

The Drew Review is the annual research journal for the undergraduates of the Drew University College of Liberal Arts, publishing undergraduate research from the previous calendar year. Our mission is to showcase the intellectual vibrancy of the students of the CLA.

Currently there are six editors, who were selected by faculty nomination, application, and invitation. Published students can to apply for a position on the board, with the remaining positions filled by faculty nomination and student application with a writing sample.

Submissions to the *Review* require a faculty nomination. Students who believe their work is exceptional should approach their professor for a nomination. The *Drew Review* accepts papers of no more than twenty-five pages in October and February from the previous semester. This year we received fifty-one submissions and have selected ten for publication.

As we are a double-blind, peer-reviewed journal, all submissions should be sent as a Word document to the corresponding editors without naming the student author or professor for whom the essay was written in the body of the essay. The student author's name and paper title should be in the faculty nomination. Images and graphs will be published in black and white and must be compatible with Word. It is the author's responsibility to ensure that all images can be reproduced. All published essays will use in-text citations referencing a works cited bibliography, and the specific citation style is dependent on the discipline of the particular essay. Students can expect to be asked to make revisions prior to publication.

Editorial Board (2016-2017)

Alef Davis (CLA 2017), Copy Editor
English Major

Samantha Lacey (CLA 2017), Public Relations
Psychology Major

Alexander Slotkin (CLA 2017), Public Relations
Philosophy and English Double Major

Jared Sutton (CLA 2018), Junior Corresponding Editor
International Relations Major

Saif Yasin (CLA 2017), Senior Corresponding Editor
Biochemistry and Molecular Biology Major

Alexa Zbieranowski (CLA 2017), Treasurer
Art History and French Double Major

Gian Domenico Sarolli, Faculty Advisor
Department of Economics and Business

Hannah Wells, Faculty Advisor
Department of English

The Drew Review
10th Anniversary Edition
May 2017

Table of Contents

All is Good for Plato the Monist Alexander Slotkin (CLA 2017) <i>Philosophy and English Double Major</i> <i>Paper Nominated by Professor Erik Anderson</i>	5
Examining the Uses of Bacteria in Oil Spill Bioremediation Josephine Emanuelli (CLA 2019) <i>Environmental Science and Economics Double Major</i> <i>Paper Nominated by Professor Alan Rosan</i>	21
Flesh and Stone, Stone and Flesh: Michelangelo and Bernini's Pygmalionesque Transformations Taylor Tracy (CLA 2017) <i>English and Art History Double Major</i> <i>Paper Nominated by Professor Margaret Kuntz</i>	46
Self-injury and pain: Analysis of non-suicidal self-injury and the effects it has on the neurological processing of pain Darci Gautam (CLA 2019) <i>Psychology Major</i> <i>Paper Nominated by Professor Alan Rosan</i>	92
Mineral Dust Aerosols as Feedback Mechanism Drivers for Global Desertification Zoe Coates Fuentes (CLA 2018) <i>Chemistry Major</i> <i>Paper Nominated by Professor Timothy Carter</i>	111
What About Bambi? An Examination of the Evidential Problem of Evil Shira Mindy Newman (CLA 2016) <i>Philosophy and Biology Double Major</i> <i>Paper Nominated by Professor Erik Anderson</i>	127

Cross-Media Metrics: The Shift in Audience Measurement
Zarina Akbary (CLA 2019) 144
Biochemistry and Molecular Biology Major
Paper Nominated by Professor Alan Rosan

“Hi Grandma, it’s me!”: Taking advantage of the aging mind
Samantha Lacey (CLA 2017) 167
Psychology Major
Paper Nominated by Professor Patrick Dolan

Barbara Strozzi’s Feminine Influence on the Cantata in 17th-
Century Venice
Rachel Rubin (CLA 2017) 187
Music Major
Paper Nominated by Professor Leslie Sprout

On the Constructability of Numbers
Jeff Moorhead (CLA 2018) 202
Mathematics Major
Paper Nominated by Professor Seth Harris

Cover Art by Shayna Miller (CLA 2019)
Art Major

All is Good for Plato the Monistⁱ
Alexander Slotkin (CLA 2017)

Abstract

We go through life believing that everything that exists is material, but there is some indication that matter presupposes the presence of things that go beyond nature. How can hot iron be red in color without the property of *being red*? Does the property of *being red* disappear as iron cools and changes in color, or does *being red* continue to exist as an independent entity indirectly observed through red objects? According to Plato's Theory of Forms, there is a form or essence of everything (e.g., *being red* or iron-ness), all of which exists in an unchanging, immaterial, and non-temporal world commonly known as "Platonic heaven." This paper argues that Plato's controversial account of reality does not go far enough. By engaging with Platonic and Neo-Platonic thought, I maintain that Plato's metaphysical doctrine leads to the conclusion that the existence of more than one thing is only an *appearance* of reality. In truth, only one thing exists, the Form of the Good, challenging the conventional wisdom of Plato's philosophy and perhaps the nature of reality itself.

Introduction

“What is there?” and “What is it like?” are the two most important questions for anyone interested in the fundamental nature of reality. Many of us dogmatically believe that everything exists as matter, but common sense insists otherwise. Basic Euclidean geometry, for example, maintains that the area of a rectangle is equal to the product of its width and length. The equation, $A = wl$, is only true for perfect rectangles: figures with four 90° angles and two pairs of sides equal in length. Perfect rectangles, however, do not exist in nature, since even the most seemingly straight line is riddled with rivets at the microscopic level, distorting the rectangle’s angles and sides. If the equation had referred to a nonexistent object, it would be meaningless. But the equation is meaningful, indicating that a perfect rectangle exists *somewhere*. Without realizing it, we posit from a young age the existence of some immaterial, abstract object, complicating the question, “What is there?”

Given that we have some idea of what exists, we might be inclined to ask the second question, “What are the things that exist like?” Are all immaterial objects geometric? If so, how do non-physical objects, if at all, interact with physical objects? These questions, as well as the two questions from the previous paragraph, lie at the intersection of metaphysics and epistemology. Metaphysics is the branch of philosophy dedicated to the nature of existence or being and the world that encompasses it, while epistemology is the branch of philosophy concerned with knowledge, namely what can be known. Plato was one of the first thinkers in the Western philosophic tradition to bring these two branches of philosophy together into a comprehensive theory of existence, the Theory of Forms.

Forms, such as perfect rectangles, are the immaterial and non-temporal essences or indispensable qualities of everything (see *Phaedo* 74a-75d). Platonic forms are eternal, changeless, and exist independently of physical objects, which have their being and properties by “participating” in forms. That is to say, physical objects have their nature or properties by attempting to copy immaterial forms. Today, the Theory of Forms is largely seen as “a many in one [theory]. It is many because it contains many Ideas [or Forms]. It is one because these Ideas constitute a single organized system of Ideas” (Stace 79). In other words, Plato’s theory posits the existence of many forms or ideas in accordance with an organized system of existence headed by a Supreme Being. My aim in this paper is to show that Plato’s Theory of Forms posits the existence of only one thing and is in this way monistic, an idea that Plato himself contradicts: “And so, in the case of all things...[we] posit a single form belonging to each” (*Republic*, Plato 458). By arguing that this theory is monistic rather pluralistic, I argue that the only thing that exists is the Form of the Good.ⁱⁱ Hence, rather than the two realms of being (material and immaterial) that Plato believes exists, the Form of the Good is itself the only realm of being.

This paper draws heavily on Plato’s writings to support the claim that the Theory of Forms’ monistic nature is self-evident when taken to its logical end despite its creator writing otherwise. Firstly, I will set out Plato’s argument for the existence of forms. Then I will explain how his character “Socrates” differentiates knowledge from opinion or belief and the role of forms as objects of knowledge, followed by the Analogy of the Divided Line. Lastly, I will defend my thesis from its two strongest objections: the Argument of Contradictory Opposites and the Objection from Broadness.

The Existence of Forms

Suppose I have two sticks, A and B, and decide to see which is longer. When I put them side-by-side I notice they are equal in length. Now suppose I take stick A and compare it to a third stick, C, and find that A is longer than C:

A. _____
B. _____
C. _____

It is clear that sticks A and C are unequal in length, but now we are in a quagmire. Earlier we concluded that stick A has the property of equality, and now we are maintaining that it has the property of inequality. In other words, it was once an equal thing and is now an unequal thing. How can something become its opposite?

For Plato, this example suggests that we “believe in the existence of equality—not the equality of pieces of wood...but something beyond that—equality in the abstract” (*Phaedo*, Plato 120). When we reconsider the stick demonstration, it becomes clear that stick A must be participating in the idea of equality because equality itself never changes. Moreover, the idea or form of absolute equality must exist for us to determine that objects are equal in any respect. As the philosopher Plotinus would later explain, without forms there is no “non-arbitrary justification” for saying anything has one property over another (Gerson). But forms cannot exist as a physical entity perceivable by the senses. If it were, equality itself, like a stick, could take on the property of inequality. Additionally, equality and other forms cannot exist as physical entities because physical objects do not interact with them.ⁱⁱⁱ They are not only ideas either because, as the dialogue *Parmenides* makes clear, forms

would then not exist beyond our minds, making it impossible for objects to participate in them (*Parmenides*, Plato 926).^{iv}

Knowledge and Opinions or Beliefs

Physical objects may adopt properties and lose them. In fact, physical objects may have different properties to different people. While my senses may tell me that two tables are the same shade of brown, Socrates, if he could, might say the opposite. As Plato asks:

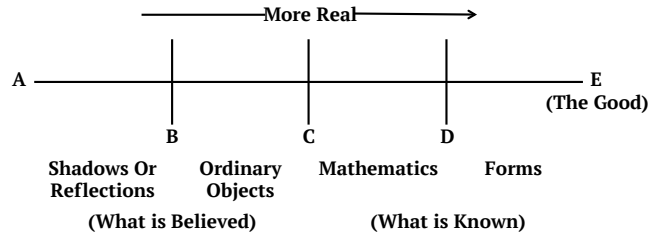
Do sight and hearing convey any real truth to men? Are not the very poets forever telling us that we neither hear nor see anything accurately? But if these senses of the body are not accurate or clear, the others will hardly be so, for they are less perfect than these, are they not (*Phaedo*, Plato 113)?

Because our senses may deceive us into having contradictory ideas of what is real and because physical objects are always changing, we cannot derive knowledge empirically. After all, knowledge must be eternal and true for everyone. Hence, we must have knowledge of forms to know anything, “otherwise we could not have compared equal sensible objects with abstract equality” (121).

Although forms are objects of knowledge, they are not intuitively understood because, to Plato, our sensory organs bewilder our souls into forgetting them. Reason, he believes, helps us remember what our souls knew before they transmigrated from the intelligible (not physical) world of forms into our bodies.^v For this reason, he champions rationalism, a theory of knowledge according to which knowledge is only achievable through reason devoid of sensory experience. The senses, on the other hand, only help us fabricate opinions.

The Analogy of the Divided Line

While Plato fleshes out his metaphysics and epistemology in *Phaedo*, one of his most important contributions to the Theory of Forms is found in the *Republic*. In this dialogue, he pens the Analogy of the Divided Line, diagramed below:



Segment AC is the physical or visible world, whereas CE is the intelligible world, such that “the division is in this ratio: as what is believed to what is known” (*Republic*, Plato 461). There are two divisions of belief or knowledge, the “lower” and “higher,” in each respective realm of being. AB (shadows or reflections) and BC (ordinary objects) remind us that the senses cannot be trusted. We have our first interaction with forms in segment CD, which concerns mathematics, a discipline that requires its practitioners to understand geometric forms, such as the form of a perfect right triangle. Lastly, DE represents knowledge of forms, which leads the knower to point E, the highest form or the Form of the Good.

Because knowledge is a good thing but not goodness itself and because all forms are objects of knowledge, the Form of the Good is the object that all other things aspire to be. Philosophers strive to understand this form because

What gives truth to the thing known and the power to know to the knower is the form of the good...Both knowledge and truth are beautiful things. But if you are to think correctly, you must think of the good as other and more beautiful than they are (460).

For Plato, to truly have knowledge or understand forms, we need to know the Form of the Good.

Thesis Defended

I. The Strongest Objections

Plato believed that the Form of the Good “illuminates” all other forms, allowing us to understand them just as the sun allows us to see the world around us (See *Republic* 507b-509c). All forms, he contends, participate in the Form of the Good but are separate from it in the same way that sunlight participates in sun-ness but exists separately from the sun (*Republic*, Plato 460). Philosophers supporting Plato’s Theory of Forms interpret the analogy as Plato had, in a one-in-many way. Plotinus, the founder of Neo-Platonism, a tradition of philosophic thought strongly influenced by Plato, wrote that the Good, which he calls the “One,” is the cause of all things in such a way that it is everything: “There is a certain necessity that the first [the One] should have its offspring, carrying onward much of its quality, in other words that there be something in its likeness” (Plotinus 813).^{vi} Plotinus’ language is abstract, but we can ground his explanation in everyday experience to make his point clearer. The One stands in relation to everything “the way

in which a properly functioning calculator may be said to contain all the answers to the questions that can be legitimately put to it” (Gerson). Just as the answers to the questions posed to the calculator depends on the calculator to exist but are not identical to it, everything in existence depends on the Good but is not identical to it. The analogy of the sun begins to break down when we study the relationship between forms and the Good, as well as the relationship between sunlight and the sun.

Unlike sunlight, which after being emitted no longer depends on the sun to exist—just as the truth-value of answers to algebraic equations (e.g., “ $|2x-1| = 5$ ”) do not depend on calculators—forms and the Good are inseparable. Plato tells us forms not only “owe their being known to the good, but their existence and being are also due to it” (*Republic*, Plato 460). If the Good disappeared, so would forms because forms participate in the Good. What grounds do we have for believing in a plurality of forms? None whatsoever. Rather, forms might be the means to interpret manifestations of the Good in seemingly contradictory instances. After all, it is easier to understand distinct forms of equality and inequality rather than a form that is both equality and inequality.

This line of thought should be applied to sensible or perceivable objects as well, since we can only conceive of ordinary objects because they are in the “likeness” of the Good. Just as the idea of *The Starry Night* depends on the original Vincent van Gogh painting, physical objects depend on the forms they participate in, such as the form of beauty itself. Because these objects depend on Plato’s forms to exist and because these forms are the Good, physical objects are by extension also the Good. If these

abstractions from Plato's own words are correct, we may surmise what he did not:

Proposition₁ All things are derived from the Good because all things participate in goodness.

Proposition₂ When we speak about forms or sensible objects, we are referring to the Good.

Proposition₃ Among competing hypotheses, the one with the fewest assumptions should be selected (Occam's Razor).

Conclusion Only one thing exists: the Form of the Good.

For clarification, the competing hypotheses mentioned in Proposition₃ are the Theory of Forms as a pluralistic theory (a one-in-many theory) requiring a dualistic world-view and positing forms for everything, and the Theory of Forms as a monistic theory, which only requires positing the Good. The controversial nature of my thesis has raised two considerable objections, which we may call the Argument of Contradictory Opposites and the Objection from Broadness.

The first objection takes issue with Proposition₁, beginning with the claim that nothing can become opposite to itself. As Plato explains in the *Phaedo*, snow cannot "ever receive heat and yet remain what it was, snow and hot: it will either retire or perish at the approach of heat" (*Phaedo*, Plato 141). Snow will melt if it takes on the opposite quality of heat. But if we agree that nothing can become opposite to itself, what should we say about benevolence and malevolence if they both

come from and depend on the Good? When answering this objection, we should refer to Parmenides, the world's first monist.

Parmenides contends that arguments for a pluralistic reality are self-defeating. Taking issue with the idea of non-being, which is entailed by a belief in opposites, he writes, "You could not know what is not—that cannot be done—nor indicate it. For the same thing is there both to be thought of and to be" (Parmenides 24). How, Parmenides asks, might we say that something can be non-being? It cannot be done. Therefore, we have no reason to believe in a pluralistic reality, in which case there are no contradictory opposites because only the Good exists.

While Parmenides's critique of pluralism is strong, rationalists may answer the same objection by simply blaming the senses for the appearance of contradictory opposites. It is worth noting that physicists define cold as the absence of heat. When we touch snow, our bodies transfer more heat to the snow than the snow transfers to our bodies, creating the appearance of "coldness." Malevolence, as well as one out of every contradictory pair, does not truly exist just as coldness does not exist; any quality that is opposed to goodness is only an absence of something.^{vii} The second objection to my thesis is the more serious of the two.

One may argue that, "Everything that exists is the Good," is too broad to mean anything substantial. How can we create an in-depth theory of the Good that not only explains its own existence but also the existence of different things, such as humanity and beauty themselves? The scope of the Good is too broad to be

defined or to account for every experience. We have no reason to subscribe to such a monistic theory.

I must admit that I cannot explain what it means to say that everything is the Good; such an endeavor would merit a book. The apparent ambiguity stems from my thesis's simplicity, however, not broadness. To argue that only one thing actually exists is the simplest way of explaining the existence of *everything*. Human beings do not naturally experience such a level of simplicity, which is why it is difficult to understand the implications of my thesis as anything other than an upheaval of our strongest beliefs about reality. But we have deduced the Good's existence and therefore have reason to believe it exists without fully understanding its nature as the only thing in existence. If this response does not overcome the objection that such a monistic reading of the Theory of Forms is not a viable metaphysical and epistemological theory, my position is not damaged in the least. I have still shown that monism is the logical end of this Platonic theory.

II. Parmenides & The Stranger

While I have shown that the two strongest objections to my proposal do not undermine my thesis, someone sympathetic to the Argument of Contradictory Opposites may object to my inclusion of Parmenides's earlier argument. These objectors would contend that Plato already answered Parmenides in the dialogue *Sophist*. In the interest of an informative discussion, I intend to answer these individuals here.

In *Sophist*, the character "Stranger" attempts to refute the idea that non-being is entailed by pluralism. Rather than saying that non-being exists, he or she maintains that only a difference between things exists, creating the

illusion of non-being (*Sophist*, Plato 1001). Hence, Parmenides's objection is answered. But all is not good for Stranger. We find a robust refutation of his or her argument in another of Plato's writings, *Parmenides*.

Parmenides is a unique dialogue in which Plato defends his Theory of Forms against Parmenides and his student Zeno's criticisms. While he did an excellent job of defending his thesis, he could not overcome Zeno's first objection. If many things exist, the argument goes, these things would be both like and unlike the things around them, which involves a contradiction (*Parmenides*, Plato 922). Sticks of equal length are not only like things in a pluralistic world, but are simultaneously unlike the shorter sticks around them.

Who could believe that something could be A and not A at the same time and in the same respect? We should not make the argument that only a difference between things exists because all physical objects are simultaneously like and unlike things in a pluralistic world. Plato concedes that Zeno's criticism is correct, adding, "When things which have a share in both characters ['alike' and 'unlike'], I see nothing strange in that" (923). While Plato may not find it strange, we have no reason to suppose that a pluralistic world exists, which would require believing that contradictory opposites exist simultaneously in physical objects.

We must accept that, in accordance with the Theory of Forms, only one thing exists: the Form of the Good, a weighty conclusion for anyone who believes in the existence of more than one thing. Plato may or may not have been a monist, but his doctrine, as I have argued, is monistic. My claim is controversial because philosophers and scholars succeeding Plato have interpreted the

Theory of Forms as a pluralistic doctrine with a monistic undertone. I hope that my argument, which I have explained and defended, will encourage scholars to rethink the standard interpretation of this doctrine if they have not done so already.

Notes

ⁱ Originally presented at the Moravian College Undergraduate Philosophy Conference in the spring of 2016.

ⁱⁱ Pluralism describes metaphysical theories that maintain the existence of more than one object, being, and or world.

ⁱⁱⁱ When I teach Plato's Theory of Forms to first-year students, I like to ask, "When was the last time you sat next to equality itself?"

^{iv} My usage of the word "idea" should not be confused with the common, alternative usage of the word to refer to Platonic forms. Whereas I am referring to our thoughts, the other usage refers to forms as ideas in the divine architect's (God's) mind. See 28a-29a of Plato's *Timaeus*.

^v Plato posits the existence of two worlds: the natural world and the intelligible world of forms. This intelligible world, he believes, houses forms as well as human souls, which leave the intelligible world to enter into our bodies. When they enter into our bodies, souls forget their knowledge of forms. Furthermore, our souls survive the death of the body, reentering the world of forms and thereby remembering the truth behind the appearance of what is real. See 69e-84b of the *Phaedo* to understand Plato's argument for the soul's immortality and its rightful place in the world of forms. Additionally, see 82b-85c of Plato's *Meno* for a demonstration of the soul's recollection of knowledge.

^{vi} See also book V 3. 15, 33 and VI 9. 5, 36 of Plotinus' book, *Enneads*, for other passages supporting Plato's interpretation of his Theory of Forms.

^{vii} Saint Augustine of Hippo makes a similar argument in his book, *Enchiridion: On Faith, Hope and Love*. Augustine contends that as far as reconciling God's goodness and the existence of evil, evil is not the opposite of good but rather the absence of it

(see chapter, “What is Called Evil in the Universe is But the Absence of Good”).

References

- Augustine of Hippo. *Enchiridion: On Faith, Hope and Love*. Regnery Publishing, Inc., 1996.
- Gerson, Lloyd. "Plotinus." *Stanford Encyclopedia of Philosophy*,
<https://plato.stanford.edu/entries/plotinus/#2>.
- Parmenides. "Fragments." *Ancient Philosophy: Volume I* 6th ed., edited by Forrest E. Baird, Prentice Hall, 2011, pp. 23-26.
- Plato. "Meno." *A Plato Reader: Eight Essential Dialogues*, edited by C.D.C. Reeve, Hackett Publishing Company, Inc., 2012, pp. 60-96.
- Plato. "Parmenides." *Plato: Collected Dialogues* 12th ed., edited by Edith Hamilton and Huntington Cairns, Princeton University Press, 1985, pp. 920-956.
- Plato. "Phaedo." *Ancient Philosophy: Volume I* 6th ed., edited by Forrest E. Baird, Prentice Hall, 2011, pp. 108-150.
- Plato. "Republic." *A Plato Reader: Eight Essential Dialogues*, edited by C.D.C. Reeve, Hackett Publishing Company, Inc., 2012, pp. 270-573.
- Plato. "Sophist." *Plato: Collected Dialogues* 12th ed., edited by Edith Hamilton and Huntington Cairns, Princeton University Press, 1985, pp. 957-1017.
- Plato. *Timaeus*. Translated by Peter Kalkavage, Focus Philosophical Library, 2001.

Plotinus. *The Six Enneads*. Translated by Stephen MacKenna and B.S. Page, Encyclopedia Britannica, 1952.

Stace, W.T. *The Philosophy of Hegel: A Systematic Exposition*. Dover Publications, 1955.

**Examining the Uses of Bacteria in Oil Spill
Bioremediation
Josephine Emanuelli (CLA 2019)**

Abstract

In a world that depends on petroleum, accidents can often occur when crude oil is being transported or extracted. These incidents are catastrophic for the environment and can cause devastation for decades to come, sometimes inflicting permanent damage on the environment. However, in some cases the environment is able to dispatch its own cleanup crew: hydrocarbon degrading bacteria. Recent improvements in gene sequencing and the creation of an international database of bacterial 16s rRNA sequences allow researchers to identify and better understand through DNA sequencing the roles specific bacteria play in hydrocarbon degradation. As a result, the versatility of these bacteria and their similarities across the globe from different oil spill sites has been revealed. Because of this, it should be possible to grow these bacteria in the lab and use them to remediate oil spills anywhere in the world regardless of whether or not the hydrocarbon degrading bacteria are native to that area. This application of bacteria in bioremediation could potentially save many ecosystems from permanent damage and save billions of dollars in cleanup costs.

Introduction

The United States and many countries around the world use petroleum oil as an important source of energy (Kelleher 2012). Drilling and extraction of oil takes place around the world both in terrestrial and marine environments (Kelleher 2012). Although petroleum is a common energy source, accidents occur regularly, with an average of 1.7 spills during transportation per year (“Oil Tanker Spill Statistics” 2012). When accidents occur in marine settings, oil can be released in very large quantities into the environment, whether in transport via tanker ships or during extraction on the rigs through mechanical failures or human error (Boesch 2012). These events, which can involve tens of millions of gallons of oil can result in catastrophic damage to marine environments for decades thereafter. Two head-line events, the Exxon Valdez oil spill off the coast of Alaska and the Deepwater Horizon oil rig explosion in the Gulf of Mexico have called attention to the methods being used to clean up oil spills and the ways that bioremediation can be used to aid with this clean-up. The Exxon Valdez spill occurred when an oil tanker ran aground on a reef off of Prince Edward Sound in Alaska in 1989, resulting in the release of millions of gallons of oil into the ocean. At the time, this was the largest oil spill in the United States, and the regulatory authorities were not prepared to deal with a spill of this size (Shigenaka 2014). Both spills were highly publicized and the cleanup efforts were under close scrutiny by industry professionals as well as the public. Both ExxonMobil and British Petroleum (BP), the two companies responsible for the Exxon Valdez and Deepwater Horizon spills respectively, put forth millions of dollars towards cleanup efforts. Despite the resources poured into the cleanup efforts of both spills, only the Deepwater Horizon spill cleanup was considered successful (King et

al 2015). A large part of that success was due to the microbes that were already present in the environment of the spill, creating an environment that was prepared to deal with the onslaught of oil (King et al 2015). Hydrocarbon degrading bacteria (HDB) proved to be essential to the clean-up process, as they consume the oil that was released in the spill. The Deepwater Horizon spill was a demonstration of how efficiently these bacteria are able to consume oil (King et al 2015).

The Exxon Valdez and Deepwater Horizon oil spills are considered to be the two worst spills to have occurred and both are regarded as unmitigated environmental disasters. However, the Exxon Valdez spill is considered to be a much greater tragedy than the Deepwater Horizon despite less oil being released into the environment during the ExxonValdez spill (Hazen et al. 2010). As displayed in the table below, the differences in the amount of oil that was released during the spill as well as the lasting impact of the spill are what makes these spills so catastrophic as well as interesting to compare.

Table 1. A comparison of the ExxonValdez and Deepwater Horizon oil spills

	Exxon Valdez	Deepwater Horizon
Date	March 24, 1989	April 20, 2010
Gallons of oil released (in millions)	11	140
Stage when incident occurred	Transport	Extraction
Results of Bioremediation	Considered to be unsuccessful after years	Successful after several months
Habitat Recovery Status	Not recovered (27 years later)	Almost completely (6 years later)
Presence of HDB	HDB were not present	HDB were already present
Time Span of Spill	Spill occurred over a few hours	Spill was continuous over several weeks

The recovery process from the Deepwater Horizon spill for the surrounding area in Alaska has been much slower than the recovery process of the environment in the Gulf of Mexico surrounding the Deepwater Horizon rig. This difference in recovery time can be credited largely to the HDBs that were present to promote clean-up after the Deepwater Horizon accident (Hazen et al. 2010). These HDBs were already present in the water of the Gulf of Mexico because of natural oil seeps on the ocean floor. Natural seeps, which are found around the world, occur

when oil from under the ocean floor is released slowly into the environment (Figure 1) (Lawrence 2010).

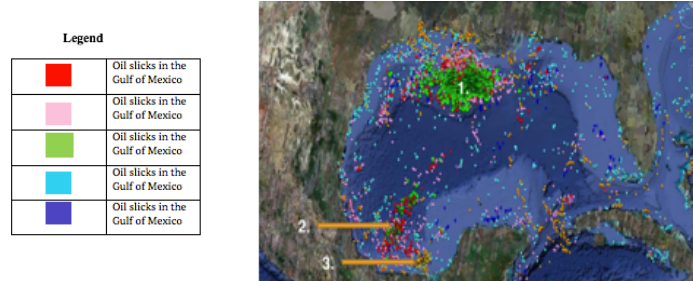


Figure 1. This map shows the location of natural oil seeps found on the floor of the Gulf of Mexico. Each dot represents a natural seep found in ocean floor. The naturally occurring HDB are likely found near these natural oil seeps found (Lawrence 2010).

As a result, the bacteria surrounding these natural seeps have evolved to use oil as a carbon source or nutrient, preventing accumulation of oil and damage to the surrounding ecosystem (Ziervogel et al. 2014). These bacteria live in the oil plumes that rise from the seeps, thereby, when the Deepwater Horizon spill occurred, the bacteria were able to move towards the oil plume coming from the oil rig (King et al 2015). Once the source was found, they began degrading the oil as it was being released from the site of the breach (Shigenaka 2014).

