus* of the notion that there is a tension between beauty and concepts. More precisely, Kant admits that there *is* a species of beauty called "adherent" beauty which involves the judgment that an object is a pleasing instance of a particular kind. However, beauty is in the first place *free* beauty, which is judged independently of concepts. That is, the *ground* of the judgment that an object is (freely) beautiful is one's disinterested pleasure and doesn't involve any concepts. While our experience of a beautiful object *may* involve concepts, we can't *infer* that the object is beautiful even partly on the basis of those concepts.

Kant therefore has a perfectly clear way of distinguishing between the aesthetically "good" and "bad" uses of concepts. This exact way isn't available to me because I put less weight than Kant on the *judgment* that something is beautiful. That is, I want to sidestep the question of whether particular objects are "really" beautiful, instead focusing on the question of which ways of experiencing objects are more *valuable* than others. Which ways of applying concepts, then, lead to more valuable experiences? In the next sections, I'll attempt to distinguish uses of concepts which get in the way of aesthetic experience from the ones which aid it.[18]

2.6 Concepts

2.6.1 Bad Concepts

In his book about Edvard Munch, [Knausgaard, 2019] *So Much Longing in So Little Space*, Norwegian writer Karl Ove Knausgaard describes a way in which concepts can blind us to beauty.

> Much of what we see, we see because we know it is there, often it is more a matter of recognition, of registering something which already exists within us. Names play an important role in this, so much of what we see is in the name; that is an apple tree, that is an elm, that is a cherry tree, that is a spruce.

You may think Knausgaard is overstating his case. I thought so too. After all, how am I supposed to view the world if *not* through concepts? At any rate, while I understood that *some* people were blinded by their concepts, *I* wasn't one of them. I was a painter; I *saw* that snow wasn't just "white" but also yellow, blue, purple, grey... I was an aesthete; *my* concept of "cherry tree" didn't get in the way of my rapture at its pink blossoms.

Then, during my meditation retreat, I was swept away by the beauty of tree branches in motion. I had never seen anything like it. Apparently, whenever I passed a tree on a windy day, my mind had been editing out the

[18] Note that my appeal to the two attentional systems has some affinity to one aspect of Kant's view: namely, the notion of the "free play of the faculties." Spelling out this connection in more detail lies beyond the scope of this essay.

visual particularities of moving branches, sending me only (or primarily) the conceptual knowledge that motion was in fact occurring. I *thought* I *saw* motion; in fact, 90% of what I "saw" was merely the concept "motion."

This experience taught me two things. First, my usual way of seeing is much more conceptual than I thought—and the conceptual "glasses" I typically wear are invisible to me. Knausgaard's claim that I was seeing primarily the name seemed intuitively, even obviously false—yet was true nonetheless. Second, less conceptual ways of seeing can be staggeringly richer and more aesthetically valuable than more conceptual ways.

And yet it would be a mistake for me to conclude that concepts always hinder appreciation. Art appreciation classes work precisely by introducing conceptual frameworks which aid perception. In fact, even the concept of "tree motion" which got me in so much aesthetic trouble can be put to such use; I carry it with me to remind me of the perceptual richness available on a windy day. Perhaps now you will too.

What's the difference between the (aesthetically) "good" and "bad" uses of concepts? In neuroscientific terms, the right type of concept-application may be the type that promotes a particular, "aesthetically right" type of interaction between the ventral and dorsal attention systems. In more practical terms, the good concepts increase accepting attention and the bad ones decrease it. For instance, the novice art-viewer without the concept "impressionism" may be at such a loss about what she is seeing that she is unable to maintain attention for more than a few seconds, or she may be so hung up on the idea that art ought to be photo-realistic that her acceptance is compromised. Here, the introduction of the concept can aid appreciation by aiding accepting attention. But in the case of my habitual use of "tree motion," it was the concept which was getting in the way of my attention, preventing me from casting more than a cursory glance at the moving branches.

In the following section, I'll discuss several ways by which concepts get in the way of accepting attention in more detail: by carrying desires, purposes, and expectations.