It is known that bacteria that degrade oil are present near natural oil seeps, but the Deepwater Horizon spill increased public as well as scientific interest in them and their use in bioremediation. HDB significantly reduced the damage to the surrounding environment. Although both spills are considered to be similar, as they are considered to be the two most ecologically devastating oil spills, their differences are what sets them apart (King et al 2015), (Hazen, et al. 2010). Compounding the problem of a lack of preparedness, the environment that the spill took place in was ill-prepared to deal with this spill as well. The HDB that are present in the Gulf of Mexico are not found in the Valdez Narrows, as there are no natural oil seeps in that area (Shigenaka 2014). If there had been natural oil seeps in the area, or if the spill had taken place today, the evidence suggests that the cleanup of the spill would be much faster and the environmental impact would be lessened (Ziervogel et al. 2014). The Deepwater Horizon spill occurred in 2010 when natural gas entered the oil drill, and a safety mechanism failed, resulting in an explosion that caused the oil flowing up into the rig to be released into the ocean (King et al 2015).

In this paper, the hydrocarbon bacteria, their functionality as well as their environmental impact will be discussed. The environmental factors that made the Exxon Valdez oil spill a much larger ecological catastrophe than the Deepwater Horizon spill will be addressed. Despite being an environmental disaster, the hydrocarbon degrading bacteria found in the Deepwater Horizon spill can be used in the future as an example of one of the most effective ways to clean up oil spills; this will serve to advantage of future cleanup efforts, if another incident like the Exxon Valdez accident were to occur.

Coupled with the advancements in sequencing techniques following the Deepwater Horizon Spill, there was an increase in the use of gene sequencing of HDB commensurate with the advancements in sequencing technology. The information gathered from the gene sequencing has helped researchers understand which organisms can be applied to future oil spills in order to decrease the level of devastation caused by these future catastrophes (McDonald et al. 2012). With that said, this paper will also explore possible bioremediation techniques such as growing the bacteria in the laboratory as a stock supply and then applying these bacteria to spills around the world to help aid with clean-up efforts. This suggestion also poses the question of what the negative impacts of these bacteria are when released into an oil spill, and if these adverse side effects can be lessened or even avoided when bioremediation is used.

Analyzing the Hydrocarbon Degrading Bacteria

Part of the reason that scientists are able to examine the genomes of bacteria so closely and compare genetic similarities is because of recent improvements in gene sequencing technology. When genes are sequenced, the 16s rRNA is the genetic sequence used. The 16s rRNA is the prokaryotic genetic sequence that is found in all species of bacteria, and codes for the part of the RNA that makes up the ribosome in the cell ("16s rRNA gene"). This sequence is used because it is different for all species of bacteria and remains well preserved throughout the extraction process ("16s rRNA gene"). This sequence appears only in bacteria and archea, as it is a section of prokaryotic DNA ("16s rRNA gene"). The ability to fully sequence a genome in less time is a recent advancement, rapidly is recent, and has resulted in the genomes of many of the species on Earth being fully or partially sequenced. As a result of this increase in sequencing and

related discovery, Greengenes (<http://www.greengenes.secondgenome.com/downloads>) has been compiling a database of the 16s rRNA gene sequencing information of many microbes, both bacteria and archaea. The 16s rRNA gene sequences are used to classify bacteria and this database allows researchers to match sequences of their subject microbes with sequences of microbes in the data base; this is beneficial for the discovery of evolutionary lineages as well as find taxa that are closely related (McDonald et al. 2012). The database helps researchers take the 16s rRNA sequences they had collected from organisms isolated from the oil plume and determine their genetic similarity to other previously sequenced bacteria. This helps to demonstrate how closely related the bacteria that responded to the Deepwater Horizon are, and how similar these bacteria are to other hydrocarbon degrading species that are found around the world.

Genetic similarities of the bacteria

Through analysis of the Deepwater Horizon spill, research showed that at many different levels or water depths of the oil plume, the most prominent family was the *Gammeobacteria* (King et al 2015). This family was detected in the oil seeps before the Deepwater Horizon rig explosion, but it was found that their population increased greatly once they were in the oil plume where the oil became plentiful; (King et al 2015). Despite being the most prominent family of bacteria in the oil plume, several other families of bacteria have hydrocarbon degrading capabilities and play a role in bioremediation of oil spills. Several families of HDB have been found to live in geographically remote and varied regions, such as northwestern Spain and the Gulf of Mexico or the Black Sea and Southeast Asia (Lamendella et al. 2014). Even though HDB are found living in such varied conditions

and environments, these bacteria are taxonomically very similar genomes based on the results from 16s rRNA sequencing. The bacteria have genetic similarities ranging from 87% to 99% (Lamendella et al. 2014), (Kostka et al. 2011) Different families of bacteria are even more closely related, as table 2 which shows the percentage similarity the genome based on the 16s rRNA sequencing of different orders of bacteria that are commonly found following oil spills (Lamendella et al. 2014). The percentage similarity to the closest match in the database shows the genetic similarity to previously sequenced hydrocarbon degrading bacteria from around the world. Table 3 further shows the percentage similarity the genome based on the 16s rRNA sequencing of different taxa of bacteria that are commonly found following oil spills (Kostka et al. 2012). The difference between the two tables is one deals with specific species of bacteria and the other one focuses on two of the main taxa found in the spills.

Table 2. Genetic similarities in different orders of bacteria

Order of Bacteria	% Similarity to closest match in databases	Where the samples were collected
<i>Psuedomonadales</i>	94.52-100	Sands from contaminated beaches on the Gulf of Mexico
<i>Vibrionales</i>	99.79-99.86	Sands from contaminated beaches on the Gulf of Mexico
<i>Alteromonadales</i>	98.32-99.93	Sands from contaminated beaches on the Gulf of Mexico
<i>Bacillales</i>	99.72	Sands from contaminated beaches on the Gulf of Mexico
<i>Chromatiales</i>	99.64	Sands from contaminated beaches on the Gulf of Mexico
<i>Oceanospirillales</i>	99.64	Sands from contaminated beaches on the Gulf of Mexico

Table 3. Genetic similarities of different taxa of bacteria

Taxa of Bacteria	% Similarity to closest match in databases	Where the samples were collected
<i>Pseudoalteromonas</i>	94.5	Oil on beaches surrounding the Gulf of Mexico
<i>Rhodobacterales</i>	99.8	Oil on beaches surrounding the Gulf of Mexico

Table 2 and Table 3 demonstrate the genetic similarities between several different orders of bacteria as well as the similarities that exist between different taxa of bacteria. These tables show that there are several different species of bacteria that are hydrocarbon degrading and perform the same functions, meaning that they are all genetically similar. Thereby, it is possible that all these species can be used in bioremediation as they all perform the same function and were all found on oil-contaminated beaches on the Gulf of Mexico (Kostka et al. 2012).

Gene expression and bacterial responses to the spill

In addition to having similar genomes, hydrocarbon degrading bacteria all utilize a unique metabolic pathway that allows them to process the chemicals that are present in oil as their carbon source for growth. Although they follow metabolic pathways similar to other strains of bacteria, the HDB have to go through an extra step when breaking down the oil before they reach their terminal oxidation state, the final step in their metabolizing of the oil (Dubinsky et al. 2013). This unique step that the bacteria take to process the oil is essential to

understanding how the bacteria are related, and without knowing the process the bacteria go through when degrading the oil, it would be impossible to create a supply of these bacteria for usage in oil spills around the world.

When breaking down the hydrocarbons, different genes in the bacteria are enriched. This was originally found in studies following the Deepwater Horizon spill, bacteria with specific genes overexpressed living in the oil plume from those living in the naturally occurring seeps (Shigenaka 2014). The genes that were overexpressed are the ones most likely to code for proteins to help the bacteria break down the oil and were activated as a result of the excess oil in the environment (King et al 2015). The genes that were enriched in the plume were also found using the same gene sequencing techniques used in taxonomy to identify similarities between microbial species based on 16s rRNA genes. In addition to enriching certain genes, the number of functional genes was also found to be almost ten times higher in some of the oil plume samples than in the non-oil plume samples (Shigenaka 2014). In the plume, the percentage of genes that were detected as functional had a widespread variation as well. Several different microbial groups were analyzed and the percentage of genes discovered to be functioning in the oil plume varied from as low as 2.2% in some species of bacteria to up to 20.5% of the genes functioning in other species (“Oil-degrading bacteria” 2014). This variation in percentage of functioning genes showed the diversity of the bacteria responding to the spill, and proved that despite this variation, they were still able to perform the same task.

Succession following the spill

When analyzing the diversity of the bacteria found in the spill, a trend of succession in the bacteria was found. The bacteria that were found in the oil plume also varied over time, as succession of the bacterial population took place as the flow rate of the oil changed (Lu et al 2012). Once the well was capped and the flow of oil ceased, the microbial community began to return to its pre-spill composition. It was established by research teams that the first responders to the spill, *Pseudomonas* and *Oceanospirillales* were found in high abundance when the flow of oil was unstoppable. However, as the flow rate and the marine environment began to change, the community of bacteria responding to the spill also started evolving.

When the amount of oil present began to decrease in early June, about one and a half months after the spill, the relative abundance of several taxa of bacteria increased. As a result, the abundance of *Pseudomonas* and *Oceanospirillales* decreased, and once there was a reduction in the rate of flow from the wellhead, levels of *Oceanospirillales* were not detected in several different samples. Later on, once partial capture of the oil began, there was an increase in abundance of *Methylomonas* bacteria, which are methane-oxidizing. These bacteria were found to be enriched within three days after the partial capture of oil from the wellhead began. This table shows the succession of three different taxa of bacteria following the Deepwater Horizon spill. The *Oceanospirillaceae* and *Colwellia* were two of the most abundant found following the spill, and their abundance in the plume increased by the fold listed in the table. The fold increase is a measure of how much the quantity of bacteria present changed from the initial value found in the area to the value found during the phase of the spill

listed at the top of the chart. The initial values are all one because that is the natural state. A twelve-fold increase means there was twelve times the number of bacteria found in samples collected than there were found to be initially. The charts are divided into three phases of the spill, the first one indicates a period when the oil flowed freely from the well and was not captured. Phase two shows the increases in abundance when the oil was being partially captured following a capping attempt. The final phase occurred when the well was completely capped and there was no more hydrocarbon flow. Values of zero show that there was no increase in bacteria detected (“Oil-degrading bacteria” 2014), (Lu et al. 2012). These data were taken from Figure 4 in Succession of hydrocarbon-degradation bacteria in the aftermath of the Deepwater Horizon oil spill in the Gulf of Mexico by Eric A. Dubinsky, Mark E. Conrad, Romy Chakraborty, Markus Bill, Sharon E. Borglin, James T. Hollibaugh, Olivia U. Mason, Yvette M. Piceno, Francine C. Reid, William T. Stringfellow, Lauren M. Tom, Terry C. Hazen and Gary L. Andersen.

Table 4. Abundance of different taxa of bacteria after the Deepwater Horizon spill

	0	I	II	III
<i>Oceanospirillaceae</i>	1	10-35	5-32	0
<i>Colwellia</i>	1	0-12	5-23	0-7
<i>Alteromonadaceae</i>	1	0	0	10-30

The rapidity of the response of the bacteria demonstrated how the environment that the spill took place in was prepared to deal with the onslaught of oil, and the environmental benefit of the bacteria that live around the natural oil seeps (King et al 2015). The rapid succession of bacteria based on the change in conditions was found

surrounding the oil plume before and after there was a change in the flow of oil (Lu et al. 2012). The structure of the microbial community completely changed with different dominant taxa of bacteria predominating at different times in the spills. This indicates a shift in the entire microbial community living in the oil plumes (Lu et al. 2012).

Despite the dramatic changes in bacterial composition in the plume, the bacterial communities found near the site of the spill both before the spill and after the oil seeps, were no longer detectable and had very similar compositions. They were more similar to each other than they were to any of the bacterial communities found in the oil plume during the time of the oil flow (Lu et al. 2012). This shows the resilience of the environment, as well as how easily it can rebound from the disaster and return to its previous state.

Growing the Bacteria in the Lab

The similarities of hydrocarbon degrading bacteria extend into their ability to be grown in an artificial environment. The HDB bacteria that are found around natural oil seeps are grown fairly easily in the lab, with half-lives ranging from 1.2 to 6.1 days in the oil plume. These long half-lives are despite the highly variable conditions that oceanic life presents, such as disturbances due to storms, predation and varying pH levels (Shigenaka 2014). The bacteria are able to grow on hydrocarbon as their primary carbon and energy source (Yergeau et al. 2015), as this is the carbon and energy source that is available around the oil seeps and in the oil plume itself during the spill (King et al 2015). The bacteria also need oxygen to power their metabolic reactions (Yergeau et al. 2015) and with those two factors it is possible that they will be able to survive and grow in

the laboratory environment. Due to their ability to grow and multiply quickly (King et al 2015), only a small supply would need to be kept in stock vials. Each taxa of bacteria should be kept in a separate vial, to avoid any competition that may occur over an extended period of time. In order for this endeavor to be successful, *Psuedomonales* and *Rhodobacterales* should be kept, along with *Psuedomonas*, *Gammeobacteria*, *Colwellia*, *Cycloclasticus*, *Pseudoalteromonas* and *Thalassosmonas* to ensure that HDB will be present at all stages of succession to aid with bioremediation. Due to the versatility of these bacteria, this method could be applied to locations where natural seeps are not present and the bacteria could be able to aid with bioremediation (Lu et al. 2012) and clean up the spill, such as where the Exxon Valdez spill occurred (Lamendella et al. 2014).

Due to the nature of oil drilling as a global business, it is essential that these bacteria be available to all countries if these HDB will be used as a bioremediation technique. The vials could be stored with the American Type Culture Collection (ATCC) to ensure timely access in case of an emergency. It is important that these bacteria be easily accessible, as they have the power to consume much of the oil spill. In addition, they should be used as quickly as possible to help with cleanup from the beginning, lessening the overall damage from the spill.

Once an environmental disaster occurs and the bacteria are needed, they would be grown to a larger population outside of their test tube environment. The different types of bacteria would need to be mixed together once their populations reach a sufficient size. Following that, the bacteria would be released at the site of the spill so that they can began consuming the oil and cleaning up the environment. This process can be used in oceanic

environments around the world, regardless of whether or not there are natural oil seeps near the site of the incident. These applications are possible because of the genetic similarities between the bacteria. The diversity of locations that the HDB are native to also make it possible for them to be successful in a wide variety of climates (Lamendella et al. 2014). In addition, in oil-related accidents around the world, the same families of bacteria respond, regardless of the climate or geographic location, indicating that these bacteria are capable of surviving anywhere in the world, not just near their oil seeps. It also does not matter if the incident occurred while the oil is in transport or when it is being drilled, and the mixture of bacteria will not need to be changed due to the versatility and adaptability of the bacteria.

Aftermath of Bioremediation Efforts

The bacteria that responded to the oil spill were very effective at cleaning up the oil in the least environmentally damaging way (Lu et al. 2012). The species were not an invasive species, and fewer additional chemicals needed to be introduced into the environment to aid with cleanup⁸. In a study done at the site of the spill one year after the spill, researchers found that there were very few differences in functional DNA of the bacteria found in the oil plume (Yergeau et al. 2015) showing that the genes that were over expressed in the plume were no longer being over expressed and the bacteria had returned to their normal functions. The *Oceanopirillales* taxa had previously been found at high frequency in the oil plume, but they were less than 3% of the bacterial community following the spill (Yergeau et al. 2015), showing that species composition was returning to pre-spill communities as well, exhibiting that there is likely an environmental dependence on these HDB.

The bacterial communities associated with hydrocarbon degradation and their functions returned to normal levels relatively quickly, however there were differences in oxygen concentrations and hydrocarbon degradation rates found in the plume area (Yergeau et al. 2015). Higher mineralization rates were discovered in the oil plume area than from the oil plume samples taken 35 km away (Yergeau et al. 2015). During the process of mineralization, a mineral is formed from the combination of a metal and another compound over time. When high rates of mineralization occur, it indicates that there is a higher presence of metals in the area. These high mineralization rates suggest that there are residuals of a higher potential of alkene degradation in the well area (Yergeau et al. 2015). Because alkene degradation was essential to the cleanup of Deepwater Horizon, these higher mineralization levels suggest that the site of the oil plume still has a unique microbial community, as alkene degradation was essential to the cleanup of the Deepwater Horizon. Although the bacteria in hydrocarbon degradation were found to return to their normal state after the spill, the high abundance of bacteria that degrade alkenes indicates that the microbial communities are still different 1.2 km away, where the sample was taken (Yergeau et al. 2015).

Furthermore, in this study, there were also differences in the oxygen concentration for the samples taken near the oil plume compared to those taken far away from the oil plume. Oxygen depletion rates of 30-50% were found as indicated by the high abundance of the *Nitrosopumilus* taxa (Yergeau et al. 2015). The bacteria that degrade the oil use aerobic reactions to break down the oil (Lu et al. 2012) resulting in large deficits of oxygen in the area surrounding the plume. Oxygen depletion can have impacts that are far reaching beyond the spill as many

aquatic organisms rely on oxygen to survive. Oxygen depletion is already an issue at the base of the Mississippi River, where a large “dead zone” has formed, killing the fish in the area and creating a toxic environment for humans with the presence of dead and decaying organisms. Oxygen depletion, or hypoxic conditions harm all levels of the ecosystem, and can result in a loss of productivity at all levels of the food chain, from plankton to larger predatory fish (Boetius et al. 2010).

Despite the adverse impacts of the bacteria and their aftermath, human cleanup efforts have proven to have even greater negative impacts. These impacts can be seen with the cleanup efforts of the Exxon Valdez oil spill. Despite best efforts and over \$2 billion spent, cleanup efforts were widely regarded as unsuccessful and it is believed that some of the ecosystems impacted may never recover (“Questions and Answers About the Spill”). The essential difference between the cleanup of the Exxon Valdez spill and the Deepwater Horizon was the presence of the hydrocarbon degrading bacteria. The Exxon Valdez did not have the bacteria present due to a lack of natural oil seeps in the area to aid with cleanup, and as a result the damage has extended for over a quarter of a century (“Questions and Answers About the Spill”). Despite the negative aftermath of the bacteria, their impact on the environment is much less than the impact of hand scrubbing otters and birds with dish soap in an effort to accelerate the cleanup process. Although the scrubbing of species is still likely to happen to try and clean up the spill as quickly as possible, using HDB in bioremediation will help lessen the impact of cleanup efforts.

Conclusion

Oil is a part of the oceanic ecosystem, and when it enters the environment naturally, as seen with the seeps in the Gulf of Mexico, it does not damage the surrounding ecosystem. As a result of the oil seeps found in oceans around the world, bacteria have evolved to consume and degrade the petroleum. Although these bacteria were known to the scientific community, the Deepwater Horizon oil spill demonstrated how powerful these bacteria are: they cleaned up almost half the oil released in the spill, significantly lessening the environmental impact of the spill (King et al 2015). Although the spill is widely regarded an environmental disaster, without the bacteria the spill would have had a much larger impact. This spill highlighted the capability of a natural phenomenon to aid with human-caused disasters. The Exxon Valdez serves as an example of the impact that oil spills can have on the environment when remediation is left to humans. To clean the spill, workers scrubbed animals and natural features to remove the oil from them. Chemicals were sprayed to disperse the oil so that it would be less concentrated in the area of the spill and tools were used to remove the oil from the surface of the water (Hazen et al. 2010). The two habitats impacted by the spills have had very different recovery periods; the Gulf of Mexico is considered to have recovered (King, et al 2015) and there are some species found in Alaska that may never recover (Hazen, et al. 2010).

Due to the power and effectiveness of these bacteria, as well as their wide geographic distribution, the evidence indicates it may be possible to grow the bacteria in a lab, and then apply them to spills around the world, regardless of whether or not there are natural oil seeps near where the spill occurred. Using these bacteria in bioremediation, the impact of oil spills can be lessened

and the cleanup can be done much faster so that the lingering impact of oil spills can be reduced. The bacteria that responded to the Deepwater Horizon spill have the potential to cleanup oil spills around the world, helping to reduce the impact of oil spills on struggling environments and lessen their catastrophic affects.

References

- “16s rRNA gene.” *Tutorial*. Greengenes. Web.
- Boesch Donald. “Deep-water drilling remains a risky business.” *Nature* vol. 484, no. 7394, 2012, pp. 484-289.
- Boetius, A, Janssen, F, Waldmann, C. “Oxygen monitoring in aquatic ecosystems,” *Earthzine* (2010). Online.
- Dubinsky, Eric, Conrad, Mark, Chakraborty, Romy, Bill, Markus, Borglin Sharon, Hollibaugh, James, Mason, Olivia, Piceno, Yvette, Reid, Francine, Stringfellow, William, Tom, Lauren, Hazen, Terry, Andersen, Gary. “Succession of hydrocarbon-degradation bacteria in the aftermath of the Deepwater Horizon oil spill in the Gulf of Mexico.” *Environmental Science & Technology* vol. 47, no. 19 2013, pp. 10860-10867.
- Hazen, Terry, Dubinsky, Eric, DeSantis, Todd, Andersen, Gary, Piceno, Yvette, Singh, Navjeet, Jansson, Janet, Probst, Alexander, Borglin, Sharon, Fortney, Julian, Stringfellow, William, Bill, Markus, Conrad, Mark, Tom, Lauren, Chavarria, Krystle, Alusi, Thana, Lamendella, Regina, Joyner, Dominique, Spier, Chelsea, Baelum, Jacob, Auer, Manfred, Zemla, Marcin, Chakraborty, Romy, Sonnenthal, Eric, D’haeseleer, Patrik, Holman, Hoi-Ying, Osman, Shariff, Lu, Zhenmei, Van Nostrand, Joy, Deng, Ye, Zhou, Jizhong, Mason, Olivia. “Deep-sea oil plume enriches indigenous oil-degrading

bacteria.” *Science* vol. 330, no. 6001, 2010, pp. 204-208.

Kelleher, Natalie B. *US Oil Sources Turmoil and Tranquility*. New York: Nova Science Publishers, 2012. Print.

King GM, Kostka JE, Hazen TC, Sobecky PA. “Microbial responses to the Deepwater Horizon oil spill: from coastal wetlands to the deep sea.” *Annual Review of Marine Science* vol. 7, 2015, pp. 377-401.

Kostka JE, Prakash O, Overholt WA, Green SJ, Freyer G, Canion A, Delgardio J, Norton N, Hazen TC, Huettel M. “Hydrocarbon-degrading bacteria and the bacterial community response in Gulf of Mexico beach sands impacted by the Deepwater Horizon oil spill.” *Applied and Environmental Microbiology* vol. 77, no. 22, 2011, pp. 7692-7974.

Lamendella R, Strutt S, Borglin S, Chakraborty R, Tas N, Mason OU, Hultman J, Prestat E, Hazen TC, Jansson JK. 2014. “Assessment of the Deepwater Horizon oil spill impact on Gulf coast microbial communities.” *Frontiers in Microbiology* (2014): doi: 10.3389/fmicb.2014.00130.

Lawrence, Geoff. *Oil Seepage an Inconvenient Truth (or the Perception of Risk in Oil Production)* TRIEC Limited. 14 May 2010. Web.

Lu Z, Deng Y, Van Nostrand JD, He Z, Voordeckers J, Zhou A, Lee YJ, Mason OU, Dubinsky EA, Chavarria KL, Tom LM, Fortney JL, Lamendalla R, Jansson JK, D’haeseleer PD, Hazen TC, Zhou

J. "Microbial gene functions enriched in the Deepwater Horizon deep-sea oil plume." *International Society for Microbial Ecology* vol. 6, no. 2, 2012, pp. 451-460.

McDonald D, Price MN, Goodrich J, Nowrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. "An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea." *International Society for Microbial Biology* vol. 6 no. 3, 2012, pp. 610-618.

"Oil-degrading bacteria." *Bioremediation- oil-degrading bacteria*. metamicrobe.com. 30 September 2014.

"Oil Tanker Spill Statistics." *ITOPF*. The International Tanker Owners Pollution Federation Limited. 2012.

"Questions and Answers About the Spill." *Exxon Valdez Oil Spill Trustee Council*. Alaska Resources Library & Information Services.

Shigenaka, Gary. "Twenty-five years after the Exxon Valdez oil spill: NOAA's scientific support, monitoring, and research." Seattle: NOAA Office of Response and Restoration (2014). 1-78.

Yergeau, Etinne, Maynard, Christine, Sanschargin, Sylvie, Champagne, Julie, Juck, David, Lee, Kenneth, Greer, Charles. "Microbial community composition, functions and activities in the Gulf of Mexico 1 year after the Deepwater Horizon accident." *Applied and Environmental*

Microbiology vol. 83, no. 8, 2015, pp. 5855-5866.

Ziervogel, Kai, D'souza, Nigel, Sweet, Julia, Yan, Beizhan, Passow, Uta. "Natural oil slicks fuel surface water microbial activities in the northern Gulf of Mexico." *Frontiers in Microbiology*, 2014, doi:10.3389/fmicb.2014.00188.

**Flesh and Stone, Stone and Flesh: Michelangelo and
Bernini's Pygmalionesque Transformations
Taylor Tracy (CLA 2017)**

Abstract

Tracy theorizes on new ways to understand sculptural works by Renaissance virtuoso Michelangelo Buonarroti (1475-1564) and prolific Baroque artist Gianlorenzo Bernini (1598-1680). This interdisciplinary essay repurposes Ovid's myth of Pygmalion's creation of a statue that comes to life as a useful interpretive lens for looking at Michelangelo's *Bacchus* and *Atlas* and Bernini's *Pluto and Proserpina* and Chigi Chapel niche statues of Mary Magdalen and Saint Jerome. Building on scholar Paul Barolsky's analysis of "Ovidian" sculpture, Tracy draws connections between Pygmalion's approach to sculpture and how Michelangelo and Bernini approached their own work. By exploring how passages from the poem can expose characteristics of these works, Tracy exposes how Michelangelo and Bernini's virtuosic abilities show transformation of form and content through illusionistic renditions of statues so naturalistic that they seem ready to emerge from a world of stone and enter into ours. Tracy also posits that viewers of sculpture serve a role parallel to the intervening figure of Venus in Ovid's Pygmalion myth, serving as the outside force who ultimately grants artworks with life behind their historical moment and ongoing interpretation.

Introduction

Transformation and illusionism are concepts that are inherent to the creation of art. In painting, sculpture or any other medium, an artist takes a subject and renders it with artifice, be it the artifice of suspended pigment on canvas, charcoal strokes across paper or chisel marks off a piece of marble. Transformation will refer to any change from one form or idea into another. Illusionism will be used to describe the representation of the human body or emotional expression through mimetic technique. By mimetic technique, I mean any visual strategy used to render the qualities of reality into artistic form. Illusionism will also be viewed as a type of transformation in this paper, the transformation from inanimate material into lifelike form through mimetic technique.ⁱ While seemingly universal across art, these themes operate differently in each artistic movement or period and within the evolving body of work of each artist. In the Renaissance, artists used these concepts to transform their materials, whether a two-dimensional surface or a block of stone into figures and scenes that encapsulated an orderly and logical notion of beauty. In the Baroque period, many artists extended these ideas and began to use the concepts of transformation and illusionism to create works that displayed a high sense of drama, emotion and movement in addition to idealized notions of beauty.

Ovid's *Metamorphoses*: Connecting Art and Literature

The ideas discussed above are especially prominent in Ovid's *Metamorphoses*, a poeticⁱⁱ narrative of interconnected myths that influenced Western art and writing beginning around the twelfth centuryⁱⁱⁱ and continue to influence art today (Barolsky "Florentine *Metamorphoses* of Ovid" 17). A Roman poet and

philosopher, Ovid combined elements of Roman and Greek mythology into the poetic saga recognized today as “the most complete and enchanting surviving guide to ancient myth” (Wilkins 383). Each myth in the *Metamorphoses* tells a different story. They are connected, as the title suggests, through the various types of transformation and all speak to the overall theme of love (Wilkins 385). The *Metamorphoses* inspired countless Ovidian^{iv} sculptures, paintings and poems that reworked its myths.^v This essay will focus specifically on Ovidian Renaissance and Baroque art and literature (Warwick 365).^{vi}

While it is difficult to discuss just one myth in the *Metamorphoses* because of their connectedness, this essay will concentrate on Pygmalion.^{vii} In the poem, Pygmalion the sculptor, disillusioned with the behavior of human women, carves the image of the perfect woman out of ivory. He falls in love with the sculpture, kisses it, gives it gifts and embraces it. Venus looks on during her feast day and hears Pygmalion’s prayer, granting his inner wish that his sculpture come alive. This particular myth’s themes of transformation and illusionism are paralleled in sculptures by the Renaissance artistic genius Michelangelo Buonarroti (1475-1564) and prolific Baroque artist Gianlorenzo Bernini (1598-1680).^{viii}

Pertinent passages from Ovid’s intense and richly poetic myth of Pygmalion offer a deeper understanding the expression of these themes (virtuosic carving, poetic transformation, illusionism and the role of the viewer in these ideas) and the simultaneously emotional and visceral impact of Renaissance and Baroque sculpture on the viewer.^{ix} I will argue in this essay that Ovid’s myth of Pygmalion’s creation of a statue that comes to life provides a useful interpretive lens for looking at the

selected works by Michelangelo and Bernini, exposing their virtuosic^x abilities to show transformation of form and content through illusionistic renditions of statues so naturalistic that they seem ready to emerge from a world of stone and enter into ours.^{xi}

Renaissance art and literary scholar Paul Barolsky has led the research of Ovidian influence on the work of artists beginning in the Renaissance.^{xii} Barolsky argues that poetic theory and Ovidian influences are deeply intertwined in the works of several Florentine Renaissance artists, including Michelangelo.^{xiii} Modern scholars like Barolsky emphasize the influence of Ovid's *Metamorphoses* on art, including Michelangelo and Bernini's works, but this is a topic that deserves further research. According to Barolsky, Ovid "played an extensive and deep role in the 'poetry' of Renaissance painting and sculpture" (Barolsky "As in Ovid, So in Renaissance Art" 451). Barolsky claims that "the very theory of Renaissance art, grounded in the concept of imitation, was often seen or described in terms of a central Ovidian fable, specifically the story of Pygmalion" (Barolsky "As in Ovid, So in Renaissance Art" 451). However, he admits elsewhere that we do not have a comprehensive understanding yet of all the ways in which Ovid influenced the art of the Renaissance (Barolsky "As in Ovid, So in Renaissance Art" 451).

Barolsky uses Lorenzo de' Medici's poetry as a gateway to understanding Ovidian influence on Florentine Renaissance art, including Michelangelo's *non finito*, which evokes the notion of working as a Pygmalionesque sculptor by releasing the soul of the figure from the rock (Barolsky "Florentine Metamorphoses of Ovid" 14-16). Elsewhere, Barolsky discusses how Pluto's hand

imprinted in Proserpina's thigh imitates Pygmalion, noting that this detail parallels the moment Pygmalion's hand sinks into his statue as if it were wax (Barolsky "Andromeda's Tears 25). Barolsky's connections between myth and artwork felt incomplete, leaving many of the questions raised by his analysis unanswered.^{xiv} Perhaps explaining the link between Pygmalion and works of art feels too self-evident because the myth is about the relationship between sculptor and his work of art. However, this lack of sufficient explanation in the existing scholarship opens opportunities to contribute new research on this topic. I will expand current research on Ovidian influence in the works of Renaissance and the Baroque artists in order to demonstrate how the Pygmalion myth elucidates key themes like illusionism and transformation in these artists' works.

Whatever the reason for the dearth of thorough discussion of this topic, I will affirm the Pygmalion myth's relevance to the work of two master artists: Michelangelo and Bernini. Additionally, I will add my voice to this evolving body of knowledge by using the myth to further explore transformation and illusionism in the artworks, which are also present in the myth of Pygmalion. Using a quote from the poem demonstrating each theme, I will analyze how the motifs in the poem are present in Michelangelo and Bernini's sculpture in order to establish a connection between the poem and the visual expression of these artists' sculpture. To be clear, this is not a paper about the connection between poetry and art. Instead, it is a paper that uses poetry influential during the time of a work's production as a lens through which to consider the art in new ways, by examining connections to the poem's imagery and the relationship between art and artist. It should also be noted that my methodology is not to produce a closed discussion

through an airtight argument, but rather to investigate these works through a series of close formal analyses that will raise new questions as I answer the ones at the heart of my argument. As such, my analysis may feel like it wanders through aspects of the works, and this method is intentional.

Returning to a discussion of the artists, Michelangelo and Bernini manipulated these key ideas of illusionism and transformation within their sculptures, drawing upon Ovidian influence. In particular, Michelangelo's early *Bacchus*^{xv} (Fig. 1) from 1496-7 and his later non-finito *Atlas*^{xvi} (Fig. 2) from 1530-4 demonstrate his adaptations of Pygmalionesque motifs of "bringing a statue to life" through transformation and illusionism. Drawing on the Classical ideals revived in the Renaissance and the emphasis on *disegno*, or the concept of a focus on line, drawing and design, these works by Michelangelo display virtuosity and several forms of transformation, including illusionism, with a focus on art theory and the perfection of depicting the male nude (Clifton 38).

Clear examples of Baroque art's most recognizable qualities, Bernini's works are emotionally expressive, combining twisting, visceral forms, a violent sense of movement and psychological intensity. Scholar Ann Wilkins cites Bernini as one of the artists "most profoundly inspired" by Ovid's myths, incorporating and transforming them in his artwork (Wilkins 383). I will discuss one of Bernini's sculptures originally intended for the Villa Borghese, his 1621-2 *Pluto and Proserpina*^{xvii} (Fig. 3), which demonstrates Pygmalionesque displays of virtuosic carving, illusionistic detail and poetic transformation. I will discuss two of his later Chigi Chapel niche statues, the *Saint Mary Magdalene* (Fig. 4) and *Saint Jerome*^{xviii} (Fig. 5) from 1661-3, with attention to

their demonstration of spiritual transformation and emotional expression. By choosing works from both earlier and later in Michelangelo and Bernini's careers, I will reveal that the inclusion of these Pygmalionesque qualities was not an isolated event that appeared in one work or one way, but instead a recurrent visual theme evident throughout their careers.^{xix}

Both sculptors manipulate the concepts of transformation and illusionism in what seem at first to be opposite ways. The subject of Bernini's *Pluto and Proserpina* comes directly from the *Metamorphoses*. While Bernini began working directly with physical manifestations of Ovidian myths such as the *Apollo and Daphne* or the *Pluto and Proserpina* early in his career, his later religious sculpture still expresses many of the same types of physical, poetic and illusionistic transformations, although it is not Ovidian in subject. On the other hand, Michelangelo transformed Ovid's ideas through physical manifestations of art theory (that is, how he approached the actual making of the object) instead of direct visual representations of the myths. At times, it is difficult to see how Ovid influenced his works, as the tangible and direct evidence of this link nonexistent in his sculpture, poetry or painting (Barolsky "Florentine Metamorphoses of Ovid" 14). Barolsky writes, "The relative absence of Ovid from Michelangelo's work should not put us off the scent of the poet's deeper meaning for Michelangelo's theory of art, as seen in his very notion of sculpture as metamorphosis" (Barolsky "Florentine Metamorphoses of Ovid" 14). A connection between Michelangelo and Ovid exists, but in a subtler way than in Bernini's sculpture. I will show how interpreting the visual expression of the selected works through a Pygmalionesque lens reveals that the transformation of stone into illusionistic subjects is

ultimately a display of virtuosic artistry, just like Pygmalion's magnum opus in ivory, his eventual wife, was a show of virtuosity for himself.

The Artistic Genius: Examining Virtuosity through Mimesis

In each of these works, Michelangelo and Bernini demonstrated their virtuosity as sculptors, demonstrating a level of skill that relies on their abilities to transform marble in a variety of ways. One of these skills is an aptitude to physically transform materials into soft flesh, hair, lips and tears. This sense of physical transformation appears in Ovid's poem as a manifestation of the sculpture's skill: "In sculpture exercis'd his happy skill; / And carv'd in ivory such a maid, so fair, / As Nature could not with his art compare" (Ovid 6-8). Pygmalion displays his virtuosity by creating a work of art so lifelike that it defies nature. He carves his statue with such attention to detail that his work seems to surpass nature and later causes him to become enamored with his own work. The demonstration of physical transformation of materials through craft is thus one parallel between the mythic poem and the works of art. This section will explore the insight Pygmalion's skill as a sculptor gives the viewer into the virtuosic abilities of Michelangelo and Bernini.^{xx} This passage of the poem detailing how Pygmalion created his statue and her visual qualities also exemplifies the idea of illusionism, the sense of lifelikeness and imitation of nature, intertwined with displays of virtuosity and physical transformation of material.

Like Ovid's mythic sculptor Pygmalion, Bernini and Michelangelo were masters at the "happy skill" of sculpting, using their abilities to create works pushing the limits of their materials, as working in ivory surely

must have tested Pygmalion's skills.^{xxi} The *Atlas*, *Bacchus*, *Pluto and Proserpina*, and the *Saint Jerome* and *Saint Mary Magdalen* are all marble statues. Both Michelangelo and Bernini worked largely in Carrara marble, the use of which is a display of virtuosity (Wallace 85-6). The natural flaws and cracks often hidden in a block of marble make it a notoriously difficult material to carve (Montagu 21). Michelangelo worked exclusively with a single block of marble as a show of his skill and process, and he controlled every aspect of acquiring the block he would use. As Montagu writes, Michelangelo's "close, almost mystical feeling for stone required him to choose the block himself and supervise the quarrying, and to work the statue from the first with his own hands" (Montagu 24). Michelangelo thus had an attachment to his materials, to the physical block of stone he would transform into a figure.

Bernini had a similarly close attachment to his sculptures, which Bolland discusses through reference to Ovid's Pygmalion. She discusses the possible influence of Pygmalion on sculptors when she writes, "Pygmalion's virtuosic artifice, which even seduces him into believing his own fiction, would doubtless make him a fitting mythic role model for any ambitious young sculptor" (Bolland 318). Bolland uses evidence from Bernini's son's biography of his father to suggest that Bernini conceived of the relationship between himself and his statues in an interesting manner. Citing two incidents in Domenico's, Bolland explains how Bernini was driven by a deep sense of adoration for his works (Bolland 318). Like Bolland, I would argue that this demonstrates a clear connection to the Pygmalion tale, but go even further to explicitly connect the ancient poem and the Baroque sculpture by positing that Bernini was a Pygmalionesque sculptor of his time who transformed his stone and pushed the limit

of his materials.^{xxii} Unlike Michelangelo, Bernini did not work exclusively with one block of marble. Instead, he would work a sculpture in separate pieces and then attach them in order to achieve greater displays of virtuosity. For example, from the works studied here, part of Proserpina's hair was worked separately and then attached later, demonstrating Bernini's ability to use sculptural techniques to achieve this striking display of artistic expertise and innovation with material (Cole 57).

Ovid does not mention how Pygmalion approached his block of ivory when carving his ideal maiden, but Michelangelo and Bernini's approaches to marble sculpture are also valuable to consider.^{xxiii} Michelangelo approached his single block of marble from the front and released the figure from the work as he moved backwards, giving him the room to make any necessary adjustments as he worked, as opposed to the traditional method of carving from four sides.^{xxiv}

In the Chigi niche statues of *Saint Jerome* and *Saint Mary Magdalen*, Chamberlain argues that based on her cross-sections of the works (Fig. 6), Bernini "conceived the frontal plane of the statue along the plane of the diagonal of the block," that although he carved them from the front, he intended them to be viewed on the corner along the blocks' diagonals (Chamberlain 80). This gave Bernini more freedom in the width of the expansive drapery around his figures, as the diagonal is longer than the longer side of the rectangular block of marble. As Chamberlain notes, this also allowed Bernini to work more efficiently (Chamberlain 80). Conceiving the sculpture in this way is thus a display of his carving expertise, as were Bernini's undercutting techniques. Bernini used elaborate undercutting, especially in the eyelids (Fig. 7) and the drapery, to maximize the impact

of light on his works in the chapel, where he planned the lighting to highlight features of his work (Chamberlain 82). Bernini cut away the base of his sculpture on an angle to further project these sculptures into the viewer's space (Fig. 8), augmenting the impact of the light and his efficient use of material (Chamberlain 80-2).

Furthering a discussion of the Chigi figures, the *Saint Jerome* and the *Saint Mary Magdalen* are carefully constructed and carved to achieve a seemingly precarious sense of balance that is structurally stable. Bernini took contemporary painting poses^{xxv} and turned them into marble poses that seemed impossible and unstable in sculpture (Chamberlain 71). About these poses, Chamberlain notes, "The poses of both statues convey tension brought about by a seemingly illogical distribution of weight" (Chamberlain 72). The seemingly unstable weights of the statues appear to be slipping to their right (the viewer's left), about to send the sculptures tumbling to the ground. Yet, the works are stable (Chamberlain 72). Thus, Bernini pushes the limits of sculptural poses by creating seemingly precarious compositions in stone, demonstrating his virtuosic abilities as a sculptor.

Bernini pushes the limits of composition and the tensile strength of the marble in the *Pluto and Proserpina* as well. Proserpina's arm extends into the air, her fingers splayed open and individually carved. This hand in and of itself is a dramatic display of virtuosity. One erroneous strike of the marble would have destroyed this fragile part of the work (Wallace 87). However, Bernini also physically manipulated the stone to show Proserpina's tears against her cheek, her leg jutting out as she struggles against Pluto and the individual curls in both Proserpina's hair and Pluto's beard (Barolsky "Andromeda's Tears" 25-6).

The execution of these highly detailed elements in the marble exemplifies Bernini's virtuosic abilities, skills in the same vein as Pygmalion's carving of an ideal ivory maiden.

While Bernini skillfully pushes the limits of material in the Chigi statues and the *Pluto and Proserpina*, Michelangelo's works also warrant analysis of the virtuosic craft demonstrated by Pygmalion in Ovid's poem. Barolsky discusses some of the formal qualities of Michelangelo's *non finito*, but this research can be expanded and then applied to other works like the *Bacchus*. Barolsky writes that Michelangelo's *non finito*, "although only a hollow persona, is marvelously rendered with so much vivacity that it seems, paradoxically, a vital, living presence. Part of the wonder of this paradox of art lies in Michelangelo's display of *poesis* or its "making," in the metamorphosis before our eyes" (Barolsky "Florentine Metamorphoses of Ovid" 14). The language that Barolsky uses here to discuss Michelangelo's sculpture parallels the language Ovid uses to describe Pygmalion's creation of a lifelike work. However, Barolsky does not follow up this discussion with a continued analysis of how the work demonstrates this sense of "making." This sense of craft is also present in the *Bacchus*. The *Bacchus*, while not idealized in the same extent as ancient depictions of the god, still demonstrates virtuosity in carving (Lieberman 66-7). The smoothness of the god's body and the subtle indications of his musculature show the marks of a hand with the same level of talent of imitating nature and not merely copying it like Pygmalion in the poem. The depiction of the clavicle and the tendons of the neck and shoulder in particular are skillfully crafted, displaying Michelangelo's extensive knowledge of anatomy.^{xxvi}

Returning to a discussion of the Chigi statues and the *Pluto and Proserpina*, these works also demonstrate virtuosic attention to anatomic detail. Bernini's works, in comparison to Michelangelo's bulky *Atlas* and feminized *Bacchus*, display anatomical accuracy, but are slightly distorted. Upon closer inspection, one can see that Bernini rendered each aspect of Mary Magdalen's anatomy supplely, from the muscles in her arms to the bones and tendons of her ankles. Bernini carved her hands clasped together, brought to the side of her face as her long, wavy hair tumbles over her shoulder, around her left breast and down to her waist. In the *Saint Jerome*, Bernini rendered the saint's signs of age, from the shadows showing the ribs under Jerome's thin skin, the sagging muscles and veins in his arm and the contorted muscles in his face as he repents, gently clasping the cross and bringing it close to his body.

In the *Pluto and Proserpina*, made earlier in his career, Bernini captured anatomical details in both Pluto and Proserpina through his virtuosic physical transformation of the marble from stone into illusionistic flesh. In *Pluto*, Bernini renders the god's muscular and angular physique down to the individual muscles in his abdomen as well as the veins accentuated in his twisted form. In *Proserpina*, Bernini captures the smooth, curvilinear lines of her arms and legs in addition to details as minute as her toenails and the flexing of her toes as she tries to wriggle free from the god. Both of these works, specifically their materials and the processes in which they were created, demonstrate how Michelangelo and Bernini acted as Pygmalionesque sculptors, rendering the physical transformation the stone into flesh through virtuosic carving, or, in the case of the *Atlas*, flesh hidden as stone.

Carving Passion: The Poeticization of Form and Emotional Transformation

Not only did Michelangelo and Bernini physically transform stone with their virtuosic sculpting abilities, but they also poeticized their sculptures by endowing them with a range of poetic meanings. These works depict a diverse range of emotional transformations. This poeticization is evident in the Pygmalion myth when Ovid writes, “Soft, and more soft at ev’ry touch it grew; / Like pliant wax, when chasing hands reduce / The former mass to form, and frame for use” (Ovid 84-6). Pygmalion’s statue transforms into flesh and the role of his sculpture is poeticized, meaning that it takes on this imaginative quality inspired by the form of the poetry itself. The craft of turning a stone into a figure is poeticized into the stone becoming soft and alive under one’s chisel and tools, reflecting theory and a level of skill inherent in the works of Michelangelo and Bernini.

Michelangelo’s approach to a block of marble is intensely poetic.^{xxvii} Barolsky explains how the *Captives* have been described as an artistic representation of Michelangelo’s view of himself, a soul trapped in poetic layers of “Earthly prisons,” derived from Michelangelo’s Neo-Platonist beliefs, which have been applied to discuss Michelangelo’s views of the self-made physical reality through his liberation of the figure from the stone (Barolsky “As in Ovid, So in Renaissance Art” 462). In this way, Barolsky explains that Michelangelo’s “Ovidian sense of carving as itself the poetic ‘metamorphosis’ of stone” (Barolsky “As in Ovid, So in Renaissance Art” 463). These ideas can be seen in the *Atlas*, one of Michelangelo’s *Captives*. The figure is sketched out in the stone enough for the viewer to tell that it’s meant to be a person, hunched over, his arm reaching up to his head. The legs appear to be stuck in the stone. The elbow juts

out, partially finished, but the hand is still encased in the rough rock. The work is thus endowed with an emotionally charged dynamism that makes it seem like the figure is wriggling free, although it is still trapped in the stone. Looking at the statue, it appears to be in transition from a rough, undefined state into a refined and smoothly polished one. The *non finito* work appears to be in a transition, in the midst of being worked on.

Though only partially finished, one can almost imagine what the ultimate outcome will be once the soul within the rock is fully set free. Michelangelo's poetry gives further insight into his philosophy behind the *non finito*. In Creighton Gilbert's analysis of Michelangelo's *non finito*, he notes that Michelangelo uses the act of sculpting as an extended metaphor for life in his writing, using the *non finito* as a symbol of man's limitations on Earth without divine intervention (Gilber 62-3). On Earth, man can only be released so far. Thus, the sense of Ovidian transformation is clear in this case. The construction of Michelangelo's partially finished male figure echoes the actions of Pygmalion in Ovid, bringing the sculpture to life by visually representing these poetic and philosophical transformations of the soul.

Similar to the clear sense of poetic transformation of the soul expressed in Michelangelo's *Atlas*, Bernini created visual representations of poetic transformations of the soul in his religious Chigi statues. The *Saint Jerome* and the *Saint Mary Magdalen*, the two niche statues by Bernini, demonstrate a poeticized sense of spiritual transformation. They complement each other through external and internal repentance to mirror the sins they are repenting to achieve salvation. In his book on Bernini, Howard Hibbard acknowledges this complementary nature when he writes, "The *Magdalen* is the traditional

penitential figure of grief and remorse who communicates with her God in an agony of physical emotion. *Jerome*, whose penance is for intellectual rather than physical sins, turns inward to a mystical experience” (Hibbard 194-5). Thus, these sculpted images of the penitents express their spiritual transformations from sinner to saint through the repentance of their sins. Expanding the formal analysis of these works begun earlier in the paper, the *Magdalen* and *Jerome* counter each other to depict two types of spiritual transformation (Hibbard 194-5). The *Magdalen* looks up and out with her open eyes, her head tilted onto her hands. Meanwhile, the *Jerome* curls in towards himself with closed eyes, eternally focused on his internal reflection and repentance to reach a state of salvation for one’s earthly sins.

Continuing this discussion of the expression of poetic transformation in Bernini’s work, the *Pluto and Proserpina* features this theme on multiple levels. First, the sculpture depicts the translation of a narrative from text to three-dimensional artwork. While Barolsky is discussing the *Apollo and Daphne* here, it applies to the *Pluto and Proserpina* as well, as both works depict myths from Ovid’s *Metamorphoses*. Barolsky notes that Bernini “transformed Ovid’s poetic version of their story in words into a marble image. The transformation of the verbal into the visual is an artistic metamorphosis, art as metamorphosis of one medium into another” (Barolsky “Ovid, Bernini, and the Art of Petrification” 157). By working from a source text, although not as explicitly influenced by the poem in the case of the *Pluto and Proserpina*, Bernini transformed the stone in order to give visual form and physical manifestation to the poem, a transition between media similar to Pygmalion’s rendering of the concept of an ideal woman into

stone.^{xxviii}

In addition to being a visual representation of a textual myth, the Ovidian content of the myth itself reveals a sense of poetic transformation. The sculpture depicts the rape of Proserpina by Pluto, the god of the underworld, who dragged the goddess into the underworld to become his queen. Her mother wandered the land, creating famines and destruction on Earth, forcing Jupiter to mediate. An agreement was reached where Proserpina would spend half the year on Earth and then return to the underworld for the other half of the year. This cycle came to represent the growth and harvest of corn, and thus the cyclical changing of the seasons (Cottrell). Thus, the rape of Proserpina can be visually interpreted as a representation of the end of the harvest, the transition between fall and winter when she would have to return to the underworld. Because of this, the rape of Proserpina has taken on a third poetic association, the transition between life and death itself, represented by Proserpina's desperate attempts to escape her fate (Wilkins 400).^{xxix}

Like the *Pluto and Proserpina*, Michelangelo's *Bacchus* plays with the visual representation of the gods. In his discussion of the sculpture, Lieberman explains how early depictions of the Bacchus in art have been classical, ordered and refined, omitting any sense of drunken revelry from their depictions of the pagan god of wine and revelry (Lieberman 66-7). Thus, while Bernini provides a visually striking and virtuosic depiction of a myth demonstrating the transformation of the seasons that represents the transformation from myth to image, Michelangelo provides a transformation of a traditional depiction of a god into something new and radically humanized, bringing the discussion to a new type of transformation: illusionism.

Illusionism fuses the sense of physical transformation through virtuosic craft and emotional expressiveness to represent a sense of lifelikeness that is inherently poetic. Ovid expresses a sense of illusionism in the poem when he writes, "One wou'd have thought she cou'd have stirr'd, but strove / With modesty, and was asham'd to move. / Art hid with art, so well perform'd the cheat, / It caught the carver with his own deceit" (Ovid 15-18). The artful lifelikeness of the statue tricks Pygmalion into believing it is real. His statue is so lifelike that the product of the artist's own hand tricks its maker. Also, the statue here is depicted as possessing emotional complexity, evoked by the idea that she could move but is modest and would be ashamed. The introduction of this psychological quality with the virtuosic execution of her physical qualities demonstrates the theme of illusionism, seen in the artworks.

The *Bacchus* is the key example of Michelangelo's display of illusionism, of physical and psychological likeness fused in a single work. Simultaneously drunk and stately, the *Bacchus* is a testament to Michelangelo's capturing of a spectrum of human emotion. He goes through its poses, showing that as one walks around the work, its subject alternates between appearing soberly heroic and sloppily drunk (Lieberman 67-9). While certainly jarring and disorienting to its intended Renaissance audience, so much so that the work was hidden away, today the viewer can fully appreciate the work's illusionistic qualities in the way that the works transform before our very eyes. With illusionism comes a sense of individualism in the particulars of representation, raising questions about whether humanism is perhaps inherent in the work and inherent in the logic of the Pygmalion myth itself. Lieberman writes, "Bacchus appears in a constant

transition from a stable pose to an unstable one and back again, losing his balance momentarily and then regaining it. As we move around him we see him sway, as anyone fuddled with too much drink would do” (Lieberman 72). In this way, the *Bacchus* is a humanized, illusionistic depiction of the god, transformed out of the stone. The *Bacchus* appears in a way that reflects anyone who has been drunk and thus raises valuable questions about humanist theory in the work and individuality at this time, which could also be reflected in the Pygmalion myth that warrant further investigation beyond the scope of this paper.

In the *Atlas*, illusionism can be seen in the emotionally charged quality of the work that results from the fusion of the virtuosic craft rendering the features of the human body in the partially completed work with the psychological quality of its incompleteness. The sense of lifelikeness comes from the way in which the work, although a roughly sketched figure, seems to come alive. The viewer can imagine the head still trapped in the upper portion of unworked stone and the emergence of the arm that then disappears again into the rough rock, creating visual and emotional tension when looking at the work. Thus, while perhaps not as explicitly as a work like the *Bacchus*, the *Atlas* also displays some of the illusionistic qualities of the Pygmalion myth, where the sculpture is poetically frozen in shame, like Michelangelo’s *non finito* figure is frozen in the unworked ethereal rough surrounding it.

In Bernini’s works, both pairs of statues include a wealth of illusionistic detail. The *Magdalen* looks up in grief, sorrow and repentance, shadows cast over her face as she leans on her beautifully carved and lifelike hands. As she looks up, representing her soul’s transformation from

sinner to forgiven, she does so in a way that that is visually expressive. The virtuosic rendering of her twisting pose and the contemplative expression exaggerated by the slight parting of her carefully carved lips blend to create an overall illusionistic figure of somber repentance.

The *Saint Jerome* displays an equal demonstration of virtuosity in rendering the aspects of the body in a way that transforms the marble from cold stone into illusionistic flesh. The physical transformation of the marble into the *Jerome's* flesh, along with his pose towards the cross, creates a psychologically poignant figure caught in religious contemplation. In these two sculptures, the drapery adds another layer of emotional expressiveness to the already passionate sculptures. The light creates strong shadows and soft gradations of light across the forms,^{xxx} which heightens the emotional expression and illusionistic rendering of forms in the work, mimicking the lifelike qualities seen in the poem.

In addition to his Chigi Chapel niche statues, Bernini also rendered the figures of the *Pluto and Proserpina* with emotionally expressive lifelikeness. When scholars discuss the illusionistic properties of this work, they cite Pluto's hand digging into Proserpina's thigh. For example, Barolsky writes, "For all the difference between the violence of Pluto's attraction to Persephone and Pygmalion's more gentle desire for his beloved, the illusion of a hand sinking into marble marvelously metamorphosed into seemingly soft skin makes of Bernini a modern Pygmalion" (Barolsky "Andromeda's Tears" 25). This is certainly an illusionistic detail, a poetic moment in the statue's narrative perfectly paralleling the moment where Pygmalion's statue comes to life. Scholars also often discuss Proserpina's tears

carved delicately onto her face (Barolsky “Andromeda’s Tears” 25).

Yet while these details are beautiful and poetic, there is a wealth of other illusionistic elements in the work. A closer formal analysis of the sculpture illuminates observations like Warwick’s even further, truly painting Bernini as a Pygmalionesque sculptor of his time. Warwick writes, “*Pluto’s* mature, muscular frame covered with a sagging skin provides a foil to the sweetness of her youthful body, firm and resistant, yet yielding in its corporeal softness” (Warwick 365). The combination of the physical forms like *Proserpina’s* outstretched hand, her wriggling pose and her hair flying backwards as she tries to escape fuses with her fearful cry and expression to create a single transformation from rough stone into a terrified girl. Bernini’s depiction of *Pluto* also reflects the transformation of stone into a physically bulky god, demonstrating his simultaneous annoyance and calm confidence in his strength as *Proserpina* pushes him so forcefully that the corner of his eye is pulled in the direction of her hand. Each of the works thus demonstrates an interest in illusionistic detail that fuses virtuosic skill of physical transformation with psychological intensity, so strong that it captivates the eye of the viewer, a necessary agent in the interpretation of these works.

Eye of the Beholder: The Viewer as Venus^{xxx}

In Ovid’s Pygmalion myth, Venus, the goddess of love and beauty, views Pygmalion’s offering at her shrine and grants life to his statue. In the poem, Ovid writes, “Almighty Gods, if all we mortals want, / If all we can require, be yours to grant; / Make this fair statue mine, he wou’d have said, / But chang’d his words for shame” (Ovid 63-6). This moment, from when Pygmalion visits the

shrine of Venus to say his prayer as he gives his offering, demonstrates that the artist does not have the power within himself to bring his statues to life. In the poem, it takes an outsider who views Pygmalion's struggle to bestow life upon his statue. One way to reinterpret this moment in the myth and adapt it to the viewing experience of art is to think of ourselves as the Venus figure, watching over the shrines of the works of art and granting them life through our continued dialogue and interpretation of them.^{xxxii} The selected works discussed in this essay demonstrate transformation, both physical and psychological, that does not exist only because of the skill of the artist or poetic motifs in the work, but through the viewer's participation in the work's expression and interpretation as well.