2.6.2 Blinding Purposes, Desires, Expectations

In a famous experiment [Simons and Chabris, 1999], subjects asked to count the number of passes of a basketball failed to notice a man in a gorilla suit walking across the court. When concepts carry purposes, they can blind us in just the same way. And I think we tend to underestimate just how many of our concepts carry subconscious purposes. My own experience of the tree branches is a case in point. I believe that one of the reasons I had failed to appreciate their motion had been that my concept of "tree" had a hidden purpose: "something to be painted." My mind had been editing out the motion in trees because I worked in a medium without a temporal dimension.

A closely related notion to purpose is that of desire. Some concepts involve or bring about desires which in turn hinder appreciation. For instance, we've seen that the concept of "weed" led Robert Wright to a desire that the plant be removed, a bias against its existence which blocked aesthetic pleasure.

A particular type of desire stands out here: the desire for or against aesthetic appreciation itself. *In Search of Lost Time* [Proust, 2003] provides a treasure trove of such examples; for instance, Proust's narrator Marcel builds so many delightful features into his concept of Balbec Cathedral or of the actress Burma that he sets himself up for disappointment. Similarly, Marion Milner's concept of "Cézanne" tends to bring about a damaging desire for appreciation [Milner, 2011]:

> I would have said: 'Here is a Cézanne, here is something one ought to like,' and I would have stood there trying to like it but becoming less and less sure what I felt about it.

Trying to like a painting is the opposite of the open, accepting attitude conducive to appreciation. My experience of Roerich's paintings was the converse of this: I labeled the paintings with the tag "sentimental," wanted *not* to appreciate them and so (temporarily) closed myself to them. In opposing ways, Milner and I were both closed to beauty by our snobbery.

In general, if we hold on to our aesthetic expectations—positive or negative—too tightly, we set ourselves up for disappointment. I believe that this fact is what led Mary Mothersill to claim that there are no laws of taste [Mothersill, 1984], i.e. no non-trivial generalizations of the form "If an object has property ϕ, then I will be (aesthetically) pleased by it."

I think there are at least *probabilistic* laws of taste. For instance, the fact that a meadow is full of flowers or that a concert is performed by the Mountain Goats increases my chances of taking pleasure in it. To claim otherwise is to make my aesthetic behavior entirely irrational: why would I purchase expensive concert tickets if the band *didn't* increase my chances of enjoyment?

And yet there is something extremely attractive about the claim that I can't know ahead of time that I'll like the concert, or even know that I have a 90% chance, say, of liking it. I think that the source of this attractiveness is a *pragmatic* truth in the vicinity of the epistemic one Mothersill is proposing. It *is* true that I'm likely to enjoy a performance by the Mountain Goats. But it's also true that as soon as I start *counting* on this regularity, I risk getting in the way of my own enjoyment. The most obviously beautiful things—sunsets, flowers, the sea on a sunny day—are precisely the ones it's easiest to see as clichéd, the ones with the most "picture-postcard garishness." There may only be a 10% chance that I'll dislike a Mountain Goats concert, but as soon as I start counting on the 90%, my chances dwindle.

To open yourself up to beauty, you have to adopt an attitude which "inhabits" uncertainty, even if it's just a 10% chance of disappointment. Seen from this perspective, Mothersill's epistemic claim is an exhortation to the sort of openness and acceptance I have been describing in this essay. Meditation is precisely the practice of inhabiting uncertainty. I start each meditation session by reminding myself that this session may be different than all the previous ones, that while overall I enjoy the practice and grow within it, this regularity can be broken during any individual session. This, I believe, is also the sort of openness that is required for aesthetic appreciation.

Marcel's concepts set his expectations too high; Knausgaard's concept of "elm" sets it too *low*. The concept is tagged with "uninteresting;" it has so little built into it that it stifles curiosity. Unlike in Marcel's case, the object isn't examined and found deficient relative to the concept—rather, the concept prevents all but the most cursory examination.[19]

It's possible to alternate between Marcel- and Knausgaard-like experiences. I glance at a cabbage, think "that's a cabbage," shrug my shoulders. Then, I see an Edward Weston photograph of a cabbage, am spellbound, and see the contents of my fridge with fresh eyes. But then I buy a book of Weston photographs, expect miracles, and see nothing but... Weston photographs. Then, my cabbage once again becomes "just a cabbage—just like in a Weston photograph."