While the *Bacchus* was hidden and thus never influenced a wide public at the time of its creation, its use of developing narrative by being viewed in the round demonstrates the interactive relationship the work was intended to have with the viewer. Lieberman comments on how Michelangelo's use of an unfolding narrative across multiple views preceded the work of Bernini, going so far to claim that Bernini learned how to make his Borghese works, with their unfolding narratives, from Michelangelo (Lieberman 72). While Wilkins discusses Bernini's *Apollo and Daphne*, her analysis can be applied to the *Pluto and Proserpina*. She writes, "Bernini's sculpture represents the metamorphosis of poetic texts into...Solid, unyielding stone has taken on the quality of fluid narrative" (Wilkins 391). In this way, the sculptor transforms the stone into a story that continues to transform as the viewer circles the sculpture. The viewer thus brings these sculptures, the *Bacchus* and the *Pluto and Proserpina*, to life through the contemplation of these multiple views, revealing the unfolding narratives

of each work.

Although the *Pluto and Proserpina* was given to the Ludovisi's villa, it was still placed in a similar cultural milieu. Works of art placed in homes at this time were viewed and discussed by guests of parties, provoking intellectual discussion (Warwick 357-60). Warwick equates illusionism in seventeenth-century Italian art with the dramatic and theatrical performance she believes was inherent to the social fabric of the time.^{xxxiii} In her interpretation, "the object acts through instigating dialogue, conversation, a performance of culture with its audiences" (Warwick 353). Proserpina's pleas for help and the dynamic composition invite conversation from the viewer. The artist alone cannot incite this intellectual contemplation of the work. It takes an outside force, like Venus in the poem, to metaphorically bring the works to life through contemplation of their craft, form and content.

However, the intended placement of this work in the Villa Borghese reveals how the work was originally meant to inspire interest and attract the viewer's attention. Kenseth refutes the notion that Bernini's sculptures were meant to have a single viewpoint. Instead, she argues that their narratives were revealed across 180 degrees, demanding the viewer to walk around them in order to obtain their full meaning.^{xxxiv} Instead, the viewer would have seen Pluto's side, where his "lunging form instantly captures the beholder's imagination and initiates a drama that unfolds in subsequent views" (Kenseth 202). She also asserts that when Proserpina cries out, she is reaching out to us, the viewer, for help. In these ways, the work activates the space of the viewer and makes us complicit in her continued suffering. We bring her pain to life by gazing upon her sculpted form.

While Warwick discusses how a culture of performance shaped how Bernini's art was made and how people received it, her ideas can be applied to how the Michelangelo works are viewed today. The curious drunken stupor of the *Bacchus*, its relatable nature of appearing like one does when drunk and the way it seems to transition in between sobriety and drunkenness spark curiosity and inquiry. Thus, the *Bacchus* draws the viewer through its odd, drunken pose and the smooth finish of its skillfully carved surface, inviting conversation and scholarship. Likewise, the *Atlas* invites conversation through its rough suggestion of the figure escaping from the stone. In both works, the viewer grants a new sense of life to the work through their interactions with them. Finally, in Bernini's Chigi niche statues, the *Saint Mary Magdalen* and the *Saint Jerome* interact with the viewer in their religious setting by protruding into the viewer's space and imposing their visually powerful displays of saintly penitence, inspiring thoughts of repentance and salvation in the viewer. These figures force themselves into the viewer's space with their striking sense of illusionistic lifelikeness. These sculptures force us to confront them and the violent emotion contained in their poses, which push the fragile marble to its utmost limits.

In all of these works, the viewer operates in a similar way to Venus in Ovid's poem. Just as she brought Pygmalion's statue to life, the viewer uses their interpretations to augment their projection of concepts from art theory onto the works and support their visual observations that these works are so illusionistic, they look ready to come to life. Barolsky claims that irony lying behind the influence of Pygmalion on art is that even if these works seem lifelike, they will never come to life, and are therefore more like the myth of Narcissus's eternally frozen image (Barolsky "As in Ovid, So in Renaissance

Art” 453). However, this argument misses a crucial point, that the suggestion of Pygmalionesque lifelikeness in these works has helped them stay alive in the Western art canon, that we regard these artists highly today and keep these works “alive” through the ongoing experience of viewing them because they possess these powerful qualities of visual expression.

Conclusion

In this essay, I have argued that the Pygmalion myth from Ovid’s *Metamorphoses* offers ways to reexamine and reinterpret the selected works by Michelangelo and Bernini, building on modern scholarship to show that they transform physical, emotional, and theoretical content through virtuosic craft and illusionistic detail. The *Metamorphoses* strongly influenced the art of the Renaissance and the Baroque. This essay has just focused on one of the many myths in the foundational text. Michelangelo and Bernini were inspired by Ovid’s work, albeit in different ways. Michelangelo’s *Atlas* and *Bacchus* as well as Bernini’s *Pluto and Proserpina* and *Saint Mary Magdalen* and *Saint Jerome* all demonstrate virtuosic carving, poetic transformation and illusionism, all of which incorporate theoretical aspects of the Pygmalion myth. However, it is the viewer who has the power to project these meanings and associations onto the works through their active engagement with the art, poetically bringing these works to life as Venus did in the myth. Barolsky notes the ubiquity of Pygmalion in art: “The myth of Pygmalion is sometimes so deeply embedded in the description of art that it can easily escape notice” (Barolsky “As in Ovid, So in Renaissance Art” 453). I have shown not only the importance of reexamining the themes of these artworks through the lens of this poetic myth, but also the literature’s potential to influence artistic expression and art’s power to visually depict

literary themes and concerns.

Images

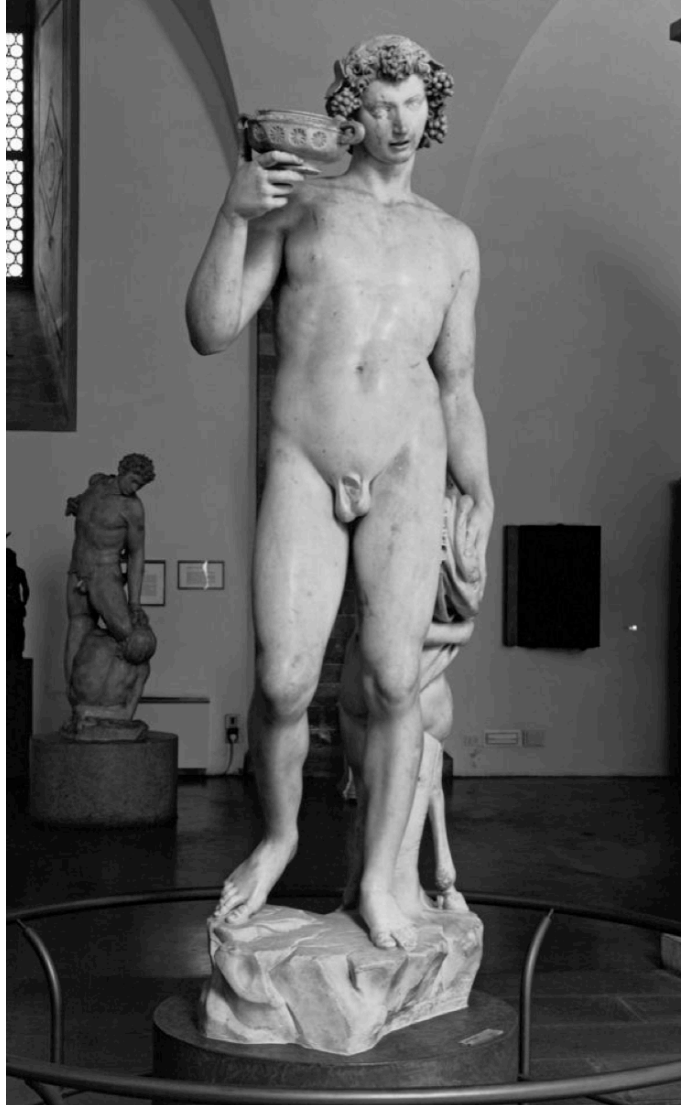


Figure 1. Michelangelo, *Bacchus*, 1496-7, marble, Museo Nazionale del Bargello, Florence. *Artstor*.

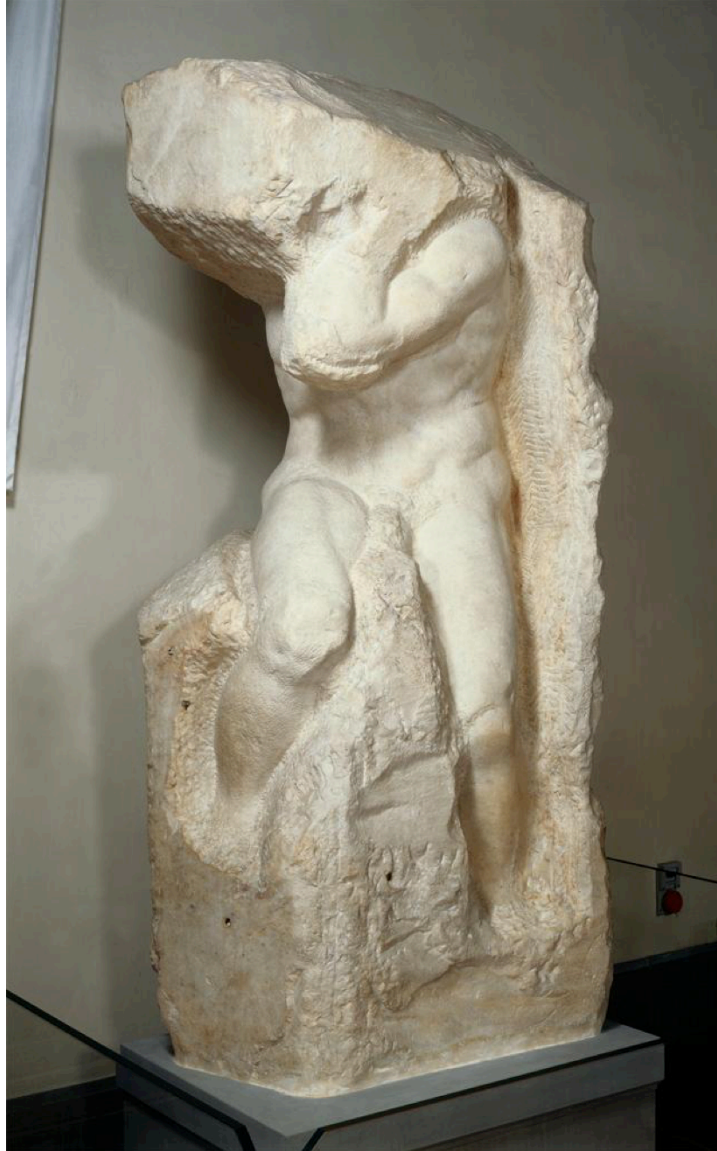


Figure 2. Michelangelo, *Atlas*, 1530-4, marble, Accademia Gallery, Florence. *Artstor*.



Figure 3. Gianlorenzo Bernini, *Pluto and Proserpina*, 1621-2, marble, Galleria Borghese, Rome. Artstor.

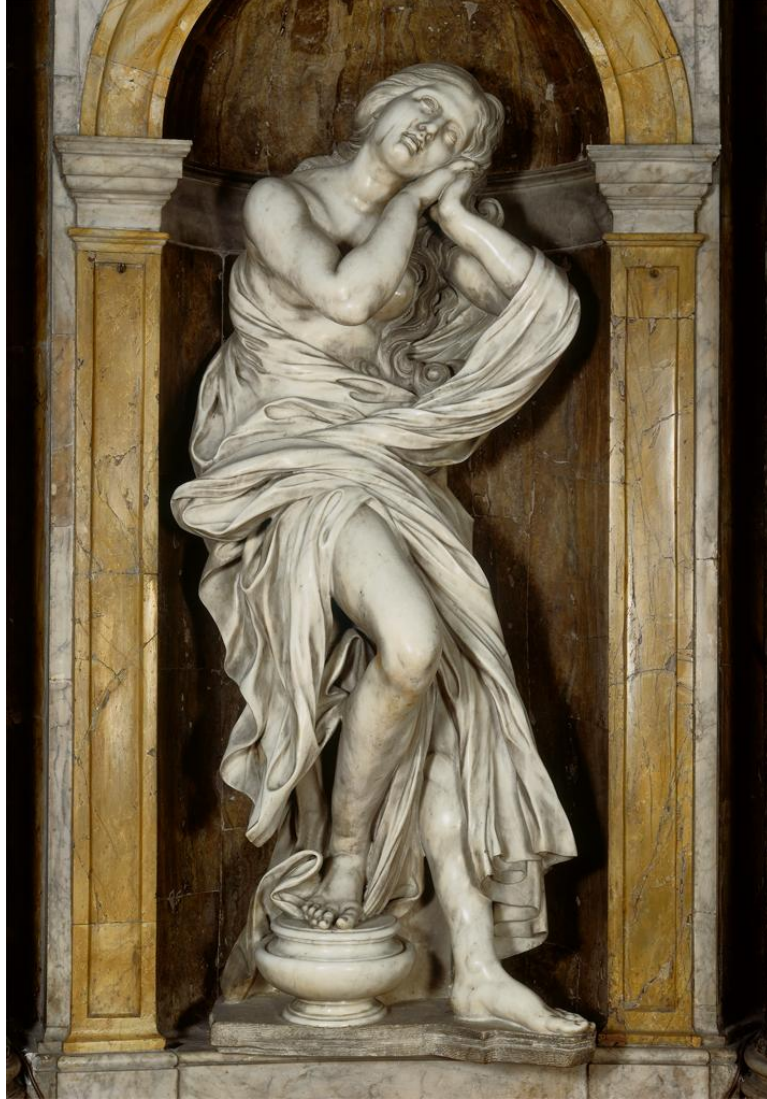


Figure 4. Gianlorenzo Bernini, *Saint Mary Magdalen*, 1661-3, marble, Chigi Chapel, Siena. *Artstor*.



Figure 5. Gianlorenzo Bernini, Saint Jerome, 1661-3, marble, Chigi Chapel, Siena. *Artstor*.

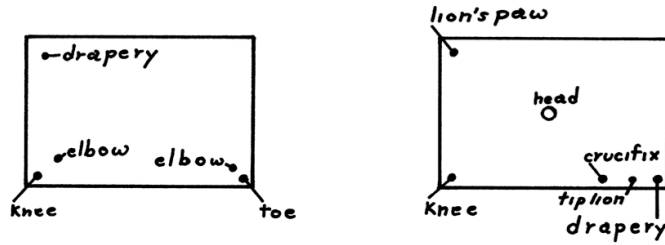


Figure 6. Diagonal of cross-sections of Bernini's *Magdalen* and *Jerome*. (Caption and diagrams from Harriet Feigenbaum Chamberlain, "The Influence of Galileo on Bernini's Saint Mary Magdalen and Saint Jerome" *The Art Bulletin* 59.1 (1977), 79).

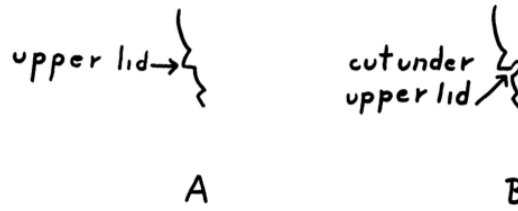


Figure 7. Diagram illustrating a double undercut (Caption and diagrams from Harriet Feigenbaum Chamberlain, "The Influence of Galileo on Bernini's Saint Mary Magdalen and Saint Jerome" *The Art Bulletin* 59.1 (1977), 82).

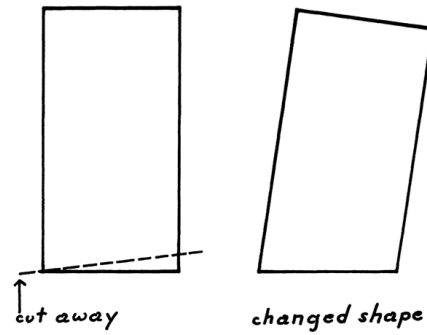


Figure. 8. Diagram of a method of altering the base plane of a block of stone (Caption and diagrams from Harriet Feigenbaum Chamberlain, "The Influence of Galileo on Bernini's Saint Mary Magdalen and Saint Jerome" *The Art Bulletin* 59.1 (1977), 81).

Notes

ⁱ The scope of this paper does not allow for a discussion of naturalism and idealism in addition to illusionism. These other sub-types of transformation of the human body also apply to the myth and works of art. They will be evoked in my discussion of the artwork, but not explicitly named and explained as separate themes. For example, my formal analysis of the visual details in several of the works focuses on naturalistic details, but for the sake of focusing on the four themes emphasized in this paper, I will not invoke a discussion of naturalism as its own theory. In an expansion of this project, it would be interesting to do more research and add these theories as separate sections and expand upon how they can be seen as parallels between Pygmalion's creation of his statue and Michelangelo and Bernini's creation of their artworks.

ⁱⁱ When I use the word "poetic" throughout this paper, I

will use it to mean possessing imaginative qualities or emotional expressiveness.

ⁱⁱⁱ For more information about the influence of Ovid's poetry on art and literature beginning the twelfth century, see the essays in: Charles Martindale, ed. *Ovid Renewed: Ovidian Influence on Literature and Art from the Middle Ages to the Twentieth Century*. (Cambridge: Cambridge University Press, 1990). For information on Ovid's influence in literature beginning in the twelfth-century, see C.W. Grocock's chapter "Ovid the Crusader" (55-70), who argues that economic and social changes may have increased Ovid's prominence and thus his influence on court poetry. For information on Ovid's influence on art beginning in the Middle Ages, see Nigel Llewellyn's essay "Illustrating Ovid" (151-67), where he provides a survey of how different artists have adopted Ovidian subject matter into their work.

^{iv} When I word the use "Ovidian" in the paper, I am using it to mean relating to Ovid's works, poetic style and myths. See: Paul Barolsky. "As in Ovid, So in Renaissance Art" *Renaissance Quarterly*, 51.2 (1998), 451-74 and Barolsky's other essays about Ovidian influence on Renaissance and Baroque art. Barolsky uses the word in this way to generally mean "Ovid-like", be it in the style of Ovid, Ovid's subjects or in the same poetic way as Ovid. I am thus adopting his established terminology as a way of further aligning my work as a continuation of the work he has started on this subject.

^v For more information on these expansions of Ovid in art, and specifically in Baroque art, see: Anne Wilkins. "Bernini and Ovid: Expanding the Concept of Metamorphosis" *International Journal of the Classical Tradition*, 6.3 (Winter, 2000), 383-408. In her essay, Wilkins also makes the point that Ovid's work had such a major influence that at times it is difficult to tell whether specific pieces of art are influenced directly by the Roman

poet or artistic re-interpretations of Ovid's myths contemporary to the time (384-5).

^{vi} In her essay, Warwick explains how intellectual circles in Rome at the time took Ovidian myths and reinterpreted them for their own agendas, whether they were poetic, artistic, philosophical or political. For interesting discussion of the Ovid's influence on Petrarchan poets, Petrarch, Marino, Caravaggio, Carracci, etc., see pp. 365-70.

^{vii} While my research focuses exclusively on Pygmalion, several of the essays I cited in my research provide useful discussion of some of the other myths in Ovid's *Metamorphoses*. For a discussion of Aglauros and a discussion of the petrification of Phineus, Perseus and Andromeda, see: Paul Barolsky. "Andromeda's Tears" *Arion: A Journal of Humanities and the Classics*, 6.3 (Winter, 1999), 24-8. For a discussion of Apollo and Daphne, see: Paul Barolsky, "As in Ovid, So in Renaissance Art," Paul Barolsky, "Ovid, Bernini, and the Art of Petrification," Andrea Bolland, "*Desiderio and Diletto: Vision, Touch and the Poetics of Bernini's Apollo and Daphne*," Genevieve Warwick, "Speaking Statues: Bernini's Apollo and Daphne at the Villa Borghese" and Ann Wilkins, "Bernini and Ovid: Expanding the Concept of Metamorphosis." For a discussion of Atlas, Medea, Lichas, Hercules, Olenos, Lethaea, Scylla, Peneus, Anaxarete, Iphis, Hecuba, Ceres, Polydorus, see Paul Barolsky "Ovid, Bernini, and the Art of Petrification." For a discussion, of Deucalion and Pyrrha, see: Paul Barolsky, "As in Ovid, So in Renaissance Art," Paul Barolsky, "Florentine Metamorphoses of Ovid," Andrea Bolland, "*Desiderio and Diletto: Vision, Touch and the Poetics of Bernini's Apollo and Daphne*," *The Art Bulletin*, 82.2 (2000), 309-30, Anne Wilkins, "Bernini and Ovid: Expanding the Concept of Metamorphoses." For a discussion of Medusa, see: Paul Barolsky, "As in Ovid, So

in Renaissance Art.” For a discussion of Narcissus, see: Paul Barolsky, “As in Ovid, So in Renaissance Art” and Andrea Bolland, “*Desiderio and Diletto: Vision, Touch and the Poetics of Bernini’s Apollo and Daphne.*”

^{viii} I chose to make this argument using these two artists because their work is metonymic for the styles of their times. Michelangelo and Bernini are the all-star artists of the Italian Renaissance and Italian Baroque respectively. Much scholarly ink has been spent examining their artistic genius, an established trend in art historical research, as will be evident in this paper. To talk about their genius is not only to address their reputation, but also to further justify connecting them to the myth and using them as representative for these larger artistic and cultural moments. In a larger and more extensive project, it might be interesting to expand this argument into other artists from the Renaissance and the Baroque and see if it is applicable. However, for the purposes of this paper, I will apply my reading of Ovid and the themes of illusionism and transformation only to Michelangelo and Bernini.

^{ix} It is worth noting here that it is not my intention in this paper to make sweeping generalizations about the relationship between Pygmalion and Renaissance and Baroque art, but rather to make an argument that there are parallels between the poem and the specific works of art, which demonstrate some of the key characteristics of these two periods of art.

^x In this essay, virtuosic will be defined as an extremely level of high skill that sets an artist apart from other artists in the same possible. Simply said, they are the best at their medium of art in their era based upon what art was trying to do at that moment, as Michelangelo and Bernini do.

^{xi} Now that I have stated my thesis, it is worth also listing what I will *not* be arguing in the scope of this paper in

order to be as clear as possible what lies within the scope of this essay. As said, I will only be discussing the myth of Pygmalion from the *Metamorphoses*, omitting any discussion of its relation with other Ovidian myths. I will not be discussing naturalism or idealism in detail. I will not be making sweeping generalizations about all of Renaissance or Baroque art. My analysis draws largely upon formal and visual analysis supplemented by sociohistorical context so patronal history and physical context will not be as important within the scope of this paper. I will also be omitting much discussion of the moralizing philosophy attached to these works in Michelangelo and Bernini's works, but will address this idea again in a later footnote. I will also not include detailed discussion of the *paragone*, the intellectual debates about the superiority of the arts, in my section on illusionism, but a footnote will address research and include some discussion of this idea. I will also not be discussing artist self-formation through these works, but for a discussion of this focused on Michelangelo and the *Bacchus*, see: Linda A. Koch. "Michelangelo's Bacchus and the Art of Self-Formation" *Art History*, 29.3 (June, 2006), 345-86.

^{xii} Barolsky has written several essays on the representation and transformation of Ovidian themes and the demonstration of Ovidian influence on Italian Renaissance and Baroque art. In addition to "As in Ovid, So in Renaissance Art" and "Florentine Metamorphoses of Ovid," see the following two articles by Barolsky addressing Ovidian influence that I cite this essay: "Andromeda's Tears" and "Ovid, Bernini, and the Art of Petrification" *Arion: A Journal of Humanities and the Classics*, 13.2 (Fall, 2005), 149-61.

^{xiii} A poet contemporary to Michelangelo evokes the idea that Michelangelo himself was a Pygmalionesque sculptor in a poem he wrote about the figures carved for

the Medici Chapel. See: Paul Barolsky, "Florentine Metamorphoses of Ovid," 15.

^{xiv} These paragraphs and references briefly mentioning Pygmalion throughout Barolsky's research raised several questions for me. Here are some of the questions arose for me from these passages: How and why does the *non finito* by Michelangelo express some of the inherent themes in the Pygmalion myth? How does using Pygmalion specifically as a way to think about the art enlighten our interpretation of the works? What other ways does the *Pluto and Proserpina* demonstrate the themes within Pygmalion? What do these connections to Ovid's myth ultimately add to my understanding of these works?

^{xv} The *Bacchus*, commissioned by Cardinal Raffaele Riario but ultimately placed in the home of Jacopo Galli, demonstrates several of the key ideas present in the Pygmalion myth through its anatomical detail, depiction of the god in a drunken stupor and the changing form of the work as one walks around it (Ralph Lieberman. "Regarding Michelangelo's *Bacchus*" *Artibus et Historiae*, 22.43 (2001), 65).

^{xvi} The *Captives* are widely believed to have been intended for Pope Julius II's tomb (Paul Barolsky, "As in Ovid, So in Renaissance Art" 461). They, and the *Atlas* in particular, have taken on a poetic and Ovidian meaning having been placed in the Medici's pastoral Ovidian grotto as well as their expression of the *non finito* (Paul Barolsky, "Florentine Metamorphoses of Ovid 16).

^{xvii} Bernini's *Pluto and Proserpina*, originally commissioned by the Borghese family for their villa, was given to the Ludovisi, the new papal family in Rome at the time, for theirs and demonstrates key parts of the Pygmalion myth through its changing views as one walks around the work and it's attention to anatomical detail as well, a theme throughout all of the works (Warwick 365).

^{xviii} Pope Alexander VII Chigi commissioned Bernini to redo the interior of the Chigi Chapel in Siena's Cathedral, of which the *Mary Magdalen* and *Saint Jerome* niche statues are a part. They reused the Chapel of Ansanus and replaced the iconography of the city's previous patron saint with that of the Virgin Mary. Bernini chose to decorate it with the theme of forgiveness. (Harriet Feigenbaum Chamberlain. "The Influence of Galileo on Bernini's Saint Mary Magdalen and Saint Jerome," *The Art Bulletin*, 59.1 (1977), 71). Like in many of his other later chapels, Bernini decided to create all of the decorations under the umbrella of a certain theme, in this case forgiveness. The *Mary Magdalen* and *Saint Jerome* are placed in niches near the entrance so that "only on turning to leave will the viewer feel compelled to stop and look" at them (Chamberlain, 71). While beyond the scope of this paper, it would be interesting to research why Bernini placed these works in this way and what he hoped to accomplish through their positioning within the chapel.

^{xix} In a longer project, it would be interesting to explore how these themes were reworked and transformed in a larger number of works by these two artists from a more diverse range of contexts and patrons.

^{xx} For more information about the evolving role of the "genius" in art, see: William Wallace, "The Artist as Genius" *A Companion to Renaissance and Baroque Art*. Ed. Babette Bohn and James M. Saslow, (Somerset, NJ, 2013), 151-67.

^{xxi} Pygmalion worked in "ivory" while Michelangelo and Bernini worked primarily in marble. It is worthwhile to consider here, while tangential to this part of my paper, the importance of the material in the myth compared with the materials of the master artists I am studying. Ivory, which comes from the tusks of living animals such as elephants and rhinoceroses, has size constraints would

appear to make carving a life-sized female figure at the very least nearly, if not completely, possible. Ivory and marble are both white, forming a visual linkage between the two. Thus, another possible explanation is that in the poem, ivory should not be taken to mean the literal material, but rather ivory as a shade of white. Using the word in this way would then simply suggest that the sculpture itself was an ivory shade of white, making it possible, and even likely, that the sculpture of his maiden was carved from marble like Michelangelo and Bernini's works, thus also demanding virtuosity in artistic genius to create a lifelike figure from this temperamental material. More research is needed to study the history of the translation into English to see if there are any linguistic clues as to which of the two definitions of ivory Ovid meant in the original text, but nevertheless these linguistic musings are beyond the scope of the paper. However, it is clear here that the use of ivory by Pygmalion, be it the material itself or a material of that color, has implications of difficulty and necessity of skill echoed in the work of Michelangelo and Bernini.

^{xxii} It should be noted here that Bernini did not work exclusively in marble sculpture, but also did a number of casted bronze works, perhaps most notably the *Baldacchino* in Saint Peter's in Rome. Because of the difficulty of casting in bronze, in ensuring that each of the pieces was casted properly, this medium too serves as a display of Bernini's virtuosity with materials even if it does not serve an exactly parallel with Ovid's mythic Pygmalion, who carved ivory. For more information about Bernini's bronze techniques and the *Baldacchino*, see: Jennifer Montagu. "Founders and Sculptors" *Roman Baroque Sculpture: The Industry of Art* (New Haven, 1989). 48-75.

^{xxiii} For more information on marble sculpture and carving, see the rest of Jennifer Montagu's chapter "From

the Quarry to the Church” from her book *Roman Baroque Sculpture*, 21-47.

^{xxiv} In his biography of the Renaissance artist, Vasari describes Michelangelo’s carving practice when he writes that Michelangelo “taught sculptors the way to make statues without spoiling them, by removing the marble so as to enable them to make such alterations as may be necessary” (Vasari, par. 19) Here, he is referring to Michelangelo’s technique of carving from one side to release the figure, a way of carving that allowed him to also make changes as he worked if needed.