2.6.3 Good Concepts

In the first two essays of this dissertation, I argued that mathematics is beautiful. Since we access mathematics through concepts, not the senses, I had better convince you that concept-application in mathematics is of the variety which doesn't typically get in the way of aesthetic experience.

Thankfully, that isn't too hard to do. Note that Knausgaard's concept of "elm" or my concept of "tree motion" got in the way of aesthetic experience by reducing the details of our *sense perceptions* of (moving) elms. In mathematics, such cases don't arise precisely *because* the discipline is non-perceptual.[20] Concepts get in the way of aesthetic appreciation when they prevent clear vision of the beautiful object; far from being blockers, concepts

[19]Note that Marcel, too, has Knausgaard-like experiences. ([Proust, 2003], Volume II).

> The names which designate things correspond invariably to an intellectual notion, alien to our true impressions, and compelling us to eliminate from them everything that is not in keeping with that notion.

[20]An exception may be the case of visual proofs, and indeed the detailed formal proofs accompanying those *can* block aesthetic appreciation by blocking clear perception of the visual aspect of the proof. (See section 4 of essay 1.)

are typically the only route to such clear vision in the case of mathematics.

While mathematical concepts don't get in the way of aesthetic appreciation merely by being concepts, concepts which carry desires or expectations *can* block appreciation in mathematics just as much as in other domains. For instance, in section 4.2 of essay 1, my desire not to have wasted effort in writing out a wrong interpretation of proof #2 of the Pythagorean Theorem briefly blocked my appreciation.

Aesthetic appreciation involves a special interaction of top-down and bottom-up attentional systems (which leans in the bottom-up direction). In the case of visible items, this may be glossed as an interaction between perceptual and conceptual seeing. This gloss is unavailable in the case of mathematics, but the top-down/bottom-up duality remains. Let me spell this out.

As I argued in "No Mathematics Without Beauty," many beautiful pieces of mathematics exhibit unity amid variety, and their experience involves awareness of the whole as well as of the parts. Even though the attention we pay to mathematics isn't *perceptual*, it appears to involve an interaction of focused attention and peripheral awareness just like what I have been describing above. Furthermore, the way the subconscious "waves" an attractive idea in front of the conscious (described in section 1.1 of chapter 1) is strikingly similar to experiences such as my perception of the Roerich painting. Peripheral awareness—of the contents of our mind instead of the contents of our visual field—appears to be at play here.

Finally, the most valuable aesthetic experiences in mathematics often involve the creation of a new concept, seeing an old thing in a new light.

Poincaré's realization that "the transformations [he] had used to define the Fuchsian functions were identical with those of non-Euclidean geometry" was of this sort, as was the invention of such core mathematical concepts as "group" and "topology."

The concepts that aid aesthetic experience enable acceptance—a freshness, a sense of seeing things with "new eyes." In Milner's words, beautiful things acquire a "gleam from the first day of creation." Speaking of the work of the painter Elstir, Proust's narrator reaches for a strikingly related metaphor:

> If God the Father had created things by naming them, it was by taking away their names or giving them other names that Elstir created them anew.[21]

A thing's new name can have all the same letters as the old one: "motion" turning to "motion," "sunset" to "sunset." What changes is how tightly we grip the aims, desires, and expectations associated with the name. What changes is whether, to borrow the words of novelist John Green, we make ourselves vulnerable to beauty.[22]

2.7 Conclusion

In my view, attention and acceptance lead to aesthetic appreciation. Distance is aesthetically important because it enables attention and acceptance, but there are cold types of distance which are aesthetically unhelpful. Similarly, desires can cause attention to be less than fully accepting, but the

[21] [Proust, 2003], Volume II.
[22] [Green, 2019]

notion of "disinterest" doesn't play the sort of core theoretical role for me that it does for Kant. In what sense, then, is my view a rehabilitation of d-theories?