^{xxv} Bernini’s capturing of the minute details of the human body, rendering flesh as soft and illusionistic, involved the *paragone* debates of the time as well, which discussed whether painting or sculpture was the better art. While this topic lies beyond the scope of this paper, Michelangelo and Bernini engaged in these debates in different ways to affirm the superiority of sculpture. For Bernini, this meant creating “a sweet, pleasing illusion of softness effected by means of imperceptibly gradual transitions from highlights to shadows” (Warwick 373). And while Warwick argues that Bernini does this in his early works, this trend continues and perhaps becomes more subtle in his later works like the niche statues in the Chigi chapel: the *Mary Magdalen* and *Saint Jerome*, where he creates poses in stone only created before in painting (Chamberlain 71). Michelangelo also participated in the *paragone* debates, which were newer and more current in his time than Bernini’s, with the goal of imitation at the center of these intellectual arguments over whether painting or sculpture was the higher art. While Michelangelo resisted participation in these debates, when he did he either promoted sculpture or *disegno* as the superior art (Judith Dundas. “The Paragone and the Art of Michelangelo” *The Sixteenth Century Journal*, 21.1 (Spring, 1990), 87-8).

^{xxvi} For more information about Michelangelo's knowledge and theories about anatomy, which go beyond the scope of this paper, see: David Summers, "Anatomy" *Michelangelo and the Language of Art*, (Princeton, NJ, 1981), 397-405.

^{xxvii} In "Ovid, Bernini, and the Art of Petrification," Barolsky emphasizes the essential role of stone within Ovid's *Metamorphoses*. He writes, "The world of Ovid, a place of marble architecture and statuary, is a site of artifice, which is inevitably measured against the originary art of nature, where stone takes on architectural form as a matter of course" (Barolsky, "Ovid, Bernini, and the Art of Petrification" 151). Thus, in Barolsky's eyes, the use of stone in Ovid reflects a poetic vision of the world and humanity.

^{xxviii} Contemplation of this kind of transformation raises new questions beyond the scope of this paper about the interdependent relationship between poet and sculpture in their own time, points difficult to prove as Kenseth notes in her analysis of the relationship between the *Apollo and Daphne* and Marino's poem, contemporary to Bernini, since it is nearly impossible to tell in that case who influenced who. For more information about these difficulties and Marino's poem, see Joy Kenseth, "Bernini's Borghese Sculptures: Another View" *The Art Bulletin*, 63.2 (1981), 191-210.

^{xxix} See Genevieve Warwick, "Speaking Statues: Bernini's Apollo and Daphne at the Villa Borghese" and Ann Wilkins, "Bernini and Ovid: Expanding the Concept of Metamorphosis" for interesting discussion of Christian moralizing attitudes projected onto these works and fusing Christian and Pagan perspectives.

^{xxx} Bernini would have had control over the lighting in the chapel where the sculptures were placed (Chamberlain 82).

^{xxxi} To my knowledge, this section of my argument is

almost entirely my own idea. It is certainly possible that Ovidian scholars beyond the scope of this paper have conjectured a similar stance, but I came up with this notion that the viewer activates the sculpture through a combination of critical analysis of the myth and paying close attention to how people interact with art.

^{xxxii} To ensure that this metaphor is clear, Pygmalion represents the artist (Michelangelo or Bernini), the statue symbolizes the works of art that I have been discussing in this essay and Venus symbolizes the viewer, the audience for the work who brings their own experiences and interpretations to the works and projects them onto their viewing experience.

^{xxxiii} For more information about the culture of performativity during Bernini's time, see Genevieve Warwick, "Speaking Statues: Bernini's Apollo and Daphne at the Villa Borghese."

^{xxxiv} Using contemporary guidebooks, Kenseth argues that before giving the work away to Cardinal Ludovisi, he put the *Pluto and Proserpina* along the west wall of the northeast room of the villa, that "the visitor, upon entering the room, would not have seen the statue's front face but its left side, where Pluto's aggressive stride is prominently displayed" (Kenseth 201).

References

- Barolsky, Paul. "Andromeda's Tears." *Arion* 6.3 (1999): 24-8. Web.
- . "As in Ovid, so in Renaissance Art." *Renaissance Quarterly* (*Renaissance Soc. of America, New York*) 51.2 (1998): 451-74. Web.
- . Barolsky, Paul. "Ovid, Bernini, and the Art of Petrification." *Arion* 13.2 (2005): 149-61. Web.
- . "Florentine Metamorphoses of Ovid." *Arion* 6.1 (1998): 9-31. Web.
- Bolland, Andrea. "Desiderio and Diletto: Vision, Touch, and the Poetics of Bernini's Apollo and Daphne." *The Art Bulletin* 82.2 (2000): 309-30. Web.
- Chamberlain, Harriet Feigenbaum. "The Influence of Galileo on Bernini's Saint Mary Magdalen and Saint Jerome." *The Art Bulletin* 59.1 (1977): 71-84. Web.
- Clifton, James. "Vasari on Competition." *The Sixteenth Century Journal* 27.1 (1996): 23-41. Web.
- Cole, Michael. "Bernini Struts." *Material Identities*. Ed. Joanna Sofaer. Oxford: Wiley-Blackwell, 2008. Web. 55-66.
- Cotterell, Arthur. "Persephone" *A Dictionary of World Mythology*. Oxford: Oxford University Press, 1997. Web.
- Dundas, Judith. "The Paragone and the Art of Michelangelo." *The Sixteenth Century Journal* 21.1 (1990): 87-92. Web.

- Gilbert, Creighton E. "What is Expressed in Michelangelo's "Non-Finito." *Artibus et Historiae* 24.48 (2003): 57-64. Web.
- Hibbard, Howard. *Bernini*. New York: Penguin Books, 1965. Print.
- Kenseth, Joy. "Bernini's Borghese Sculptures: Another View." *The Art Bulletin* 63.2 (1981): 191-210. Web.
- Koch, Linda A. "Michelangelo's *Bacchus* and the Art of Self-Formation." *Art History* 29.3 (2006): 345-86. Web.
- Lieberman, Ralph. "Regarding Michelangelo's "Bacchus"." *Artibus et Historiae* 22.43 (2001): 65-74. Web.
- Martindale, Charles. *Ovid Renewed: Ovidian Influences on Literature and Art from the Middle Ages to the Twentieth Century*. Cambridge: Cambridge University Press, 1988. Web.
- Montagu, Jennifer. *Roman Baroque Sculpture: The Industry of Art*. New Haven: Yale University Press, 1989.
- Ovid. "The Story of Pygmalion and the Statue." *The Internet Classics Archive | Metamorphoses by Ovid*. Daniel C. Stevenson. Web. 15 Nov. 2015. <<http://classics.mit.edu/Ovid/metam.10.tenth.html>>.
- Summers, David. "Anatomy" *Michelangelo and the Language of Art*. Princeton, NJ: Princeton University Press, 1981. Print. 397-405.
- Vasari, Giorgio. "Vasari's Biography of Michelangelo." *Vasari's Biography of Michelangelo*. SUNY Oneonta.

Web. 09 Dec. 2015.
<https://www.oneonta.edu/faculty/farberas/arth/Arth213/michelangelo_yasari.html>.

Wallace, William E. "La Bella Mano: Michelangelo the Craftsman." *Artibus et Historiae* 32.63 (2011): 85-99. Web.

---. "The Artist as Genius." *A Companion to Renaissance and Baroque Art*. ed. Babette Bohn and James M. Saslow. Somerset, NJ: Blackwell-Wiley, 2013. Web. 151-67.

Warwick, Genevieve. "Speaking Statues: Berninis Apollo and Daphne at the Villa Borghese." *Art History* 27.3 (2004): 353-81. Web.

Wilkins, Ann Thomas. "Bernini and Ovid: Expanding the Concept of Metamorphosis." *International Journal of the Classical Tradition* 6.3 (2000): 383-408. Web.

Self-injury and pain: An analysis of non-suicidal self-injury and the effects it has on the neurological processing of pain
Darci Gautam (CLA 2019)

Abstract

Non-suicidal self-injury (NSSI) is a prevalent but poorly understood syndrome. Despite the number of documented cases of NSSI, there is a paucity of research explaining the neurological mechanisms behind this behavior. Much of the prior research on this topic has focused on the emotional reasoning behind NSSI and on finding treatments for this syndrome. This paper examines the literature on NSSI from a neurological perspective in an attempt to further an understanding of the causes behind this distressing behavior. In addition, a close examination of the research demonstrates a correlation between emotional and physical pain, as well as altered pain processing in individuals who engage in NSSI. However, much of the research that has been conducted in this field is often isolated and does not appear to be holistic, often pertaining to a singular aspect of this syndrome and studying smaller subsets of the population that do not provide for generalizable results. Therefore, in order to gain a better understanding of this prevalent syndrome and to eventually reach viable treatment and preventative options, more research must be conducted that addresses NSSI more representatively from a broader perspective.

Introduction

Non-suicidal self-injury (NSSI) is a syndrome defined as the intentional harming of oneself without suicidal intent. The most common form of NSSI is the act of cutting the skin followed by intentional bruising. Other forms of NSSI include burning and wound excoriation (Wolff et al.; Haines et al.; *Diagnostic and Statistical Manual of Mental Disorders*). NSSI is cited as the most frequent reason for psychiatric visits to emergency departments, and although it is a serious and widespread issue, startlingly little is known about the neurological abnormalities correlated with this behavior (Stanley et al., 2010). Greater knowledge about NSSI and the neurological processing of pain in those who engage in self-injurious behaviors could lead to improved treatments and possibly an overall reduction in self-injury. The literature review presented in this paper provides a comprehensive description and analysis of various studies related to a number of aspects of NSSI and related subjects. The results of this review emphasize the need for more holistic research on this topic and provide suggestions for how this research could be designed.

fMRI Studies of NSSI

Historically, much of the research investigating NSSI was carried out in the form of self-reported surveys. This approach has formed the foundation for much of our understanding about this syndrome. Self-reports are especially useful for identifying an individual's motivation for engaging in this behavior and for determining the levels of effectiveness of NSSI for each individual. Such studies have identified the primary driver of NSSI cited by patients as a reduction in negative affect (Nock and Prinstein, 2004; Klonsky, 2007; Bresin et al., 2010).

Although these studies have been instrumental in examining the emotional stressors that trigger NSSI, they have not provided some critical information concerning the biological factors that contribute to this behavior. In recent years, there has been an increase in research using functional magnetic resonance imaging (fMRI) as a means to study NSSI (Ogawa and Sung, 2007). Through fMRI studies, the amount of knowledge on the neurological processing that facilitates NSSI and those who engage in it has greatly increased. These studies have enabled the discovery of abnormalities in the regulatory pathways of people who engage in NSSI (Bonenberger et al., 2015; Plener et al., 2012; Schmahl et al., 2006).

Through the combination of self-reports and fMRI scans, researchers have been able to simultaneously gain knowledge of an individual's emotional experience during NSSI and of the activity in the brain that occurs during an approximation of NSSI created in a laboratory setting by the application of painful stimuli. In one such study, participants with and without NSSI were given two stimuli rated as cool and cold in self-administered and experimenter-administered conditions (Osuch et al., 2014). Participants were scanned using fMRI technology during these instances, and were asked to give self-reports on the level of pain each stimulus provided and the level of relief that they felt after the instance was over. Results from the self-reports showed that in the NSSI group, there was increased relief after the cold stimulus, especially when self-administered. Results from the fMRI scans showed that participants in the NSSI group had increased activity in the reward processing centers of the brain and in the dorsal striatum, the area of the brain associated with addiction. NSSI participants also showed poor connection between the anterior

cingulate cortex (ACC) (Figure A), orbital frontal cortex (OFC), and the medial prefrontal cortex (MPFC). These areas are often associated with affective regulation, and deficits in these areas may explain the emotional reactivity often found in individuals who engage in NSSI.

Social Pain and NSSI

Because individuals who engage in NSSI report a reduction in negative affect caused by physical pain (Bresin et al., 2013; Chapman, Gratz, and Brown, 2006; Klonsky, 2007; Nock and Prinstein, 2004), understanding the relationship between emotional pain and physical pain may be imperative to understanding this syndrome. Virtually all humans experience social and emotional pain, and yet until recently, there was a minute amount of knowledge about the neurological mechanisms that incite emotional distress. Researchers have suggested that emotional and physical pain may be more closely connected than what was originally hypothesized (Kross et al., 2011). Through investigating the relationship between the two types of pain, researchers found common usage across languages of physical pain words to describe emotional discomfort, *broken heart*, *hurt feelings* (Eisenberger and Lieberman, 2004).

Eisenberger and colleagues expanded this idea further, hypothesizing that affective pain, or the cognitive and emotional reaction to pain, is felt in the same region of the brain for both types of pain. To test this hypothesis, the researchers scanned participants using fMRI while they participated in a virtual ball tossing game. Participants were led to believe that they were playing with two other individuals when they were in fact playing with the central processing unit of a computer. After a period of time, these computerized players began passing the ball between themselves, excluding the participant.

When compared to fMRI scans of people experiencing physical pain, the exclusion scans showed similar brain activity in two key areas, the ACC and the right ventral prefrontal cortex (RVPPFC), confirming the hypothesis that the affective dimensions of social and physical pain have a shared neural circuitry.

Since Eisenberger and colleagues' original study, more research has been conducted concerning the overlap between emotional and physical pain and evidence has supported Eisenberger's original hypothesis (DeWall et al., 2010; Eisenberger and Lieberman, 2004; Eisenberger and Lieberman, 2005; Eisenberger, 2012a; Eisenberger 2012b; Eisenberger, 2015; Masten, Morelli and Eisenberger, 2011). Although there has been some controversy surrounding this issue, recent studies suggest that all parts of the ACC, the dorsal anterior cingulate cortex (dACC), the ventral anterior cingulate cortex (vACC), and the pregenual anterior cingulate cortex (pgACC), are active during emotional pain, discrediting theories that only the dACC is active in such instances (Eisenberger, 2015). The research also shows that children show more vACC activity than adults, suggesting that the dACC develops later in life. This information could be important for understanding why NSSI is more prevalent in younger populations; however, more research examining the functions of the different parts of the ACC as well as the development of the brain is necessary to make such a claim.

Despite the apparent similarity between the two topics, little research has connected physical and emotional pain to NSSI. In one study, Eisenberger found that individuals with lower thresholds for physical pain are also more susceptible to emotional pain and are likelier to experience "hurt feelings" (Eisenberger, 2012a). This

correlation suggests that individuals who engage in NSSI, who are known for their poor emotional regulation, would be more sensitive to physical pain.

Borderline Personality Disorder

Until recently, NSSI was considered a criterion for borderline personality disorder (BPD) (Glenn and Klonsky, 2013); thus it is difficult to discuss NSSI without also discussing BPD to some extent. Patients with BPD are characterized by emotional instability, low self-esteem, traumatic stress, and impulsivity (*DSM-5, 2000*; Schmahl et al. 2003).

Due to the increased prevalence of NSSI in populations with BPD, and the fact that these two syndromes have been historically paired, many of the studies investigating NSSI specifically researched BPD populations. One such study examined the emotional reactivity of BPD patients when viewing images related to NSSI (Plener et al., 2012). In this study, females with BPD were matched with individuals without a history of NSSI and without an axis I or II psychiatric disorder who were part of a control group. Each group underwent an fMRI scan while being shown several images, some of which were random and some of which were related to NSSI. Participants were later shown the images again and asked to rate the emotional content of each picture. The BPD group rated the NSSI images to be more emotionally arousing than the control group did. Data from the fMRI scans showed that NSSI participants had higher blood oxygen level dependent (BOLD) (Ogawa & Sung, 2007) activity in the left superior parietal cortex, amygdala, and ACC bilaterally and in the right cerebellum than did control participants. These results were consistent with the fact that BPD patients generally show more emotional reactivity than controls (Schmahl et al., 2006).

Contrary to Eisenberger's hypothesis of increased emotional and physical pain sensitivity, one study showed that patients with BPD actually have higher pain thresholds than psychiatric controls (Schmahl et al., 2006). In this study, twelve BPD patients and twelve controls were given a variety of heat stimuli. Participants were asked to rate each stimulus for painfulness on a scale of zero to one hundred. The ratings were used to form an individual heat stimulus that matched each participant's pain threshold. After this procedure, participants were scanned using fMRI while being given a heat stimulus ten different times: five stimuli at a fixed temperature of forty-three degrees Celsius, used as a baseline painful stimulus, and five at the temperature matched to the participant's pain threshold. Results showed that individuals in the BPD group had less overall brain activity at the fixed temperature of forty-three degrees than members of the control group. When the stimulus matched an individual's pain threshold, fMRI activity was similar between both groups; however, BPD patients did show slightly altered patterns of activity, with the medial thalamus being activated bilaterally rather than contralaterally, as it was in the control group.

Although all members of the BPD group had engaged in NSSI, this study only tested individuals with BPD and, therefore, the results may not be indicative of all individuals who engage in this behavior. Still, this study suggests that Eisenberger's correlation between physical and emotional pain sensitivity may not apply in all situations.

Pain Perception

As it was established that individuals with BPD have more emotional reactivity as well as higher pain

thresholds than controls (Schmahl et al., 2006), studies assessing the pain thresholds of those who engage in NSSI but are not diagnosed with BPD are necessary to gain more information about the effects of NSSI on pain processing.

In the study conducted by Osuch and colleagues discussed above (Osuch et al., 2014), the NSSI and control groups rated pain almost exactly the same, both reporting more pain in the experimenter-administered scenario. This suggests that people who engage in NSSI do not have altered pain perception, but that their brains do react differently to physical pain than controls. Another study examining individuals who engage in NSSI, but were not diagnosed with any other specific psychiatric disorder, showed that participants in both the control and NSSI groups rated an electrical stimulus similarly in terms of least possible perceptible stimulus and greatest possible endurable painful stimulus (Bonenberger et al., 2015). The groups also showed similar activity levels in the posterior insula, the area of the brain associated with perceiving and discriminating intensities of painful stimuli. Where the groups differed was in the affective perception of pain. In the NSSI group, bilateral anterior insula activity was significantly reduced. NSSI participants showed the same level of activation in this area across all stimulus intensities, where control groups showed modulated activity dependent upon intensity. These studies suggest that it is not the perception of pain that is altered in individuals who engage in NSSI, but rather the affective differentiation of pain.

Physical Syndromes Related to NSSI

Certain physical syndromes have been shown to correlate with NSSI. Further investigation of this correlation may give insight into the neural properties of this behavior. Studies have shown that both migraine headaches and epilepsy correlate positively with NSSI (Colman et al., 2016; Singhal et al., 2014). Individuals who suffer from migraine headaches are almost 50% likelier to engage in NSSI than individuals without migraines. Chronic epilepsy has been related to both NSSI and suicidality, with epileptics being three times likelier to attempt suicide than those who do not suffer from the disease.

Depression and chronic pain are also known to correlate positively (Averill et al., 1996; Elliott, Renier and Palcher, 2003; Haythornthwaite, Sieber and Kerns, 1991), and a similar relationship may exist between physical illness and NSSI. However, certain physical syndromes such as cancer, sickle cell anemia, and Down's syndrome have a negative correlation to NSSI (Singhal et al., 2014). Researching why some physical syndromes seem to predispose individuals to NSSI and others seem to lessen the risk of this behavior may provide insight into the biological factors of NSSI.

Endogenous Opioids

Recent studies suggest that endogenous opioid levels may play a key role in NSSI (Bresin and Gordon, 2010; Sher and Stanley, 2008; Stanley et al., 2010). In one such study, self-harming individuals' endogenous opioid levels were measured by the extraction and analysis of cerebrospinal fluid, and then they were compared to endogenous opioid levels of a control population. Due to a correlation between decreased opioid levels, as well as NSSI behaviors, and suicide attempts, the researchers ensured that all participants in the study had made at

least one attempt; this ensured that any deficiencies in opioid levels in the NSSI group could definitively be attributed to self-injurious behavior.

Results from the cerebrospinal fluid samples showed that members of the NSSI group had significantly lower levels of β -endorphin and met-enkephalin than members of the control group. Mu-opioids have been associated with the suppression of the affective qualities of pain and negative internal affective states. Further testing is required to determine whether there is a correlation between altered ACC activity and lowered levels of endogenous opioids.

Discussion

The available literature suggests that there are significant biological differences between those who engage in NSSI and healthy controls. Individuals who self-injure have altered cortical patterns, including increased activity in the reward centers of the brain and poor connections in the areas associated with affect regulation. These individuals may also have higher pain thresholds and may experience greater negative affect reduction after instances of self-injury. Physical syndromes, such as migraine headaches and epilepsy, have been shown to correlate with NSSI, and those who engage in self-injury have lower levels of endogenous opioids than healthy controls.

While these findings provide a degree of insight into the neurophysiological factors related to NSSI, the dearth of available data precludes a deeper understanding of the disorder. Furthermore, the literature is often disconnected from the topic as a whole, instead focusing on a single specific factor related to NSSI. In addition to a lack of research on the subject, research currently being conducted on NSSI suffers from some major limitations.

One such limitation is the fact that most research on NSSI does not specify the particular form of NSSI in which the research participants have engaged. Because NSSI covers a wide variety of acts, such as cutting, bruising, and excoriating the skin, studying individuals who engage in a particular type of injurious behavior may be advantageous in determining whether certain types of NSSI correlate to different brain alterations from other forms of the behavior.

Another downside to the current research on NSSI is that many of the studies discussed in this paper relied on patients with BPD as a sample of individuals who engage in NSSI (Jovev et al., 2008; Kraus et al., 2010; Ludäscher et al., 2009; Plener et al., 2012). However, BPD patients exhibit a specific set of symptoms, and any neurological abnormalities seen in these individuals may not be directly related to NSSI. NSSI has recently been added as a separate syndrome to the fifth edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5), as opposed to merely being a symptom of other psychiatric disorders (DSM-5). Although NSSI is now considered a separate condition with its own parameters for diagnosis, little research has been conducted on individuals who engage in NSSI and are not diagnosed with any other major psychological disorder.

NSSI is most prevalent in the adolescent population, affecting 14-39% of adolescents in the general community and 40-61% of adolescents in psychiatric inpatient settings (Nock & Prinstein, 2004). Despite the prevalence of this maladaptive behavior, few studies have been conducted using adolescents as subjects. This lack of research may be attributed to complications involved in using these individuals as research subjects, such as the issue of obtaining parental consent. However,

studying adolescents who engage in NSSI, both with and without other psychological disorders, and comparing the data, between these groups and to earlier studies, could provide valuable insight into NSSI in relation to adolescent brain development. Therefore, despite barriers that researchers may face in obtaining appropriate and representative samples of the population, more research would provide insight into the neuropathology of individuals who engage in NSSI.

Another type of research that would be instrumental in understanding the mechanisms of NSSI, but has scarcely been employed, is a longitudinal study. A study involving adolescents who engage in NSSI revealed a negative correlation between age and relief after NSSI (Osuch et al., 2014). This finding demonstrates the prevalence of NSSI in adolescence and is congruent with the fact that many people abandon this behavior in their twenties. One longitudinal study of adolescents beginning at a mean age of 15.9 years and ending at a mean age of 29 years showed a significant decrease in self-harm over time (Moran et al., 2012). The study also showed a correlation between NSSI and alcohol use, cannabis use, and cigarette-smoking, suggesting that, although NSSI decreased over time, individuals who have engaged in NSSI may continue to engage in self-destructive behavior throughout their lives. More longitudinal research could strengthen this claim and give insight into the relationship between NSSI and other forms of addiction. Taken together, a combination of longitudinal and fMRI studies would yield more data on the effects of NSSI over a period of time. Studies using fMRI have shown altered activity in the ACC and other regulatory pathways of patients who engage in NSSI as much as two years after their latest incident of this behavior (Bonenberger et al., 2015). If longitudinal studies were conducted using fMRI,

more information could be gathered about the abnormal brain activity of individuals who engage in NSSI. If abnormal brain activity ceased once this behavior was discontinued, a causal claim could be made about the relationship between NSSI and abnormal brain activity. However, if abnormal brain activity persisted after the discontinuation of NSSI, particularly in the dorsal stratum, this finding would support the finding that those who engaged in NSSI were likelier to show addictive behaviors later in life; which suggests the use of alternative coping mechanisms rather than the complete abolition of regulatory impairments and other brain abnormalities (Moran et al., 2012).

Conclusion

More investigation into the biological factors related to NSSI and the populations affected by this syndrome is imperative to the advancement of knowledge and treatments for NSSI. Through a combination of self-reports and fMRI techniques, more information can be gathered regarding the brain abnormalities. Although this combination of research approaches has become more prevalent in recent years, even more studies using this methodology that specifically examine the adolescent population and individuals over time are necessary to gain an understanding of NSSI. The combination of self-reports and fMRI provides a deeper and more holistic understanding of the neurological processes and affective aspects of this syndrome than any other study design that has been utilized thus far. By studying adolescents, more data can be collected on the most representative population for this syndrome; combining the study of adolescents with the study of those who engage solely in NSSI, separate from other psychological disorders, could enrich our understanding of this syndrome.

Acknowledgements

A special thanks to Dr. Evans, Dr. Kongsamut, and the entire RISE program at Drew University for encouraging and facilitating student research.

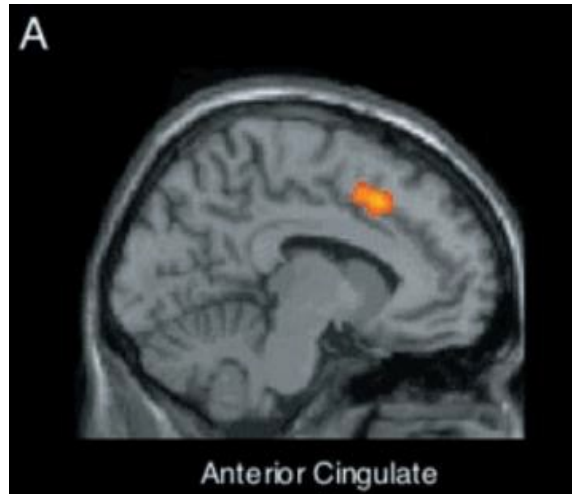


Figure A: The anterior cingulate cortex is associated with error detection, conflict monitoring, and emotional regulation. Abnormalities in the anterior cingulate cortices of individuals who engage in non-suicidal self-injury may explain the emotional dysregulation characterized by the syndrome (Eisenberger, Lieberman, & Williams, 2003).

References

- Dsm-IV*. DSM Library: American Psychiatric Association, 2000. Print.
American Psychiatric Association., and American Psychiatric Association. DSM-5 Task Force. *Diagnostic and Statistical Manual of Mental Disorders : Dsm-5*. Fifth edition. ed. Arlington, VA: American Psychiatric Association, 2013. Print.
- Averill, Patricia M., et al. "Correlates of Depression in Chronic Pain Patients: A Comprehensive Examination." *Pain* 65.1 (1996): 93-100. Print.
- Bonenberger, Martina, et al. "Differential Neural Processing of Unpleasant Haptic Sensations in Somatic and Affective Partitions of the Insula in Non-Suicidal Self-Injury (Nssi)." *Psychiatry Research: Neuroimaging* 234.3 (2015): 298-304. Print.
- Bresin, Konrad, and Kathryn H. Gordon. "Endogenous Opioids and Nonsuicidal Self-Injury: A Mechanism of Affect Regulation." *Neuroscience & Biobehavioral Reviews* 37.3 (2013): 374-83. Print.
- Bresin, Konrad, et al. "No Pain, No Change: Reductions in Prior Negative Affect Following Physical Pain." *Motivation and Emotion* 34.3 (2010): 280-87. Print.
- Chapman, Alexander L., Kim L. Gratz, and Milton Z. Brown. "Solving the Puzzle of Deliberate Self-Harm: The Experiential Avoidance Model." *Behaviour Research and Therapy* 44.3 (2006): 371-94. Print.
- Colman, I., et al. "Migraine Headache and Risk of Self-Harm and Suicide: A Population-Based Study in Ontario, Canada." *Headache* 56.1 (2016): 132-40. Print.
- DeWall, C. Nathan, et al. "Acetaminophen Reduces Social Pain:

- Behavioral and Neural Evidence." *Psychological Science* 21.7 (2010): 931-37. Print.
- Eisenberger, N. I. "Broken Hearts and Broken Bones: A Neural Perspective on the Similarities between Social and Physical Pain." *Current Directions in Psychological Science* 21.1 (2012 a): 42-47. Print.
- Eisenberger, N.I. "Meta-Analytic Evidence for the Role of the Anterior Cingulate Cortex in Social Pain." *Social Cognitive and Affective Neuroscience* 10.1 (2015): 1-2. Print.
- Eisenberger, N. I. "The Pain of Social Disconnection: Examining the Shared Neural Underpinnings of Physical and Social Pain." *Nature Reviews Neuroscience* 13.6 (2012 b): 421-34. Print.
- Eisenberger, N. I., and M. D. Lieberman. "Why It Hurts to Be Left Out: The Neurocognitive Overlap between Physical and Social Pain." *The Social Outcast: Ostracism, Social Exclusion, Rejection, and Bullying*. Eds. Williams, K. D., J. P. Forgas and W. von Hippel. Sydney Symposium of Social Psychology Series. New York, NY, US: Psychology Press, 2005. 109-27. Print.
- Eisenberger, N.I. "Why Rejection Hurts: A Common Neural Alarm System for Physical and Social Pain." *Trends Cogn Sci* 8.7 (2004): 294-300. Print.
- Eisenberger, N. I., M. D. Lieberman, and K. D. Williams. "Does Rejection Hurt? An Fmri Study of Social Exclusion." *Science* 302.5643 (2003): 290-92. Print.
- Elliott, Thomas E., Colleen M. Renier, and Jeanette A. Palcher. "Chronic Pain, Depression, and Quality of Life: Correlations and Predictive Value of the Sf-36." *Pain Medicine* 4.4 (2003): 331-39. Print.

- Glenn, Catherine R., and E. David Klonsky. "Nonsuicidal Self-Injury Disorder: An Empirical Investigation in Adolescent Psychiatric Patients." *Journal of Clinical Child & Adolescent Psychology* 42.4 (2013): 496-507. Print.
- Haines, Janet, et al. "The Psychophysiology of Self-Mutilation." *Journal of Abnormal Psychology* 104.3 (1995): 471-89. Print.
- Haythornthwaite, Jennifer A., William J. Sieber, and Robert D. Kerns. "Depression and the Chronic Pain Experience." *Pain* 46.2 (1991): 177-84. Print.
- Jovev, M., et al. "An Mri Study of Pituitary Volume and Parasuicidal Behavior in Teenagers with First-Presentation Borderline Personality Disorder." *Psychiatry Res* 162.3 (2008): 273-7. Print.
- Klonsky, E. David. "The Functions of Deliberate Self-Injury: A Review of the Evidence." *Clinical Psychology Review* 27.2 (2007): 226-39. Print.
- Kraus, A., et al. "Script-Driven Imagery of Self-Injurious Behavior in Patients with Borderline Personality Disorder: A Pilot Fmri Study." *Acta Psychiatrica Scandinavica* 121.1 (2010): 41-51. Print.
- Kross, Ethan, et al. "Social Rejection Shares Somatosensory Representations with Physical Pain." *Proceedings of the National Academy of Sciences* 108.15 (2011): 6270-75. Print.
- Ludäscher, P., et al. "A Cross-Sectional Investigation of Discontinuation of Self-Injury and Normalizing Pain Perception in Patients with Borderline Personality Disorder." *Acta Psychiatrica Scandinavica* 120.1 (2009): 62-70. Print.

- Masten, Carrie L., Sylvia A. Morelli, and Naomi I. Eisenberger. "An Fmri Investigation of Empathy for 'Social Pain' and Subsequent Prosocial Behavior." *NeuroImage* 55.1 (2011): 381-88. Print.
- Moran, Paul, et al. "The Natural History of Self-Harm from Adolescence to Young Adulthood: A Population-Based Cohort Study." *The Lancet* 379.9812 (2012): 236-43. Print.
- Nock, Matthew K., and Mitchell J. Prinstein. "A Functional Approach to the Assessment of Self-Mutilative Behavior." *Journal of Consulting and Clinical Psychology* 72.5 (2004): 885-90. Print.
- Ogawa, Seiji, and Yul-Wan Sung. "Functional Magnetic Resonance Imaging." *Scholarpedia* 2.10 (2007): 3105. Print.
- Osuch, Elizabeth, et al. "Functional Mri of Pain Application in Youth Who Engaged in Repetitive Non-Suicidal Self-Injury Vs. Psychiatric Controls." *Psychiatry Research: Neuroimaging* 223.2 (2014): 104-12. Print.
- Plener, Paul L., et al. "Prone to Excitement: Adolescent Females with Non-Suicidal Self-Injury (Nssi) Show Altered Cortical Pattern to Emotional and Nss-Related Material." *Psychiatry Research: Neuroimaging* 203.2-3 (2012): 146-52. Print.
- Schmahl, C., et al. "Neural Correlates of Antinociception in Borderline Personality Disorder." *Archives of General Psychiatry* 63.6 (2006): 659-66. Print.
- Sher, Leo, and Barbara H. Stanley. "The Role of Endogenous Opioids in the Pathophysiology of Self-Injurious and Suicidal Behavior." *Archives of Suicide Research* 12.4 (2008): 299-308. Print.

Singhal, Arvind, et al. "Risk of Self-Harm and Suicide in People with Specific Psychiatric and Physical Disorders: Comparisons between Disorders Using English National Record Linkage." *Journal of the Royal Society of Medicine* 107.5 (2014): 194-204. Print.

Stanley, Barbara, et al. "Non-Suicidal Self-Injurious Behavior, Endogenous Opioids and Monoamine Neurotransmitters." *Journal of Affective Disorders* 124.1-2 (2010): 134-40. Print.

Wolff, Jennifer C., et al. "Negative Cognitive Style and Perceived Social Support Mediate the Relationship between Aggression and Nssi in Hospitalized Adolescents." *Journal of Adolescence* 37.4 (2014): 483-91. Print.

**Mineral Dust Aerosols as Feedback Mechanism Drivers for
Global Desertification
Zoe Coates Fuentes (CLA 2018)**

Abstract

Discussions on habitat modification due to anthropogenic causes have risen exponentially as coral reefs suffer from drastic bleaching and former fertile regions become increasingly arid. Although the discussion on desertification has been dominated by concerns of reduced land productivity, one of the less-mentioned side-effects of soil nutrient depletion has been the production of mineral dust aerosols. These aerosols are particulates that exist within the atmosphere due primarily to dust storms in arid regions. Given their high abundance, they have significant impacts on atmospheric chemistry, human health, climate, and biogeochemical cycles. Throughout this paper I argue that the increase of desert regions throughout the world has spurred an influx of aerosols into the atmosphere, which is of great concern, given these aerosol's effect on global temperature changes via radiative forcing and pollution. Furthermore, as their atmospheric abundance increases, aerosols will be correlated not only with exacerbated climate change through positive feedback mechanisms but also with reduced land fertility.

Introduction

Desertification, or the spread of infertile and arid ecosystems, is a problem that brings with it severe implications not only for ecosystem stability and climate change rates, but also for international relations and socio-economic stability. Among the direct implications of encroaching deserts is reduced land productivity for agriculture, which has severe effects on some region's economy and overall survival. However, one of the less discussed issues is the production of mineral dust aerosols, an issue that presents dramatic ramifications for the nature of climate change in the years to come. Despite their role within climate modification, current international policy fails to address the broader influence of these aerosols in both atmospheric chemistry and climate change, partially because of the level of uncertainty that is associated with their generalized effects. Throughout this paper, we will argue that desertification may be correlated with an increased influx of mineral aerosols in the atmosphere, which in turn acts as a driver for feedback mechanisms that exacerbate the rates of desertification worldwide.

To exemplify increased desertification and possible routes toward its mitigation, we will be analyzing China's current approach toward combating encroaching arid regions. In particular, we will focus on their implementation of the framework established in the United Nations Convention on Combatting Desertification (UNCCD) through comprehensive law reforms, technological developments, and local leadership initiatives. This analysis of their approach will allow for a better understanding of how other countries can tackle the reduction of land degradation and foster more sustainable social and agricultural practices on their terrain.

Mineral Aerosols: Positive Feedback Mechanism Drivers in Climate Change

Desertification refers to the spread of infertile terrain, usually through either "desert creep" or the degradation of farmland due to unsustainable land practices (Kassas et al., 1995). More

specifically, this deterioration of fertile terrain results from “excessive human exploitation that oversteps the natural carrying capacity of the land” and the “inherent ecological fragility of the [ecosystem]” (Kassas et al., 1995). The combination of both these factors yields a strong negative influence over the sustainability of a land’s fertility. Of particular concern is the fragility of the ecosystem, which is highly dependent on water availability, plant coverage, and organic matter content in the soil. When all of these are at a low, added overexploitation of the land from socio-economic drivers (primarily changes in agricultural patterns) leads to further loss of nutrients, soil erosion, and organic matter, which are instrumental for vegetation growth and consequently animal survival (Kassas et al., 1995).

With reduced land fertility comes a number of issues, both societal and ecological. At a societal level, lower land fertility yields a scarcity of resources, thus spurring possible conflicts over land and water use as well as mass migrations to areas that are less affected. One region that is particularly impacted by these issues is the West African Sahel, where nomadic cattle herders have started grazing on farmer’s ancestral land, spurring a wave of intense conflict. This emerging tension has fed on pre-existing antagonism between the nomadic Fulani and the sedentary local farmers, exacerbating the rate of violence throughout the region (Oduah, 2016). Simultaneously, at an environmental level, reduced fertility and poor farming practices translate to higher levels of erosion and exposed topsoil, which, when combined with strong winds, results in an influx of suspended aerosols in the atmosphere.

The concept of an aerosol spans a number of different species, anything from desert dust to smoke from biomass burning and air pollution. They are suspended particulates that are free to interact with other chemicals and agents within the atmosphere, such as other pollutants, solar radiation, and greenhouse gases. Depending on their size and chemistry, aerosols are able to remain in the atmosphere for longer or shorter periods of time,

which also affects their overall effect on atmospheric regulation and balance. The longer a particulate's lifetime in the atmosphere, the more likely it is to react and contribute to atmospheric conditions. Though multiple types of aerosols have been shown to affect atmospheric interactions, this paper focuses on mineral dust aerosols, which are direct products of increased desertification and the largest mass contributors to the atmosphere (Buseck and Posfai, 1999).

Mineral aerosols are particulates that exist within the atmosphere due primarily to dust storms in arid regions; as greater wind comes in contact with dry soil, the influx of these dusts in the lowest layer of the atmosphere, the Troposphere, sees a drastic increase (Woodward et al., 2005). In the year 2003 alone, the amount of mineral dust loading into the atmosphere was estimated at an average of 1000-3000 Tg per year, greatly depending on the location being studied and the season, with the Saharan Desert acting as the largest contributor globally (IPCC, 2001). The likelihood of any given soil being ejected into the atmosphere is highly dependent on various factors, among them: wind velocity, physical characteristics of the soil such as moisture and cohesiveness, and surface characteristics of the terrain, including vegetation coverage. Usher et al. cites that human activities, including "improper agricultural and grazing practices, can account for 20-50%" of the dust influx in the atmosphere, while "expanding [the] desertification of land" (2003). Furthermore, with continued anthropogenic drivers, atmospheric mineral dust loading is expected to increase by about 10% before the year 2100 (IPCC, 2001).

At present, most dust storms occur "on the west coast of North Africa and extending through the Middle East and into Central Asia," with 90% of the dust generation currently occurring in the Northern Hemisphere (Usher et al., 2003). However, current models of aerosol loading predict an increase by up to a factor of ten, with new source areas expanding primarily in the Southern Hemisphere, including the southern Pacific, South America, and Australia (Woodward et al., 2005). Woodward et al.'s model, for

instance, argues for the combined effect of fluctuating vegetation coverage and changes in climate, which together increase the likelihood of dust loading through intense winds. Furthermore, changes in vegetation that make dust loading more likely also highlight a change in localized wind speed and precipitation, as forested areas tend to mitigate the harsh effects of strong winds.

Once in the Troposphere, mineral dusts have significant impacts on atmospheric chemistry, human health, climate, and biogeochemical cycles. Moreover, as they participate in environmental processes, they can exacerbate the effect of current climate change drivers through feedback mechanisms such as dieback of vegetation and animals, variable rainfall, and increased fluctuations in convection patterns through a changing albedo. We will be examining these three feedback mechanisms throughout the remainder of this section.

Atmospherically, mineral dusts can serve as the reactive site for numerous species, thereby serving as a possible sink and source of compounds such as ozone, which is associated with high levels of Tropospheric pollution. Additionally, they can serve as possible sites for hazardous chemicals, organisms, or allergens to adsorb, which in turn causes a rise in the spread of disease and a spike in allergic reactions. Furthermore, depending on the size of a particular aerosol, it can cause “respiratory damage [as] they penetrate deep into the alveoli of the human lungs, produce scarring, and potentially lead to conditions such as emphysema” (Usher et al., 2003). Their inherent ability to serve as a surface for different compounds to adsorb allows them to carry possibly toxic species to areas where large populations will be exposed.

Mineral dust deposition in ocean waters can also exert a significant role over oceanic conditions and marine life. For instance, soluble iron compounds can be deposited into the oceans through fluctuating wind patterns. When dissolved, these compounds serve as a source of nutrients for microorganisms such as phytoplankton, which have been connected with algal blooms or red tides (Usher et al., 2003). This correlation between

iron compound deposition and algal blooms was documented by Walsh and Steidinger in *Saharan dust and Florida red tides: The cyanophyte connection* (2001). They highlighted an increase in phytoplankton species, particularly *Gymnodinium breve*, as a function of increasing Saharan dust in the Gulf of Mexico. This particular species is responsible for toxic, large-scale algal blooms that have measurable effects on: fish and shellfish mortality, respiratory conditions for humans exposed, water discoloration, and reduced tourism to affected areas (Walsh and Steidinger, 2001).

Aside from the larger direct implications to human and animal health, mineral dust may also exert drastic effects on radiative forcing and ecosystem survival. In general, atmospheric aerosols can have either net heating or cooling effects, depending greatly on size and physical characteristics (Usher et al., 2003). As an example, sulfates have a high albedo or reflectivity, which translates to a higher rate of radiation being reflected away from the atmosphere. For this reason, sulfates have been researched as a possible solution for increased atmospheric temperatures through geoengineering. On the opposite end of the spectrum, however, we find particles such as soot, which absorb radiation and therefore increase the temperature of the atmosphere around them. When it comes to mineral dusts, their overall reflectivity depends greatly on their size. Smaller particulates tend to have greater albedo both directly, reflecting radiation back into outer space, and indirectly, through increasing the lifetime of clouds, which have a high albedo (Buseck et al., 1999). In contrast, larger particulates tend to have lower reflectivities and therefore are more likely to have net heating effects. These particulates also remain in the atmosphere for shorter periods of time, as they are too heavy.

As argued in Woodward et al., mineral dusts overall possess a scattering albedo which could have a net cooling effect in the Troposphere and near the surface of the terrain. However, they argue that this may actually produce a positive feedback, or increased effect, in processes such as desertification. The main

reason for this is that mineral dusts remain in the Troposphere, having a net cooling effect, while temperatures in higher levels of the atmosphere are relatively unaffected. As the difference in temperatures between the surface of the terrain and the atmosphere increases, there is a shift in convection and wind patterns, which further intensifies dust production. (Woodward et al., 2005)

On top of being able to exert an effect over atmospheric temperatures, aerosols often act as cloud formation nuclei, possibly altering precipitation patterns (Buseck et al., 1999). As Buseck et al. explains, “clouds with more and smaller droplets are less prone to rain and drizzle formation,” which allows them to remain in the atmosphere for longer, thus exerting a cooling effect. That being said, extended time in the atmosphere also means that there is reduced rainfall in those regions (Buseck et al., 1999). Rosenfeld et al. analyzed the effect of Saharan dust on cloud condensation and they concluded that, given the small size of the particles in the atmosphere, they are able to condense water droplets. Nevertheless, due to the scattered condensation as a function of multiple nuclei, these droplets are unable to precipitate quickly, as Buseck et al. concluded in their research.

Overall, mineral dusts exhibit a strong effect over environmental and atmospheric conditions, particularly through their participation in feedback mechanisms that exacerbate organism death, atmospheric convection patterns, and reduced rainfall. For this reason, possible routes of suppression or mitigation are necessary to reduce the impact of mineral dusts on further desertification and loss of arable land.

China: A Case Study of Desertification Mitigation

Despite a strong need for solutions, as well as numerous initiatives to address the problem, such as the UN Convention to Combat Desertification (UNCCD), little has been done toward a cohesive international scientific agreement to correct and mitigate the effects of desertification worldwide. Greenhouse gases have been prioritized in numerous international

agreements for their net heating of the atmosphere, among them the 1997 Kyoto Protocol, the Montreal Protocol of 1989 and its corollary, the Kigali deal, and the 2016 Paris Climate Change Agreement. Nonetheless, anthropogenic aerosol emissions have remained largely unaddressed. This is of particular concern given their possible role in the feedback mechanisms mentioned throughout this paper (i.e. changes in convection patterns, organism death, and reduced rainfall).

In light of the increase in desertification, the UNCCD, originally held in 1994, presented a strategy for tackling the encroaching aridity through a wide range of approaches. Although the strategy focused on African nations, the convention provides a framework for all nations affected to follow in their development of national initiatives through “prevention and/or reduction of land degradation,” “rehabilitation of partly degraded land,” and “reclamation of desertified land,” while encouraging local initiatives (UNCCD, 1994). We find China among the countries that have attempted to implement some of these solutions. In recent years, China has started dominating global industry and has thereby become a major player in international markets and politics.

Throughout the last 20 years, the Chinese government has cooperated with a number of international and local institutions in monitoring and evaluating the level of desertification in China. Among the most reliable of these institutions is the Chinese Committee for Implementing UN Convention to Combat Desertification (CCICCD), a government-organized committee that relied upon a bioclimatic map developed by the UNCCD to assess the aridity of Chinese terrain. A great deal of their time and effort has been allocated to the assessment of current desert regions within China, culminating in the map presented in Figure 1. Throughout their studies, they have evaluated the effects of wind, water, and other processes on land erosion, focusing on different terrains and vegetation coverage. For their evaluation, the Chinese government based themselves on four main indicators of aridity: “percentage of deflation land or mobile

sand”, annual mean expanding rate of said land or sand, annual decreasing rate in land productivity, and change in vegetation coverage (Yang et al., 2005). This extensive research has allowed them to determine that the total desertification area in China is 2.622 million km², with an expanding rate of 2460 km² per year (Yang et al., 2005).

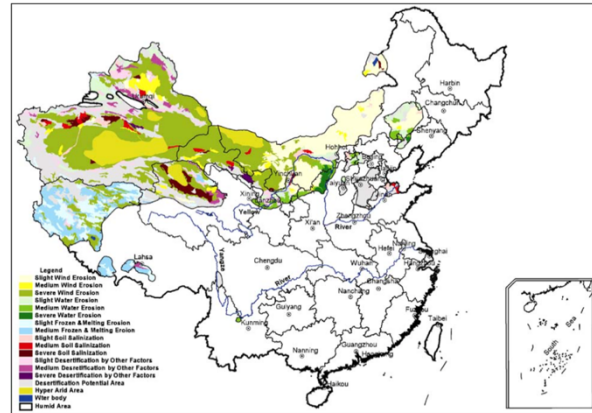


Figure 1. Desertification distribution map of China redrawn from CCICCD (1997) (Yang et al., 2005).

Given the rapidly encroaching desert and the fact that 60% of their population lives in affected areas (CICCD, 1996), the Chinese government has started concerted efforts to mitigate the effect of this changing ecosystem on the Chinese population. Throughout the process, the Chinese government has cooperated with the UNCCD to implement a strategic three-phase plan setting goals for each stage’s set time frame. Among their strategies are “[relocating] residents, [planting] trees, and [limiting] herding”, the combination of which has led to slowed and even reversed desert growth in certain areas (Haner et al., 2016). Primarily, however, the Chinese government has prioritized the establishment of protected “ecological and economic forests” and reforestation in viable regions (CICCD, 1996).

In addition to the direct short-term methods of alleviating the effects of desertification on the Chinese people, the government

has created a program to target possible long-term consequences, relying on comprehensive law reforms, technological development, and local leadership initiatives. Among their law changes lie alterations to current forest, grassland, mineral resources, land management, water and soil conservation, and agriculture laws (CICCD, 1996). Their primary goal in amending these laws is to reduce land practices that overuse the soil and consume excess nutrients, thereby hoping to curtail soil erosion, improve land productivity, maintain water source availability, and limit species extinction (Fu, 2000). As a complement to these reforms, the Chinese government has followed the UNCCD's push for local leadership initiatives and technological consulting, which consist in developing and implementing more sustainable practices for agriculture and land tenure through partnerships with regional organizations. At the root of both the UNCCD and the CICCD's strategic plan is the realization that both implementation of better technologies and enforcement of law reforms rely strongly on close connections with local leaders that are better suited to drive change in remote communities.

Another important application of the UNCCD that the Chinese government has stressed is the incorporation of research and education into a cohesive government movement. In particular, the Chinese government has pushed for the establishment of more scientific research institutes, enhancement of current facilities, and incorporation of the research obtained into new national policies. Moreover, they have encouraged improvement in current agriculture and forestry programs at colleges and universities so that recent graduates understand the extent of modern technology and its benefits in counteracting land degradation (CICCD, 1996).

Through an integration of all three approaches and cooperation with the UNCCD, the Chinese seek to reduce erosion and maintain moist soil, thereby reducing the risk of contributing further to mineral aerosols in the atmosphere, seeing as they would exacerbate desertification through aerosols' multiple feedback mechanisms. In order to fulfill their goals, two of their

primary mechanisms of action are large-scale foresting campaigns and pasture enclosure. The former, when coupled with farmland development, entails the planting of vegetation in areas where desertification is starting to display itself while encouraging sustainable farming practices that reduce population pressure or overexploitation on the soil. The latter refers to the closing off of grazing areas in order to allow the terrain to recover from grazing naturally, thereby containing the effect of cattle on the soil (Shengyue and Lihua, 2001).

Among the Chinese government's largest projects is the "Green Wall of China," an initiative to plant vegetation as a barrier to stop encroaching deserts. This approach seeks to increase "human-made tree cover from five per cent to 15 per cent of the country's [...] landmass," which, aside from maximizing the intake of carbon dioxide from the atmosphere, limits current deserts' ability to spread into neighboring soils (Alvarez Tudela, 2012). By planting more vegetation, the soil is kept together, leaving less terrain exposed to violent winds.

Overall, the Chinese government, in collaboration with local leaders, has been able to counteract the effects of desertification on Chinese territory through concerted plans to increase vegetation coverage while educating their population on proper farming techniques. Moreover, their approach to ameliorating the problem has consisted of both short-term and long-term initiatives that mitigate current damage and reverse pre-existing desertification. As Luc Gnacadja, the executive secretary of the UNCCD, stated: China has set a precedent for combatting desertification, teaching the world that "restoration of even a severely degraded world is possible," "that the role of government [in] creating the right policy incentives [...] is indispensable," and that approaches to land tenure must be drastically modified to account for the current global climate (Alvarez Tudela, 2012).

Though the Chinese have already implemented significant changes to regulate desertification within their country, there are

still areas that require further action, as highlighted by Fu: establishing “institutional frameworks for combating desertification”, perfecting enforcement and assessment mechanisms, researching into technologies that reduce dependence on unsustainable land practices, increasing the public’s education on the issue, and developing a more cohesive international approach to dealing with these situations (Fu, 2000). Furthermore, scarcity of funds invested in the initiative has restricted progress in such a way that the breadth of programs available cannot be applied.

Conclusion

Throughout the debate on climate change, discussion on desertification has been dominated by concerns of reduced land productivity for farming and communities’ livelihoods. However, the production of mineral dust aerosols and their role on atmospheric chemistry have been largely ignored. These aerosols, which are a direct result of dust storms in arid regions, exert a strong effect on atmospheric chemistry, animal and plant health, climate, and biogeochemical cycles. Furthermore, they exhibit significant contributions to environmental and atmospheric conditions through their participation in feedback mechanisms that exacerbate convection patterns, organism death, and reduced rainfall. Throughout this paper, I have argued that the increase of desert regions throughout the world has spurred an influx of aerosols into the atmosphere, leading to heightened radiative forcing in the atmosphere through greater cloud formation, mineral dust albedo, reduced precipitation, and exacerbated weather patterns that increase the release of dust in the Troposphere.

Given the dramatic consequences of both desertification and mineral dust in the atmosphere, current approaches to combating land degradation have focused on minimizing soil erosion through improved agricultural practices, as well as reforestation campaigns. China, as both a key player in international politics and a country severely affected by desertification, has been a leader in desertification mitigation, relying primarily on

comprehensive law reform, the improvement of farming technology, and the promotion of local involvement in large-scale mitigation initiatives.

References

- Alvarez Tudela, R. "Fighting Desertification in China". *Aljazeera.com*. 2012. [Online]
- Buseck, P. R; Posfai, M. "Airborne minerals and related aerosol particles: Effects on climate and the environment". *Proceedings of the National Academy of Science*. 1999, p. 3372-3379.
- CICCD. "China National Committee for the implementation of the UN Convention to Combat Desertification". *United Nations: China*. 1996.
- Fu, C. "Combating Desertification: Action Programme in China." *Royal Swedish Academy of Science*. 2000, p. 359-361.
- Haner, J.; Wong, E.; Watkins, D.; White, J. "Living In China's Expanding Deserts." *Nytimes.com*. 2016. [Online]
- IPCC. "Climate Change 2001: The Scientific Basis. Contribution of Working group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change." *Cambridge University Press: Cambridge, U.K.* 2001.
- Kassas, M. "Desertification: a general overview". *Journal of Arid Environments* 1995, p. 115-128.
- Nicholson, S. E; Tucker, C. J.; Ba, M. B. "Desertification, Drought, and Surface Vegetation: An Example from the West African Sahel." *Bulletin of the American Meteorological Society*. 1998, p. 815-829.
- Oduah, C. "Nigeria: Deadly nomad-versus-farmer conflict escalates". *Aljazeera.com*. 2016. [Online]
- Prospero, J.M. "Long-term measurements of transport of African mineral dust to the southeastern United States:

- Implication for regional air quality". *Journal of Geophysical Research*. 1999, p. 15917-15927.
- Rosenfeld, D.; Rudich, Y.; Lahav, R. "Desert dust suppressing precipitation: A possible desertification feedback loop." *Nature*. 1996, p. 681-683.
- Shengyue, F.; Lihua, Z. "Desertification Control in China: Possible Solutions". *Ambio: A Journal of the Human Environment*. 2001, p. 384-385.
- Sokolik, I.M; Toon, O.B. "Direct radiative forcing by anthropogenic airborne mineral aerosols." *Nature*. 1996, p. 681-683.
- UNCCD. "United Nations Convention to Combat Desertification in Those Countries Experiencing Serious Drought and/or Desertification, Particularly in Africa." *United Nations*: Paris, France. 1994.
- Usher, C. R; Michel, A.E; Grassian, V.H. "Reactions on Mineral Dust." *Chemical Reviews*. 2003, p. 4883-4939.
- Walsh, J.J.; Steidinger, K.A. "Saharan dust and Florida red tides: The cyanophyte connection." *Journal of Geophysical Research*. 2001, p. 11597.
- Woodward, S.; Roberts, D. L.; Betts, R. A. A simulation of the effect of climate change-induced desertification on mineral dust aerosol. *Geophysical Research Letters*. 2005. [Online]
- Yang, X.; Zhang, K.; Jia, B.; Ci, L. "Desertification assessment in China: An Overview." *Journal of Arid Environments*. 2005, p. 517-531.

Zhang, X. Y; Gong, S.L.; Zhao, T. L.; Arimoto, R. Wang, Y. Q;
Zhou, Z. J. "Sources of Asian dust and role of climate
change versus desertification in Asian dust emission."
Geophysical Research Letters. 2003, p. 2272. [Online]

**What About Bambi? An Examination of the Evidential
Problem of Evil
Shira Mindy Newman (CLA 2016)**

Abstract

Does the existence of pointless or meaningless evil provide us with enough evidence to conclude that God does not exist? This question, posed by the philosopher William Rowe, is the basis for the evidential argument of evil. According to Rowe, there appears to be instances of pointless suffering, such as a fawn suffering horrible burns before death, which achieve no good or prevent no worse evil. This leads him to the contentious conclusion that God does not exist. Stephen Wykstra, a skeptical theist, objects to Rowe's claim that God does not exist on the basis that we are in no position to know God's choices. In this paper, I aim to show that Wykstra's skeptical response to Rowe's evidential argument is valid; from this we can conclude that the alleged problem of evil is not a problem for theists. I achieve this goal by creating a logical proof to demonstrate a potential flaw in Wykstra's argument and defend Wykstra's position using the works of other philosophers. From this I conclude that the existence of natural or pointless evils does not provide us with enough evidence to conclude that God does not exist.

A Brief History of Evilⁱ

Why do bad things happen to good people? It is natural to begin to wonder if there is a higher being determining our fate. Inevitably, the question, known as the evidential problem of evil, takes on a far more disconcerting form: How can an omnipotent (all-powerful), omniscient (all-knowing) and omnibenevolent (wholly-good) God exist while there is evil in the world? Natural evils, such as tsunamis, earthquakes, and volcanic eruptions, exist independently from human free will because they are produced by nature. Could an all powerful, wholly good God prevent natural evils? If so, theists appear to be faced with a genuine problem.