I think we can reframe d-theories as roadmaps to the broad categories of of things which get in the way of aesthetic appreciation. Interest (desire for or against an object's existence), attachment, closeness (but also distance of the repressing kind), even concepts—all of these things really do get in the way of appreciation. Furthermore, these "blockers" are much more pervasive than we'd like to think.

If we think of d-theories in this way, we'll start to see that the "smallness" of the temple of beauty which they propose is only illusory. It's not that there are few beautiful things; it's that our (potential) blindness to beauty is staggering. There are countless ways to shield ourselves from beauty, and there is almost no beauty strong enough to reliably break through our shields.

Cataloging the blockers to aesthetic experience is important not just for the theoretical aim of figuring out what the d-theories get right about aesthetic experience, but also for the practical aim of living a fuller, more joyous life. As examples such as the tree in motion illustrate, we tend to underestimate the extent of our blindness to beauty, and without realizing it settle for merely partial aesthetic experiences. Unless we learn what shields we habitually put up, we won't be able to set them aside.

The realm of aesthetic pleasure has its characteristic disappointments. Sometimes, like Marcel, we set our expectations are too high; sometimes, like Knausgaard, we set them too low. Ironically, sometimes a beautiful item disappoints by being precisely how we imagined it: another way to describe Milner's "picture-postcard Mediterranean" is to call it "picture-perfect." At my first Mountain Goats concert, the songs sounded so precisely like what I had hoped for that, like Milner, I found myself "weary and bored."

But if aesthetics has its characteristic disappointments, it has an equally characteristic habit of rising out of the ashes of disappointment. The second time he watches Berma perform, Marcel realizes that the very simplicity which so disappointed him during the first performance had been Berma's greatest skill. As soon as Milner relinquishes her desire for appreciation, her "picture-postcard garishness" turns to a "gleam from the first day of creation." And as soon as I located the source of my disappointment during the Mountain Goats concert, I realized that in fact the concert wasn't "exactly how I expected;" I hadn't expected *disappointment*! Just like that, the music shifted from something out of a dream to beautiful reality. Half of the time, we can remove our "shields" simply by noticing that they are there.

Aesthetic pleasure is a shy creature which shuns expectations and has to be approached obliquely. I hope that these pages can serve as a manual for such an (accepting, attentive) approach.

Bibliography

[sta, 2010] (2010). What is the most elegant proof of the Pythagorean Theorem? https://math.stackexchange.com/q/803. Accessed 2019.

[Aigner and Zieglier, 2010] Aigner, M. and Zieglier, G. (2010). *Proofs From the Book*. Springer-Verlag Berlin Heidelberg.

[Austin, 2009] Austin, J. H. (2009). *Selfless Insight. Zen and the Meditative Transformations of Consciousness*. The MIT Press.

[Baumgarten, 2007] Baumgarten, A. G. (2007). *Aesthetica/Ästhetik*. Felix Meiner Verlag.

[Benacerraf, 1973] Benacerraf, P. (1973). Mathematical truth. *Journal of Philosophy*, 70(19):661–679.

[Binkley, 1977] Binkley, T. (1977). Piece: Contra aesthetics. *Journal of Aesthetics and Art Criticism*, 35(3):265–277.

[Blåsjö, 2018] Blåsjö, V. (2018). Mathematicians versus philosophers in recent work on mathematical beauty. *Journal of Humanistic Mathematics*, 8(1).

[Breitenbach, 2015] Breitenbach, A. (2015). Beauty in proofs: Kant on aesthetics in mathematics. *European Journal of Philosophy*, 23(4):955–977.

[Bullough, 1912] Bullough, E. (1912). 'Psychical Distance' as a Factor in Art and an Aesthetic Principle. *British Journal of Psychology, 1904-1920*, 5(2):87–118.

[Deikman, 1963] Deikman, A. (1963). Experimental meditation. *Journal of Nervous Mental Disorders*, 136:329–343.