The philosopher David Hume cites Epicurus as one of the first philosophers to clearly state the problem (Hickson 8). His formulation, a trilemma known as the “Epicurean paradox” or the “Riddle of Epicurus,” is as follows:

Is God willing to prevent evil, but not able? Then he is not omnipotent.

Is he able, but not willing? Then he is malevolent.

Is he both able and willing? Then whence cometh evil?

Is he neither able nor willing? Then why call him God (8)?

Epicurus presents this paradox to show that God’s omnipotent and goodness is inconsistent with the existence of evil. Because omnipotence is one of the fundamental characteristics attributed to God, the rejection of God’s omnipotence challenges the existence of God. Similarly, the second premise maintains that God must be wholly good. If God has the power to prevent evil but chooses not to, God is not omnibenevolent and, therefore, does not exist.

Refutations of the Epicurean paradox begin to occur within a Christian framework. In accordance with the Bible, Saint Augustine of Hippo maintains that God created everything in the world *except for evil*. Evil for Augustine is the *privation boni* or “the privation of good” and therefore cannot be created. Saint Thomas

Aquinas uses Augustine's perception of evil to develop his own argument on the existence of natural evils. He characterizes God's perfection as *actus purus* or a pure act, stating that God can never be lacking in any respect including wisdom or goodness (Reichenbach 131). Because God is absolutely perfect, it is logically impossible for God to sin or create evil. Bruce Reichenbach incorporates this idea in his book, *Evil and a Good God*, in which he presents an argument that he believes exonerates God for the existence of natural evils. In his book, Reichenbach claims that we must believe that God, being the essence of perfection, has an unknown morally sufficient reason for creating a world where natural evils exist.

Philosopher William Rowe, however, will come to pose the modern form of the problem under discussion: Does the existence of pointless or meaningless natural evils in the world provide us with enough evidence to conclude that an omniscient (all-knowing), omnipotent (all-power), and omnibenevolent (all-good) God does not exist? This question stems from natural disasters that seemingly achieve no greater good and prevent no worse evil from occurring. Although Rowe is credited with the development of this modern form of the problem of evil, it is important to note that Saint Thomas Aquinas was the first to discuss the religious implications of the existence of natural evils. Rowe argues that this evidential problem of evil gives us enough evidence to doubt God's existence, whereas Stephen Wykstra, a skeptical theist and one of Rowe's contemporaries, objects to the claim that God does not exist. For Wykstra, we are in no position to understand God's choices and thereby cannot conclude that God's existence should mitigate "evil" in the world. My aim in this paper is to show that Wykstra's skeptical response to Rowe's evidential argument is robust, leading to the conclusion that the alleged problem of evil is not a problem for theists.

The Evidential Problem of Evil

The evidential problem of evil, otherwise known as the inductive argument from evil or the *a posteriori* argument from evil, centers on the question: Does the existence of pointless or meaningless evils in the world provide us with enough evidence to conclude that it is probable that God does not exist? Formulated by William Rowe, this argument may be rewritten as follows:

Proposition₁: There appears to be instances of pointless or meaningless evils in the world.

Proposition₂: An omnipotent and supremely good God would not allow pointless or meaningless evils to exist.

Conclusion: An omnipotent and supremely good God does not exist (Peterson 133-137).

In order to understand how he arrives at this highly controversial conclusion, we ought to review the premises used to substantiate his conclusion.

Because Proposition 2 is more intuitive, Rowe begins his discussion of the problem of evil by analyzing God's reason for allowing evil to occur:

Let "s₁" refer to some instance of extreme human or animal suffering which an omnipotent, wholly good being, "OG," could prevent.

Either:

(i) there is some greater good, G, such that G is obtainable by OG only if OG permits s₁,

or

(ii) there is some greater good, G, such that G is obtainable by OG only if OG permits either s₁ or some evil equally bad or worse,

or

- (iii) s_1 is such that it is preventable by OG only if OG permits some evil equally bad or worse (Rowe 265).

In the first alternative (i), the greater good can only be obtained when God allows s_1 to occur. The second possibility (ii) is similar to the first in that a greater good can be achieved by s_1 . However, the goods in this case are also achievable when some event worse than s_1 occurs. And in the last instance (iii), no greater good is achieved, but s_1 can be prevented if God allows something potentially worse to occur. Only in conditions (i) and (ii) are instances of greater goods achieved by some evil. No greater good is achieved in the third.

This way of thinking can be confusing, so let us ground his thesis in a commonplace experience. Let G be the greater good which is the child's health, s_1 be the pain a child feels receiving a flu shot and let OG be the child's parent such that either,

- (i) there is some greater good, CHILD'S HEALTH (G), such that CHILD'S HEALTH (G) is obtainable by the PARENT (OG) only if the PARENT (OG) permits the CHILD'S PAIN (s_1),

or

- (ii) there is some greater good, CHILD'S HEALTH (G), such that CHILD'S HEALTH (G) is obtainable by the PARENT (OG) only if the PARENT (OG) permits either the CHILD'S PAIN (s_1) or some evil equally bad or worse,

or

- (iii) the CHILD'S PAIN (s_1) is such that it is preventable by the PARENT (OG) only if the PARENT (OG) permits some evil equally bad or worse.

In the first instance, the parent allows the child to experience pain to maintain good health. In the second instance, the child maintains good health when the parent decides to allow the child to experience the pain of getting the shot or the parent decides to

allow the child not to get the shot, but in doing so allows the child to contract the flu. And in the third instance, the child does not experience the pain of the shot, but experiences similar pain or contracts the flu.

The main point Rowe tries to convey is that there are instances where a greater good can be achieved only by necessary evils. By treating our relationship with God analogously to the relationship between a child and his or her parent, we come to understand that just as a good parent would not allow his or her child to experience unnecessary pain or suffering, a good God would not allow humans to experience unnecessary pain or suffering. The parent analogy appears to make the second premise acceptable to theists and atheists alike.

Because there is usually agreement on Rowe's second premise, we will turn our attention to the first premise, the major source of contention between theists and atheists. There are instances where evils do not occur to bring about greater goods or prevent worse evils from transpiring. These are what we would refer to as meaningless evils. We may think of Rowe's "Bambi" example, in which a fawn caught in a forest fire "is [meaninglessly] trapped, horribly burned, and lies in terrible agony for several days before death relieves its suffering" (266). Unlike moral evils, which are caused by free human agents, natural evils such as the fawn's cannot be solved by the free will defense, which is used to answer the logical problem of evil. Therefore, we must hold God responsible for these evils.

Skeptical Theism

Skeptical theism is the thesis that human beings are in no position to argue against the existence of God since any such argument depends on accepting unknowable premises. For instance, William Rowe contends that pointless or meaningless evils exist, but his claim is unprovable:

So according to the skeptical theist, we simply are in no position to reasonably judge that God could have prevented the fawn's five days of

terrible suffering without losing some outweighing good or having to permit some equally bad or worse evil. Our limited minds are simply unable to think of the goods that the mind of God would know (qtd. in Allen 121).

For the sake of simplicity, this position can be regarded as one of optimistic agnosticism.

I. CORNEA

Wykstra's "Condition of Reasonable Epistemic Access" (CORNEA) was the first of many cogent skeptical solutions offered by philosophers to Rowe's "The Problem of Evil and Some Varieties of Atheism" (1979). Recall that Rowe's argument is based on the premise that there are instances of "apparently pointless suffering," such as the fawn's narrative (Rowe 266). But despite feeling hard-pressed to accept that God does not exist, Rowe acknowledges that without omniscience, we are unable to say with certainty if the suffering we see is at all pointless. Wykstra's criticisms stem from this unverifiable epistemic claim.

In "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding Evils of 'Appearance'" (1984), Wykstra introduces CORNERA, a guideline to be used when making epistemic claims based on what *appears* to be the case. Wykstra believes that a human ("H") can claim to have knowledge of "p" through some cognized situation, "s:"

On the basis of cognized situation s, human H is entitled to claim 'It appears that p' only if it is reasonable for H to believe that, given her cognitive faculties and the use she has made of them, if p were not the case, s would likely be different than it is in some way discernible by her (Quoted by Adams 152).

Consider the following example: A six year old named Tom walks outside and notices the sidewalk in front of his house is wet, which prompts him to tell his parents, "It appears to have been raining." It is perfectly reasonable for Tom to believe that the

sidewalk is wet because it was raining earlier, but when CORNEA is applied to this situation it becomes clear that there are a number of subtle factors to take into account.

The person making the epistemic claim here is a six year old, who by the very nature of his age is limited in his cognitive abilities. Perhaps the sidewalk is wet because it was sprayed by a sprinkler. Conceivably, a neighbor with a hose could have drenched the sidewalk. It is also not unimaginable that an underground sewage pipe burst open near Tom's house, resulting in the wet sidewalk. There are many unstated possibilities as to why the sidewalk looked wet, but the main conclusion to be drawn from this example is that unless Tom can distinguish between a sidewalk that is wet from the rain and a sidewalk that is wet from sprinklers, he is not epistemically justified in claiming that, "It appears to have been raining."

Wykstra uses CORNEA to criticize Rowe's "Bambi" example. According to Rowe, the fawn's suffering seems to serve no outweighing good. Therefore, it is an instance of apparently pointless suffering. Equipped with CORNEA, Wykstra asks the burning question: "If there were an outweighing good of the sort at issue, connected in the requisite way to instances of suffering like this, how likely is it that this should be apparent to us" (qtd. in Adams 155)? It is from here that we can begin to analogize God as a parent as is oftentimes done by theists.

Wykstra conceives of God as a father to show how our intellect compares to God's intellect. He asks rhetorically: "How much greater [is God's intellect than ours]? A modest proposal might be that his wisdom to ours, roughly as an adult human's is to that of a one-month old infant's" (qtd. in Adams 156). If our understanding of the world, as Wykstra suggests, is that of a one-month old infant, we are in no position to judge God's actions when they seem to cause pointless evil. After all, there may in fact be an outweighing good which our limited minds prevent us from grasping. At the end of his paper, Wykstra concludes that,

If CORNEA is correct, such a concession is fatal to Rowe's case: for by CORNEA, one is entitled to claim 'this suffering does not appear (i.e., appears not) to serve any Divinely-purposed outweighing good' only if it is reasonable to believe that if such a Divinely-purposed good exists, it would be within our ken (Quoted by Adams 157).

By including the notion of limited cognition in his argument (the phrase, "does not appear"), Rowe makes his epistemic claim subject to CORNEA, providing theists a means of escaping from his criticism. The only way, according to Wykstra, Rowe can rescue his position from CORNEA is if he shows that it is reasonable for a theist to believe that Divinely purposed goods are within our understanding.

Conclusion

Shortly before Rowe published his first paper on the evidential argument, philosopher Thomas Nagel published a seminal article, "What is it like to be a bat?" Two important principles discussed in this paper can be applied to the problem of evil and may be used in a "Rowe-like" objection to the parent analogy.

I. What it is Like to be a Bat

Thomas Nagel discusses the problem of consciousness through a thought experiment in which a human brain is placed in a bat's body. According to Nagel, even if a human brain is transplanted into the body of a bat, the brain will not be able to think like a bat. Regardless of what body houses what brain, both humans and bats will have the perception of their respective species because humans cannot have the subjective experiences of bats, such as knowing what it is like to use echolocation since birth, and bats cannot have a human's subjective experiences, like learning to walk. No matter what body a human brain is held in, a human mind can at best only be a human mind thinking about what it is like to be the animal whose body it is in.

How does Nagel's bat case pertain to the evidential problem of evil? Wykstra's CORNEA principle is valid only if the universe is morally obscure, which means humans cannot identify the greater goods that justify evil. But if the world is morally transparent, meaning these greater goods are understood by humans, CORNEA fails. Although it may seem reasonable to believe that a wholly good God would create a transparent universe, Wykstra asserts that the parent analogy supports the idea of God creating an obscure universe on the basis that God's intellect and creation would always go beyond our cognitive reach. The logical conclusion of Wykstra's argument can be obtained as follows:

Take C to mean "CORNEA is valid," E to mean "the evidential argument fails," O to mean "the universe is morally obscure," and P to mean "the parent analogy is valid."

Argument #1:

1. $C \supset E$
2. $C \equiv O$
3. $P \supset O$
4. P
5. $\therefore E$

Proof for Argument #1:

1. $P \supset O$
2. P
3. O 1,2 Modus Ponens
4. $C \equiv O$
 - 4.1 $(C \supset O) \bullet (O \supset C)$ 4, Equivalence
 - 4.2 $(O \supset C) \bullet (C \supset O)$ 4.1, Commutativity
 - 4.3 $O \supset C$ 4.2, Simplification
5. C 3,4.3 Modus Ponens
6. $C \supset E$
7. E 5,6 Modus Ponens

Logical Name and Symbol	Definition
Horseshoe \supset	If ____, then ____.
Triple Bar \equiv	If and only if, propositions equivalent
Dot \cdot	And
Tilde \sim	Not (negation of proposition)
Therefore Sign \therefore	Therefore (indicates end of proof)

Thus, if the parent analogy is valid, the universe is morally obscure and CORNEA is valid, invalidating the evidential argument.ⁱⁱ

It would appear as though Wykstra's argument against Rowe's position is irrefutable, but when his argument is posed in conjunction with Nagel's analogy principle, the result proves fatal. If "A" is symbolic for "all analogies are exhaustive" and, once again, "P" stands for "the parent analogy is valid," we are left with the following proof:

Argument #2:

1. $\sim A$
2. $\sim A \supset \sim P$
3. $\therefore \sim P$

Proof for Argument #2:

1. $\sim A \supset \sim P$
2. $\sim A$
3. $\sim P$ 1,2 Modus Ponens

Upholding Nagel's principle that all analogies, especially the bat case, are "incompletable," it can be concluded that the parent analogy is incomplete as well. What consequences does this conclusion have for Wykstra's argument? Picture Wykstra's argument as the Eiffel Tower. At the top of the tower is CORNEA, the level below that obscurity, and the parent analogy at the bottom level. When we apply Nagel's analogy principle to Wykstra's position, we are showing that the building's foundation is built on sand, leading to the failure of skeptical theism. Nagel's point, however, is compatible with Rowe's evidential argument.

Whereas Wykstra uses the parent analogy to support the idea of a morally obscure universe, Rowe only uses the parent analogy to show that its use is arbitrary for Wykstra. Rowe believes that if we treat God as a parent, it is reasonable to suppose that God would make the world morally transparent. Because the analogy can be used to support moral obscurity and moral transparency—that God can act as a good parent by creating a world where the goods that justify evil is unknown to us and by creating a world where we can know the goods that justify our suffering—it would seem that the analogy is a weak foundation for skeptical theism. And because Rowe’s use of the parent analogy has no bearing on the validity of his argument, it does not follow that the parent analogy being invalid undermines the evidential problem.ⁱⁱⁱ

II. Wykstra Saved

An analogy can be understood as a means of comparison between two similar things. In his *Metaphysics*, Aristotle categorizes words as equivocal when “they have a common name, [but] the definition corresponding with the name differs for each,” and univocal when they “have both the name and the definition answering to the name in common” (Aristotle7). Consider this pair of statements, labeled “A” and “B:”

Example A:

Jack and Harrison are humans.

All humans are rational.

Example B:

Max swung the bat.

Dracula is a bat.

The word “humans” have the same meaning in both sentences in Example A, so they may be regarded as univocal. Although the word “bat” appears in both sentences in Example B, the meaning of the word is not used in the same fashion (i.e, “bat” in the first statement refers to a piece of athletic equipment whereas “bat” in the second describes a nocturnal animal). Hence, “bat” is equivocal in meaning.

Not all words, however, are univocal or equivocal in meaning. With this in mind, Aristotle created a third category of meaning—

pros hen. Words denoted as pros hen (*pros* meaning “related to” and *hen* meaning “one”) are used equivocally with a common meaning. Healthy is one such word, with many possible uses:

Your complexion is healthy

Broccoli is healthy.

Exercising is healthy.

A patient is healthy.

Each use of the word is different, but as Aristotle smartly points out, “Everything which is healthy is related to health” (Aristotle “Metaphysics” 732). Aristotle later applies this concept to his discussion of “being as being,” which he uses to explain substance or God. Saint Thomas Aquinas eventually builds on Aristotle’s idea of pros hen equivocally with a rich religious-linguistic tradition, making him one of the most notable Christian philosophers (Baird 307).

In his *Summa Theologica*, Aquinas argues that “no name is predicated univocally of God and of creatures,” but at the same time, names are not “applied to God and creatures in a purely equivocal sense” (Quoted by Baird 329). For instance, I can use the word “good” to describe my friend Adam or to describe God, but the meaning of the word is not exactly the same in both cases. Adam is good because he exhibits the best qualities that he can possess and God is good because God exhibits the best qualities that God can possess. Undeniably, the goodness that God exhibits is far greater than the goodness that Adam exhibits, so the word differs in that respect. Initially, this may seem problematic. After all, how can Aquinas say that names applied to God and creatures not be purely equivocal or purely univocal? Only by analogies is this possible.

Aquinas refers to analogies as “a mean between pure equivocation and simple univocation” (qtd. in Baird 331). He goes on to say that, “in analogies the idea is not, as in univocal, one and the same, yet it is not totally diverse as in equivocals; but a term which is thus used in a multiple sense signifies various proportions to some one thing” (qtd. in Baird 331). An analogy is when the same words have different meanings, but those

meanings all relate to one central concept. Recall that Aristotle refers to words that are used equivocally with one common meaning as being *pros hen* equivocal and that his example of a word being used *pros hen* equivocally is “healthy.” In his discussion of analogies, Aquinas explains how Aristotle’s healthy example can be used to draw the connection between *pros hen* equivocality and analogies; *pros hen* equivocality to Aristotle is analogy to Aquinas.

From Aquinas’ explanation of analogies, we can conclude that analogies are not “incompletable” as suggested by Nagel, but rather that using analogies are a means by which humans can bridge the gap between ourselves and God. Ultimately, in the context of the evidential argument, this conclusion can be used to validate Wykstra’s parent analogy. Because the argument proposed by Aquinas allows us to use analogies, we are justified in projecting our own views on what qualities make a parent “good,” onto God. By this admission, the parent analogy can be regarded as complete^{iv} and the idea of a morally obscure universe is adequately supported. Because we have grounds for believing that the universe is morally obscure, we can consider CORNEA when making epistemic judgments. Finally, because CORNEA is valid, the evidential argument proposed by Rowe fails.

III. Final Thoughts

Cognitive limitations and living in a morally obscure world prevent us from knowing all the goods or evils that exist. Wykstra uses the parent analogy to support a morally obscure universe. By doing this, Wykstra makes himself vulnerable, since his argument could be defeated by invalidating the parent analogy which he uses as its foundation.

Thomas Nagel’s paper “What is it like to be a bat?” focuses on subjectivity and provides evidence against the completeness of analogies. After considering Nagel’s bat case it would seem as though the parent analogy fails to fully describe God as a parent since no analogy is exhaustive. Because the parent analogy fails by Nagel’s standards, it would appear that Wykstra is not justified

in claiming that the world is morally obscure and if the world is world is not morally obscure then CORNEA is not valid.

Wykstra's position is saved by the linguistic-religious work of Aristotle and Saint Thomas Aquinas. Aquinas provides support for the use of analogies when describing God, which allows us to conclude that the parent analogy is valid. Because the parent analogy is valid, Wykstra is justified in claiming that the world is morally obscure and because the world is morally obscure, CORNEA is valid.

Notes

ⁱ For the purposes of this paper, I adopt the classical theist definition of God. According to classical theism, there is one God who is omnipotent (all powerful), omniscient (all knowing), and omnibenevolent (wholly good). Christians, Jews, and Muslims uphold the classical notion of God. I chose to define God in accordance with classical theism to remain consistent with the definition of God used in philosopher William Rowe's original argument on the problem of evil.

ⁱⁱ It is worthwhile noting that the second proposition shows that CORNEA only succeeds if the world is morally obscure.

ⁱⁱⁱ The concept of subjectivity discussed in Nagel's bat argument is traditionally used to support theism. Theists have used the bat case to argue that just as the subjective mind of a bat is inaccessible to humans, the mind of God is inscrutable to all living creatures. Although there is merit to this interpretation of Nagel's argument, I demonstrate a different understanding of this analogy. I believe that Nagel uses the bat example to show that as convenient as analogies are, they are by no means exhaustive.

^{iv} Complete is used here to mean that this analogy successfully portrays the relationship between God and humans. To put it in context of the argument over the incompleteness of analogies, analogies discussing the attributes of God are exhaustive. It is important to note that characterizing this analogy as valid does not mean we have full knowledge of God's actions or attributes.

References

- Aquinas, Saint Thomas. "Summa Theologica." *Medieval and Renaissance Philosophy*. Edited by Forrest E. Baird. Upper Saddle River: Prentice Hall, 2011.
- Aristotle. "Metaphysics." *The Basic Works of Aristotle*. Translated by E.M. Edghill. Edited by Richard McKeon. New York: Random Press, 1941.
- Aristotle. "Organon." *The Basic Works of Aristotle*. Translated by E.M. Edghill. Edited by Richard McKeon. New York: Random Press, 1941.
- Baird, Forrest E. "Thomas Aquinas." *Medieval and Renaissance Philosophy*. Edited by Forrest E. Baird. Upper Saddle River: Prentice Hall, 2011.
- Hickson, Michael W. "A Brief History of Problems of Evil." *The Blackwell Companion to The Problem of Evil*. Oxford: Wiley-Blackwell, 2014.
- Nagel, Thomas. "What is it like to be a bat?" *The Philosophical Review* (1974) 83: 435-450.
- Peterson, Hasker, Reichenbach, and Basinger. "The Problem of Evil: The Case Against God's Existence." *Reason & Religious Belief: An Introduction to the Philosophy of Religion*. New York: Oxford University Press, 2003.
- Reichenbach, Bruce. "Why is God Good?" *Evil and a Good God*. New York: Oxford University Press, 1982.

Rowe, William L. "The Problem of Evil." *Philosophy of Religion: An Introduction*. Edited by Holly J. Allen. Belmont: Cengage Learning, 2007.

Rowe, William L. "The Inductive Argument from Evil against the Existence of God." *Philosophy of Religion: An Anthology*. Edited by Michael Rea and Louis Pojman. Stamford: Cengage Learning, 2015.

Wykstra, Stephen J. "The Humean Obstacle to Evidential Arguments from Suffering: On Avoiding the Evils of 'Appearance.'" *Oxford Readings in Philosophy: The Problem of Evil*. Edited by Marilyn McCord Adams and Robert Merrihew Adams. New York: Oxford University Press, 1990.

**Cross-Media Metrics: The Shift in Audience
Measurement
Zarina Akbary (CLA 2019)**

Abstract

Within the past few years, the media landscape has seen remarkable changes in the technologies made available to consumers and the ways in which consumers access content. As a result, audience behavior has changed drastically, giving rise to three critical issues: cannibalization, convergence, and fragmentation. These three issues present a challenge to audience measure in addition to the issues of dealing with Big Data, and audience measurement companies must find a way to address them, either by adjusting current metrics or developing new ones. Otherwise, the audience marketplace risks compromising accuracy. Although these obstacles have yet to be overcome, the audience marketplace is undergoing a huge evolution, one where competing metrics have arisen and success is being redefined. Most notable in this shift in audience measurement is the re-evaluation of exposure-based metrics and the increasing value of audience engagement. However, amidst this adaptation and innovation, the consumer perspective—the powerlessness that consumers feel in regard to their current media consumption and what they want from the audience marketplace in the future—is not being discussed and must be taken into greater account.

Introduction

The introduction of new media platforms, such as tablets and smartphones, has revolutionized the way consumers access content. Naturally, these momentous changes in media consumption garner a great deal of attention. From the humanities, to cultural studies, to the sciences, conversations about media consumption and its effects persist (Hassoun 272). However, throughout those discussions, few engage in discourse on cross-media metrics, which are crucial to evaluating the current media landscape and consumer usage. The exceptions are those industries in which cross-media metrics play a central role: advertising, TV networks, and audience measurement companies (Napoli). In this paper, these industries are collectively referred to as the audience marketplace (Napoli).

While cross-media metrics and its issues encompass a myriad of complexities, they are essentially the measurement systems used to measure audience viewership across multiple media platforms (Enoch and Johnson). Audience measurement companies directly determine these metrics and utilize them to collect data on audiences. This information is sold to advertising agencies, networks, and numerous other companies to guide business decisions and influence the content produced for consumers (Napoli). Since the information must be as current and relevant as possible, new data is constantly collected and analyzed by these audience measurement companies. New programs, new media platforms, and new trends in audience viewership create a challenging dynamic for audience measurement companies to keep pace with. In this rapidly changing media landscape, the three critical issues—cannibalization, convergence, and fragmentation—in cross-media metrics arise, and various methods have

been employed to address those issues. The more pressing questions for the audience marketplace is whether the goals of existing metrics have changed and what repercussions those different goals could have, especially with regard to defining success.

This paper begins with an overview of issues in cross-media metrics: starting with accuracy and its significance in the audience marketplace, and continuing on with a discussion about the challenges of Big Data. Next, I dissect the three critical issues of cannibalization, convergence, and fragmentation. This leads into a discourse on audience exposure and audience engagement, the competing measurement systems based on each, and their value in the audience marketplace. Lastly, I explore the consumer perspective and the juxtapositions between the consumer and the audience marketplace.

Accuracy

The goal of cross-media metrics is to accurately measure audience viewership, and any standard set forth by the audience marketplace aims to better achieve that accuracy (Napoli). The audience marketplace deals with three types of audiences: the predicted audience, the measured audience, and the actual audience (Napoli). Advertisers have an audience of a certain size and demographic that they intend to market their product and or service to, and media companies approach these advertisers claiming their media product (shows, magazines, etc.) will reach that audience. Together, media companies and advertisers predict the size and demographic of a media product's audience. That predicted audience is an educated guess on the part of both parties, and the success of that media product is determined by how close that predicted audience is to the

measured audience. Audience measurement companies, such as Nielsen, provide data on the measured audience, which the audience marketplace views as an accurate representation of the actual audience. Media companies and advertisers cannot control the media consumption of the actual audience, nor can audience measurement companies measure the media consumption of every individual in the country (legally or practically). Since the actual audience is unknowable, the measured audience is regarded as the accurate representation, a useful equivalent of the actual audience.

Since the success of media products and advertising campaigns are based on the difference between the predicted audience and the measured audience, and the reliability of the measured audience is based on how closely it mirrors the actual audience; the accuracy of metrics is crucial (Napoli). In one sense, accuracy is a matter of perception. If members of the audience marketplace agree that the measured audience is accurate enough, then the measured audience is recognized as a reasonable proxy for the actual audience. However, if the members of the audience marketplace do not agree, competing metrics arise. Often, the perception of accuracy is affected by how favorable metrics are to certain audience marketplace participants. If new metrics favor digital content but not cable television, cable television executives are more inclined to view the new metrics as less accurate.

Opinions on the accuracy of metrics are influenced by what the audience marketplace believes are trends in the actual audience. In another sense, accuracy is better achieved by minimizing sources of error (Napoli). Prior to the people meter (a device attached to TV cable boxes to track what programs are being watched), audience

measurement companies obtained data on TV viewership by sending out surveys and TV diaries for households to record their TV viewing (Steel). Individuals recording their own viewing habits are much less reliable than the people meter, which notes what shows were viewed and for how long passively without interfering with the individual's viewing. Besides measurement techniques, other sources of error include non-representative samples and the challenges that arise when dealing with Big Data.

Big Data in Audience Measurement

Ultimately, the fact that cross-media metrics are Big Data poses the greatest challenge for audience measurement companies. Specifically, the characteristics of Big Data, encompassed in the HACE theorem, is the reason why the data collected by audience measurement companies are difficult to obtain and analyze. HACE stands for heterogeneous, autonomous sources, and complex and evolving relationships (Wu et al. 98). In the context of audience measurement, heterogeneous refers to the various sample features by which an audience may be represented. Some examples of sample features include age, race, gender, and income. Audience measurement companies must not only decide which sample features to use, but also what range they encompass. For example, it is not practical to group age in one-year increments. Individuals in a panel are anonymous, and to maintain that anonymity, uncertainty must be present in the data. Knowing the range of an individual's age infringes less on their privacy than knowing their exact age. Additionally, categorizing the individuals in a sample by one-year increments is also difficult because an individual will only remain in that group for a year and individuals become a year older at different points in the year (Wu et

al. 101). That being said, an age range that is too big compromises accuracy.