[Detlefsen and Arana, 2011] Detlefsen, M. and Arana, A. (2011). Purity of methods. *Philosophers' Imprint*, 11(2).

[Dewey, 1934] Dewey, J. (1934). *Art as Experience*. G. Allen & Unwin.

[Dickie, 1964] Dickie, G. (1964). The Myth of the Aesthetic Attitude. *American Philosophical Quarterly*, 1(1):56–65.

[Doyle, 2016] Doyle, G. (2016). *Love Warrior: A Memoir*. Flatiron Books.

[Eisenstein, 1844] Eisenstein, G. (1844). Geometrischer Beweis des Fundamentaltheorems für die quadratischen Reste. *Journal Für reine u. angew. Math. 28*, 28:246–249.

[Farb et al., 2010] Farb, N., Anderson, A., Mayberg, H., Bean, J., McKeon, D., and Segal, Z. (2010). Minding one's emotions: Mindfulness training alters the neural expression of sadness. *Emotion (Washington, D.C.)*, 10:25–33.

[Feferman, 1996] Feferman, S. (1996). Kreisel's 'unwinding program'. In Odifreddi, P., editor, *Kreiseliana. About and Around Georg Kreisel*, pages 247–273. A K Peters.

[Gardner, 1973] Gardner, M. (1973). Mathematical Games. *Scientific American*.

[Gauss, 1808] Gauss, C. F. (1808). Theorematis Arithmetici Demonstratio Nova. *Commentationes Societatis Regise Scientiarum Collingensis*, 16.

[Giaquinto, 2005] Giaquinto, M. (2005). From symmetry perception to basic geometry. In Mancosu, P., editor, *Visualization, Explanation and Reasoning Styles in Mathematics*. Springer.

[Goldman, 1997] Goldman, J. (1997). *The Queen of Mathematics: A Historically Motivated Guide to Number Theory*. A K Peters/CRC Press.

[Green, 2019] Green, J. (2019). Capacity for wonder and sunsets. `https://www.wnycstudios.org/podcasts/anthropocene-reviewed/episodes/anthropocene-reviewed-capacity-wonder-and-sunsets`.

[Hardy, 1992] Hardy, G. (1992). *A Mathematician's Apology*. Cambridge University Press.

[Henri, 2007] Henri, R. (2007). *The Art Spirit*. Basic Books.

[Hume, 1757] Hume, D. (1757). *Four Dissertations*. New York: St. Augustine's Press.

[Hurwitz and Rudio, 1895] Hurwitz, A. and Rudio, F. (1895). Briefe von G. Eisenstein an M. Stern. *supplement to Z. Math. Phys.*, 40:169– 203.

[Hutcheson, 2008] Hutcheson, F. (2008). *An Inquiry Into the Original of Our Ideas of Beauty and Virtue: In Two Treatises*. Liberty Fund.

[Jiang et al., 2006] Jiang, Y., Costello, P., Fang, F., Huang, M., and He, S. (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences of the United States of America*, 103(45):17048–17052.

[John Dawson, 2006] John Dawson, J. (2006). Why Do Mathematicians Re-Prove Theorems? *Philosophia Mathematica*, 14(3):269–286.

[Kant, 2000] Kant, I. (2000). *Critique of the Power of Judgment*. Cambridge University Press.

[Kivy, 1980] Kivy, P. (1980). *The Corded Shell: Reflections on Musical Expression*. Princeton University Press.

[Knausgaard, 2019] Knausgaard, K. O. (2019). *So Much Longing in So Little Space. The Art of Edvard Munch*. Harvill Secker.

[Lange, 2016] Lange, M. (2016). *Because Without Cause: Non-Causal Explanations in Science and Mathematics*. Oxford University Press USA.

[Leddy, 1995] Leddy, T. (1995). Everyday Surface Aesthetic Qualities: "Neat," "Messy," "Clean," "Dirty". *Journal of Aesthetics and Art Criticism*, 53(3):259–268.

[Levinson, 2012] Levinson, J. (2012). Musical beauty. *Teorema: International Journal of Philosophy*, 31(3):127–135.

[Mancosu, 2008] Mancosu, P. (2008). *The Philosophy of Mathematical Practice*. Oxford University Press.

[Matravers, 1998] Matravers, D. (1998). *Art and Emotion*. Oxford University Press.

[McAllister, 1996] McAllister, J. W. (1996). *Beauty & Revolution in Science*. Cornell University Press.

[Milner, 2011] Milner, M. (2011). *A Life of One's Own*. Routledge.

[Moran, 1994] Moran, R. (1994). The expression of feeling in imagination. *Philosophical Review*, 103(1):75–106.

[Mothersill, 1984] Mothersill, M. (1984). *Beauty Restored*. Clarendon Press.

[Nanay, 2015] Nanay, B. (2015). Aesthetic attention. *Journal of Consciousness Studies*, 22(5-6):96–118.

[Nelsen, 1993] Nelsen, R. (1993). *Proofs Without Words: Exercises in Visual Thinking*. Classroom resource materials. Mathematical Association of America.

[Poincaré, 1910] Poincaré, H. (1910)). Mathematical Creation. *The Monist*, 20(3):321–335.

[Poincaré, 1914] Poincaré, H. (1914). *Science and Method*. Dover Publications.

[Proust, 2003] Proust, M. (2003). *In Search of Lost Time*. Modern Library.

[Reinhard C. Laubenbacher, 1994] Reinhard C. Laubenbacher, D. J. P. (1994). Gauss, Eisenstein, and the “third” proof of the Quadratic Reciprocity Theorem: Ein kleines Schauspiel. *Mathematical Intelligencer*, 16:67–72.

[Rota, 1997] Rota, G.-C. (1997). The phenomenology of mathematical beauty. *Synthese*, 111(2):171–182.

[Scarry, 2001] Scarry, E. (2001). *Dreaming by the Book*. Princeton University Press.

[Seed, 2019] Seed, J. (2019). *Disrupted Realism: Paintings for a Distracted World*. Schiffer.

[Shelley, 2017] Shelley, J. (2017). The concept of the aesthetic. In *Stanford Encyclopedia of Philosophy*.

[Simons and Chabris, 1999] Simons, D. J. and Chabris, C. F. (1999). Gorillas in our midst: Sustained inattentional blindness for dynamic events. *Perception*, 28(9):1059–1074. PMID: 10694957.

[Smith, 1929] Smith, D. E. (1929). *A Sourcebook In Mathematics*. McGraw Hill Book Company, Inc.

[Tappenden, 2008] Tappenden, J. (2008). Mathematical concepts and definitions. In Mancosu, P., editor, *The Philosophy of Mathematical Practice*, pages 256–275. Oxford University Press.

[Walton, 1990] Walton, K. L. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts.* Harvard University Press.

[Wikipedia contributors, 2020] Wikipedia contributors (2020). All along the watchtower — Wikipedia, the free encyclopedia. `https://en.wikipedia.org/w/index.php?title=All_Along_the_Watchtower&oldid=946009736`. [Online; accessed 14-April-2020].

[Wright, 2017] Wright, R. (2017). *Why Buddhism Is True.* Simon & Schuster.

[Yates, 2015] Yates, J. (2015). *The Mind Illuminated: A Complete Meditation Guide Integrating Buddhist Wisdom and Brain Science.* Dharma Treasure Press.

[Zangwill, 2001] Zangwill, N. (2001). *The Metaphysics of Beauty.* Cornell University Press.

[Zeidan et al., 2015] Zeidan, F., Emerson, N. M., Farris, S. R., Ray, J. N., Jung, Y., McHaffie, J. G., and Coghill, R. C. (2015). Mindfulness meditation-based pain relief employs different neural mechanisms than placebo and sham mindfulness meditation-induced analgesia. *Journal of Neuroscience*, 35(46):15307–15325.

[Ziff, 1979] Ziff, P. (1979). Anything Viewed. In Saarinen, E., Hilpinen, R., Niiniluoto, I., and Hintikka, M., editors, *Essays in Honour of Jaakko Hintikka*, volume 124. Springer, Dordrecht.