Although the term “millennial” refers to people age 18-34, the lifestyle of young dependent millennials, and by extension their media consumption, is vastly different from the lifestyle of those who are older and starting a family (“How Much TV Do Millennials Watch?”). Nielsen took this into account and released a section of their fourth quarter “Total Audience Report” that distinguished subgroups within the millennial generation, revealing the various trends within those subgroups (“How Much TV Do Millennials Watch?”). When millennials age 18-34 are grouped together, the average person in that age group spends four hours and eight minutes a day using the TV. However, dependent millennials use the TV for three hours and forty-four minutes a day on average, whereas millennials starting a family use the TV for four hours and forty minutes—a full hour more (“How much TV Do Millennials Watch?”). Moreover, societal changes alter the validity of certain sample features. Race, for example, becomes increasingly arbitrary in a multicultural nation. Interracial households will fall under multiple categories of race while certain groups, such as those of Middle Eastern descent, are not given a category at all. Despite these flaws in sample features, their use is commonplace in audience measurement systems and useful for the audience marketplace, which relies heavily on demographics (Kosterick and Napoli 4).

The second characteristic of Big Data, as indicated by the HACE theorem, are autonomous sources, which refers to the fact that data gathered from individuals in a sample are obtained independently (Wu et al. 98). In a panel with of thousands of households, each one must have its own

measurement device installed to collect data from (Hassoun 276). The result is a large volume of data containing complex and evolving relationships (Wu et al. 98). That complexity is the most distinguishing feature of Big Data.

In the case of audience measurement, complexity arises from assumptions, biases, and the three critical issues of cannibalization, convergence, and fragmentation, which are discussed in the next section. Audience measurement companies are aware that any given TV program or video may not have the audience's undivided attention, but there is no way to take into account certain distractions in the data, such as a person who sleeps with the TV on (Hassoun 273). Hence, present within the data is the assumption that the audience member watched the program. The presence of bias within a data set is also inevitable. Though audience measurement companies put their best effort to selecting panels as randomly as possible, samples are not perfect. When measuring viewership on devices like tablets, the audience is individuals of higher income who can afford a tablet. Likewise, certain media platforms, such as traditional TV, are losing favor among younger demographics (The Nielsen Company).

Audience measurement companies have had a history of bias, with measurements that better serve white males than females and other minority groups (Kosterick and Napoli 14). Again, assumptions and biases are inevitable, but audience measurement companies try their best to mitigate them with more representative samples and better measurement devices and techniques. Furthermore, the strategies for dealing with Big Data are also improving and, although Big Data presents itself as a challenge, it is an opportunity to learn more than ever

before about the audience. Similarly, more recent challenges, such as the critical issue of fragmentation, are difficult to deal with but present new opportunities for the audience marketplace.

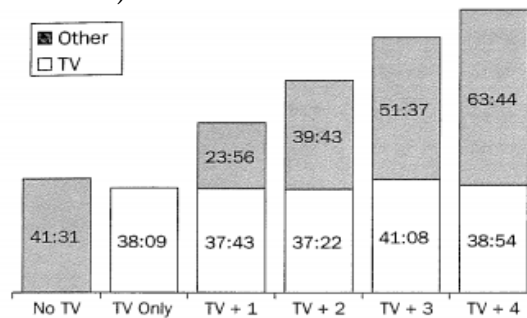
The Three Critical Issues

I. Cannibalization

Anytime new media platforms enter the market, people predict that it will “consume” all the time dedicated to older media. This phenomenon is known as cannibalization, and despite persistent speculations about it, the extent of cannibalization is limited, especially as far as digital media are concerned (Enoch and Johnson 127). For example, when the DVR entered the market, people predicted the death of live TV-viewing. But not every household adopted DVR; only about one-third of TV-viewing households adopted the DVR (Enoch and Johnson 126). Some households do not use a DVR every day or for all their TV-viewing. In 2009, DVR composed just six percent of all TV viewing during the fourth quarter. The same principle applies to other technologies like tablets and smartwatches, which have not replaced smartphones or laptops. When new media platforms are introduced, they do not completely replace or cannibalize current media platforms. If that were the case, audience measurement would be, in some ways, much easier (Enoch and Johnson 126). Audience measurement companies would lose money on established panels, but once they get rid of one they would not need to continue utilizing the older ones like they do now.

Aside from the misconception that new technology replaces older ones, cannibalization encompasses the idea that the use of new media “eats into” the usage of older media (Enoch and Johnson). For example, if the use

of smartphones increased, we may expect that TV usage would decrease. However, this logic is flawed and not supported by existing data. The time that individuals spend using media continues to increase, especially as new technologies open “new markets of time” (Enoch and Johnson 134). “New markets of time” refers to the fact that advertisers and content producers can reach users for a larger fraction of their day. For example, the portability of smartphones means that consumers can access content in settings they previously could not, like in a classroom or on a bus. That new market of time on the bus does not come at the expense of the time a consumer spends watching TV at home. In fact, one trend evident by the data provided by audience measurement companies is that heavy users of one medium tend to be at least above-average users of another medium (Enoch and Johnson 130).



ESPN's analysis first cut of Knowledge Network's Multimedia Mentor data made TV the primary medium. The network created cells of users based on those who viewed TV only, TV plus one other medium, two other media, and so forth up to four additional media. Additional media use had no effect on the amount of TV viewing.

Source: Knowledge Networks Multimedia Mentor, Spring 2009

Figure 1. Time TV viewed + Time Additional Media Were Viewed. Source: Enoch and Johnson.

In households that view TV, usage of television remains consistent, even when additional media are also used

(Enoch and Johnson 130). As shown in Figure 1, households that use TV and four other mediums have almost the same TV usage as households that only watch TV. There is no cap on the amount of time a given individual will spend consuming media, and they are free to increase the total time spent using media so that the usage of one medium will not come at the expense of another. Given the data, cannibalization seems to be no more than a myth perpetuated by the fear of obsolescence. But it must be acknowledged that, while new markets of time have opened, multiple media platforms still compete for the audience's attention within the same time frame, especially if the audience is using two mediums at once.

II. Convergence

Simultaneous media usage, or the use of two media platforms at the same time, is referred to by the audience marketplace as convergence (Hassoun 271). In previous years, the audience marketplace ignored the phenomenon, but in recent years convergence has become so prevalent that it is impossible to ignore (Hassoun 273). Some studies show that more than half of a consumer's total media time is spent using more than one media platform (Hassoun 276). Given its prevalence, convergence both excites and terrifies the audience marketplace since it splits the audience's attention but offers the potential for "the second screen" (Hassoun 279). As previously mentioned, audience measurement companies are aware that media platforms do not always possess the audience's undivided attention (Hassoun 275). In the past, it was assumed that the audience's attention was split between the medium and manual labor (e.g., domestic chores), but ultimately the consumer was receiving content from only one media platform. However, the phenomenon of convergence is

distinct from previously assumed distractions because instead of the attention being split between the medium and a chore, it is split between two mediums. When the audience's attention is split between two screens, the message that advertisers try to convey becomes less effective and the quantity of attention that advertisement receives from the audience becomes dramatically reduced (Hassoun 274). Because convergence has such an effect on the efficacy of advertisements, the audience marketplace demands metrics to measure it and reduce uncertainty.

This is a challenge for audience measurement companies because convergence must be measured differently than the way most media platforms are usually measured. Companies such as Nielsen use data fusion sources, panels that only measure the audience for one media platform (Enoch and Johnson). The panel that Nielsen uses to measure radio usage (the Personal People Meter) only measures a radio-listening audience, whereas a TV-viewing audience is measured by another panel (the National People Meter) (The Nielsen Company). To measure convergence, Nielsen uses a convergence panel that utilizes single source data collection (Enoch and Johnson 126). Unlike data fusion, which measures the audience for each media platform separately, single sources are a group of case studies in media consumption.

All of the consumers' media consumption (TV usage, radio usage, internet usage, etc.) in the panel is measured. Single sources are used mainly to measure convergence and are not as widely used or relied upon as data fusion (Enoch and Johnson 126). It is harder to obtain a representative sample for a single source and even more difficult to find a representative sample that

engages in the use of all media platforms. For instance, producers of radio programs are first and foremost interested in the radio-listening audience. If a number of people in the panel do not listen to radio, they are not a good sample for representing the radio-listening audience. The same applies for all other media platforms (Enoch and Johnson 126). Additionally, audience measurement companies have huge investments in their data fusion panels and so it is not feasible to convert or replace them with convergence panels.

The demand for convergence panels may increase in the near future as the audience marketplace figures out how to take advantage of simultaneous media usage (Hassoun 272). As previously mentioned, convergence concerns advertisers because they fear that, when the audience's attention is split between two screens, the advertisement becomes less effective. On the other hand, if advertisements engage audiences on both screens, the message's effectiveness increases. This idea provides the basis for "the second screen," a strategy for the audience marketplace to take advantage of simultaneous media usage to coordinate content across multiple media platforms (Kattukaran). In this way, convergence may work in their favor.

Networks have already capitalized on the potential for coordinated media content through the creation of companion apps, which allows mobile phones or tablets to act as the second screen to television, the primary screen (Kosterick and Napoli 10). For instance, Fox released a companion app for its program, *Sons of Anarchy*, which allowed viewers to learn about and purchase any of the products featured on the show (Hassoun 279). In 2011, the Academy Awards launched the "Oscar Backstage Pass," an app that streamed live

backstage footage during commercial breaks. Other second screen strategies involve the coordination of social media platforms such as Facebook and Twitter where, in the case of the Academy Awards, viewers were encouraged to post on the site as winners were announced (Kattukaran). In short, a great deal of experimentation is being done with the second screen and, while convergence puts a strain on traditional metrics, it offers a host of new possibilities for the audience marketplace.

III. Fragmentation

Over the past few years, the introduction of new media platforms to consumers changed the way audiences access content. In addition to these new platforms, the abundance of content itself continues to grow (Kosterick and Napoli 7). Consequently, audiences are divided widely across these various platforms and programs, resulting in a phenomenon known as fragmentation. Even when observing one platform (e.g., television), there are hundreds of channels that broadcast a wide array of programs (Kosterick and Napoli 7). For instance, in 2015, FX C.E.O. John Landgraf reported that the network produced 371 scripted shows that year alone and projected the production of more than 400 in 2016 (Nussbaum). With audiences so widely distributed, viewing audiences become smaller, and it becomes more and more challenging to detect a given audience for a given show (Kosterick and Napoli 7). In response to fragmentation, the audience marketplace demanded new and improved metrics to measure the audience more accurately. The issue of fragmentation is so pressing that it drives the shift in audience measurement and the rise of social media analytics.

Audience Engagement vs. Audience Exposure

The audience marketplace has come to a point where it must decide the value of audience exposure and audience engagement and, consequently, which matters more when defining success. Traditional metrics measure audience exposure, or the amount of people that viewed a program (Kosterick and Napoli 2). As previously mentioned, new media platforms have been introduced in recent years, but new metrics have not been developed to measure the audiences for these new platforms. Admittedly, these new metrics were given different names, but regardless these metrics addressed questions solely about audience exposure (Napoli). For example, even though the number of viewers for a television show is referred to as a Reach and the number of viewers for an online video is referred to as a Unique, both measure the number of viewers (Enoch and Johnson 126, The Nielsen Company). By contrast, social media analytics, or social TV analytics, sets out to measure audience engagement (how invested the audience is in that program). Therefore, the rise of social media analytics requires new metrics because the question of audience engagement is fundamentally different from that of audience exposure.

To quantify audience engagement, audience measurement companies have turned to social media as a possible source for new metrics (Kosterick and Napoli 2). In 2012, Nielsen partnered with Twitter to develop the "Nielsen Twitter TV Rating" (NTTR) to measure audience engagement according to actions on Twitter (Kosterick and Napoli 12). Likewise, other social media companies, such as Facebook, support the push for social media analytics and offer potential measurement systems based on their respective social media platform. For example, Facebook offered weekly reports to various networks about how many "actions" (e.g., likes or posts) were

motivated by the networks' shows (Kosterick and Napoli 9). In addition to social media companies and audience measurement companies, television networks have also embraced the idea of social media analytics, mainly because it provides an alternative form of success (Kosterick and Napoli 9).

Even if audience engagement and exposure are measured for the same program, a different but accurate conclusion may be drawn about the success of that program if different metrics are utilized (Kosterick and Napoli 2, Wu et al. 98). This concept also holds true for various other industries. For instance, an X-ray and a blood test can be conducted on an individual with a broken bone. From the X-ray, a physician easily identifies that the person has an injury. From the blood test, a physician knows nothing about the injury but immediately recognizes that the individual is not anemic (Wu et al. 98). In the same sense, traditional exposure-based metrics can reveal the size of a program's audience if it is big enough to detect, but it cannot detect how invested that audience is or the impact that the show has on the lives of the audience members (Kosterick and Napoli 10). Now, with two fundamentally different ways to measure success, networks with low traditional ratings still have an audience to present to advertisers and a choice of which measurement system they value more (Napoli). However, despite the rise of social media analytics, traditional exposure-based metrics still hold the greatest value in the audience marketplace (Kosterick and Napoli 12).

The Problems with Social Media Analytics

Part of the reason why exposure-based metrics persist is precedent (Napoli). Marketers have advertised on TV for decades and know how to do it well, whereas advertising on the Internet and on new media platforms such as smartphones is still in its infancy (Kattukaran). Naturally, advertisers are comfortable applying exposure-based metrics, which they know well, to territory as unfamiliar as the Internet. Nielsen, the audience measurement company with a monopoly over audience measurement, is greatly invested in exposure-based metrics and, if audience engagement suddenly became more valuable, Nielsen risks losing its dominant status in the audience marketplace (Kosterick and Napoli 2). To be sure, Nielsen's investment in exposure-based metrics and the audience marketplace's history of reliance on them partially explains why social media analytics have not been established as the norm in audience measurement, but these are not the only reasons.

With social media analytics comes a host of new questions and issues. First and foremost, not all social media is created equal. 72% of all American internet users are on Facebook whereas 23% are on Twitter (Duggan). Of all American female internet users, 21% are on Twitter while 44% are on Pinterest (Duggan). If the audience marketplace chooses to rely more on social media analytics, they must decide whether a "like" on Facebook is equivalent to a "favorite" on Twitter. Additionally, social media analytics may provide the audience marketplace with a better idea of how engaged audience members are with a particular show, but they cannot report how long the show was viewed or on what platform (Enoch and Johnson 126). Currently, most of the audience marketplace's anxiety stems from an increase

in uncertainty due to the three critical issues previously discussed, and social media analytics certainly provide their own uncertainties. But instead of rejecting social media analytics outright, the audience marketplace values social media analytics, and by extension audience engagement, as a supplementary metric to traditional exposure-based metrics.

Discussion

To remove uncertainty between the measured audience and the actual audience, the audience marketplace continues to demand new metrics and more from existing metrics. Some proposals for new metrics are intrusive, such as biometric devices that monitor viewers' physical responses to evaluate emotional reactions (Hassoun 278). Other techniques involve installing cameras on televisions to monitor viewers and verify whether or not their eyes are on the screen. Practical and legal constraints prevent the implementation of these new metrics (Hassoun). Similarly, practical and legal constraints limit the granularity of existing metrics. But these demands are the product of an audience marketplace in flux.

In the past, ratings defined success ("What Counts as a Hit?"). The overnight viewer numbers determined whether a show would live or die while the quality of a program was determined by its ratings (Nussbaum). As George W.S. Trow, a notable media critic in the 1980s put it, "What is loved is a hit. What is a hit is loved" (qtd in Nussbaum). Since success meant acquiring the greatest number of viewers, the idea of creating niche content was absurd and not economically viable (Kosterick and Napoli 10). However, now that the assessment of metrics have changed, so too has the definition the success.

Ratings are still a widely used metric, but it is not the only measurement and not as definitive to success as it had been (“What Count as a Hit?”). Low ratings do not always result in cancellation; audience engagement on social media sites or on companion apps are now valued measurements (Kosterick and Johnson). The perception of quality and niche audiences has also undergone a shift. A show that lacks high ratings can still be considered high quality (“What Counts as a Hit?”). Furthermore, the production of niche content is taken seriously and minority audiences have more value in the audience marketplace than before (Kosterick and Napoli 11). A large part of this shift can be attributed to fragmentation, which forces the audience marketplace to find new ways to detect, evaluate, and reach their audience, taking note of audiences and audience behavior that was previously unnoticed. In addition, the critical issue of cannibalization, though not as dire as some believe, alarms participants in the audience marketplace, pushing them to evolve technologically and artistically. Likewise, the issue of convergence also motivated a great deal of experimentation.

The audience marketplace continues to seek new ways to better utilize the second screen. It is difficult to assess the current state of audience measurement, let alone predict its future, but clearly there is a huge shift in audience measurement. These challenges demand different evaluations of data. In a way, this has liberated the audience marketplace from self-imposed constraints, and with new uncertainties arise new opportunities for innovation.

Future Suggestions: Integrating the Consumer Perspective

So far, this paper has focused on the audience marketplace and changes in media consumption that have spurred a shift in audience measurement. Missing in this discussion, and in most discussions on this topic, however, is the consumer's perspective. The audience marketplace has its own ideas of the future and a relationship with data that juxtapose consumers' attitudes. One point often overlooked by the audience marketplace in their speculations of a future with personalized advertisements, ads that utilize specific data about viewers to market to them, is the present reality that ad blockers are chart-topping downloads (Kattukaran, "Ad Blocker's Dilemma"). There are over 198 million ad-block users, costing publishers 22 billion in wasted ad dollars ("Ad Blocker's Dilemma"). The audience marketplace dismisses the consumers' resentment and annoyance with advertisements, claiming that, as data quality improves, advertisements will be more targeted and more relevant to consumers, giving them a better experience with advertisements (Kattukaran). However, gathering more personalized data raises concerns about privacy and does not alter the consumers' annoyance with ads they find interruptive, regardless of the ads' relevance ("Ad Blocker's Dilemma").

Another juxtaposition between the audience marketplace and consumers is their differing perspectives on the power of the consumer over media. To the audience marketplace, viewers are uncontrollable, free to dedicate however much time they wish to any content on any media platform (Napoli). The very issues driving the shift in audience measurement arose because the consumer has the freedom of choice. Knowing this, it

is astounding that the consumer feels a lack of control over their media consumption. Consumers report feeling overwhelmed and even powerless (“The Case for Infomagical”) Many of the conceptions that people have about media consumption are not supported by the data. Nowhere is this more evident than the consumers’ misconception that they only have two options: use technology or do not (“The Case for Infomagical”).

One means of empowering the consumer is to change their relationship with data. In the past, data and metrics trapped the audience marketplace by narrowing the definition of success and limiting possibilities for innovation. In a like manner, the consumer seems to trap themselves with metrics. Consumers use media platforms to collect data on how many steps they have taken, how many calories they have eaten, how many followers they have, how many views their videos receive, etc (“The Case for Infomagical”). They are, at once, oblivious to and acutely aware of their dependence on devices and their reliance on the data it provides. However, just as the audience marketplace has been liberated of self-imposed constraints without abolishing metrics, so too can the consumer.

A good start would be to introduce consumers to the data gathered by audience measurement companies, beyond using it to alarm people about how much time they spend on their phones. A consumer can look at the data as a point of comparison to evaluate their media consumption against an accurate average instead of self-evaluating their current media consumption according to what they perceived it to be two years ago. Then they can decide whether they want to increase or decrease their media consumption, or allocate their time to a different media platform. But the most important information for

a consumer to gather from audience measurement data is the knowledge that they have choice over their media consumption and that, as far as metrics are concerned, there is more than one value of worth.

As previously mentioned, cross-media metrics exist to accurately measure audience viewership. In light of the changes in the media landscape, the audience marketplace has experimented, innovated, and evolved to realize that goal. In the same token, the consumer is also adapting to these changes. As this evolution is underway, it would benefit the consumer to see the reality of media consumption, as supported by the data, and engage in discourse about what they want from media. Likewise, moving forward, the consumer perspective needs to be given greater weight in the audience marketplace.

References

- Duggan, Maeve. The Demographics of Social Media Users: Pew Research Center, 2015. Web.
- Enoch, Glenn, and Kelly Johnson. "Cracking the Cross-Media Code." How to Use Single-Source Measures to Examine Media Cannibalization and Convergence 50.2 (2010): 125-36. Print.
- Hassoun, Dan. "Tracing Attentions: Toward an Analysis of Simultaneous Media Use." Television and New Media. 15.4 (2014): 271-88. Print.
- Kattukaran, Anish. Evolution of TV Measuring across All Screens: Double Click, 2016. Web.
- Kosterich, Allie, and Philip M. Napoli. "Reconfiguring the Audience Commodity the Institutionalization of Social TV Analytics as Market Information Regime." Television & New Media (2015): 1527476415597480. Print.
- Napoli, Philip M. Audience Economics : Media Institutions and the Audience Marketplace. New York, NY, USA: Columbia University Press, 2003. Print.
- Nussbaum, Emily. "The Price Is Right." The New Yorker 2015: 95-99. Print.
- Steel, Emily. "Nielsen Plays Catch-up as Streaming Era Wreaks Havoc on TV Raters." The New York Times (2016). Web. 2 February 2016.
- The Nielsen Company. Total Audience Report Q1 2015:

The Nielsen Company, 2015. Web.

Wu, Xindong, et al. "Data Mining with Big Data." Knowledge and Data Engineering, IEEE Transactions on 26.1 (2014): 97-107. Print.

Yahr, Emily. "How Much TV Do Millennials Watch a Day? Depends on What Kind of Millennial You Are." The Washington Post (2016). Web. March 31 2016.

Yahr, Emily. "What Counts as a 'Hit' TV Show These Days? Don't Worry—Even the Experts Have No Idea." The Washington Post (2016). Web. 13 April 2016.

Zomorodi, Manoush. "The Ad Blocker's Dilemma: Sell Your Soul or Destroy the Internet." *Note to Self*. WNYC. New York, New York, 23 Sept. 2015. Radio.

Zomorodi, Manoush. "The Case for Infomagical." *Note to Self*. WNYC. New York, New York, 26 Jan 2016. Radio.

**“Hi Grandma, it’s me!”: Taking advantage of the
aging mind
Samantha Lacey (CLA 2017)**

Abstract

In light of a growing aging population and increasing numbers of older adults falling victim to financial scams each year, it is vital to determine what makes this population so highly targeted and susceptible to scams. In order to begin making progress on this problem, this paper explores the age-related cognitive processes of older adults, namely: attention, memory, decision-making, and executive functioning. Through synthesizing and analyzing research on these various processes, it is suggested that cognitive changes related to natural aging may contribute to the older populations increased likelihood to be targeted and susceptible to scams. Based on this suggested relationship between age-related cognitive changes and fraud, recommendations for future research directions are proposed. All in all, this paper aims to motivate future research that may be applied in the coming years to decrease the number of older adults that fall prey to financial fraud.

Introduction to fraud and the aging mind

One in five Americans age 65 and older fall victim to financial abuse and, in sum, are robbed of more than \$3 billion a year. In 2008, the Federal Trade Commission reported that people 60 years and older made up 10% of all fraud complaints, the lowest of any age group. Over a short span of four years, however, this percentage rose drastically. By 2012, older adults made up 26% of all fraud complaints, the highest of any age group (Browning, 2013). Unfortunately, there are no signs to suggest that this trend will change course and regress in the coming years. In fact, it is speculated that this percentage will only continue to rise as the number of people age 65+ is expected to more than double by the year 2060 (44.7 million in 2013 to 98.2 million in 2060) (Administration on Aging, Administration for Community Living, & U.S. Department of Health and Human Services, 2014).

Thereby, in light of this growing aging population, and increasing numbers of older adults falling victim to financial scams, it is vital to determine what makes this group so vulnerable. By analyzing what is known about age-related cognitive changes, this paper aims to motivate future research that can hopefully be applied to decrease the number of older adults whom fall prey to financial abuse; especially before this population begins to burgeon.

Of significance, although the elderly population currently only comprises 14% of the U.S. population, it disproportionately holds 34% of the nation's wealth (AoA, ACL, & HHS, 2014; Browning, 2013). Coupled with the natural cognitive aging that this group stereotypically experiences, older individuals age 65 and older become prime targets of financial abuse, whether it is by an unknown scammer, or a trusted relative. Older

adults generally exhibit age-related changes in certain cognitive functions, namely: attention, memory, decision-making, and executive functioning—it is hypothesized that changes in these domains may contribute to increased scam susceptibility.

Attention

With regard to attentional processes in the aging brain, older adults begin to experience difficulty on tasks that require them to divide or switch their attention among multiple inputs or tasks (Glisky, 2007). As a result, they become more inclined to attend to tasks that are personally relevant and meaningful rather than to those that appear unimportant or emotionally irrelevant. This change is often explained by the socioemotional selectivity theory that holds that: as people age, they perceive future time to be increasingly limited, and, as a result, they tend to place a greater emphasis on emotionally meaningful goals than younger individuals (Fung & Carstensen, 2003). Fung and Carstensen (2003) presented evidence of this goal change theory in a series of studies that examined how older and younger adults respond to advertisements.

In their first study, Fung and Carstensen (2003) presented a diverse sample of participants with advertisements for products using three different kinds of appeals (slogans): (1) an emotionally-relevant version in which the product's slogan related to love and caring (i.e. "Capture those special moments"); (2) a knowledge-related version in which the product's slogan emphasized the future (i.e. "Capture the unexplored world"); and (3) a neutral version in which products were presented without a slogan. The researchers found that older adults rated all of the products presented more positively than younger adults, exhibiting a common positivity bias that

will be explained in more detail below (Fung & Carstensen, 2003). Furthermore, in support of the aforementioned socioemotional selectivity theory, the researchers found that: although younger adults remembered more information about the advertisements overall, older adults were found to have best remembered information from emotionally meaningful advertisements rather than from knowledge-related or neutral advertisements (Fung & Carstensen, 2003).

Considering these findings, Fung and Carstensen (2003) conducted a second study to examine how a person's perception of time (limited or extended) affects advertisement preferences in younger and older adults. As hypothesized, they first found that older adults tended to prefer advertisements with emotionally meaningful appeal, while younger adults preferred advertisements with knowledge-related appeal. However, researchers discovered that this preference could be manipulated in older adults when one's perception of time was extended. Thereby, when older adults were made to perceive their life expectancy as higher, their inherent preference for emotionally meaningful advertisements decreased to the level generally shown by younger adults (Fung & Carstensen, 2003). Taken together with the results of their first study, this supports and points to the effect of the attentional shift associated with aging: as a person grows older, he or she tends to view time as increasingly limited. As a result, emotional information becomes more salient and relied upon, and subsequently affects memory and decision-making processes.

Memory

As aforementioned, memory for information that is relevant to an older individual's emotionally meaningful goals is better preserved than memory for information

that is personally or emotionally irrelevant; specifically, knowledge-related information (Fung & Carstensen, 2003). Although this is often supported by the socioemotional selectivity theory described above, brain mechanisms provide an alternative, potentially more tangible explanation for this phenomenon. Specifically, declines in working memory have been found to often account for divided attention deficits and the attentional shift from knowledge-related information to more emotionally relevant information. With age, a person's working memory has difficulty actively manipulating, reorganizing, and integrating information from various sources to complete complex everyday tasks (Glisky, 2007). Accordingly, older adults tend to experience reduced information-processing speeds (Salthouse, 1994). Given that this would make it more difficult—requiring extra time and cognitive effort—for new knowledge-related information to be encoded, stored, and recalled, it is tenable that emotionally relevant information becomes more salient with age (Glisky, 2007; Salthouse, 1994).

In addition to changes in working memory, although semantic memory, memory of general world knowledge, largely remains intact in normal cognitive aging, episodic memory, memory of autobiographical events, is affected. Glisky (2007) reported that older adults maintain the “gist” of older memories, however, they lack detail, as well as spatial and temporal context. As mentioned above in Fung and Carstensen's (2003) first study, older adults are more likely to recall more positive information, exhibiting a positivity bias. This is especially evident when older individuals recall previous personal choices/experiences (Mather & Johnson, 2000 as cited in Fung & Carstensen, 2003). Given that older adults have a tendency to distort their autobiographical memories in a

positive direction (Kennedy, Mather, & Carstensen, 2004 as cited in Mather, 2006), older adults tend to be more susceptible to making poor judgments or decisions. For example, if an individual knows a piece of information, but cannot recall the source of that information, he or she could make a decision based on, potentially, faulty information from an unreliable source.

Decision-making

Some research, however, has suggested that compared to younger adults, the positivity bias most frequently exhibited by older adults may actually enhance decision-making, leading to more regular and consistent choices (Strough, Mehta, McFall, & Schuller, 2008; Tentori, Osherson, Hasher, & May, 2001). In their study, Strough, Mehta, McFall, and Schuller (2008) examined older and younger adults' decision-making processes in relation to the sunk-cost fallacy: a decision-making bias in which people tend to invest more resources in a situation that they made an investment in, versus a similar situation in which they did not make a personal, emotional, or monetary investment. For example, imagine that a couple is channel surfing on their personal television for a movie to watch. They start watching a romantic comedy that they've never seen before, and within the first 15 minutes they start falling asleep because it's a snooze. As a result, they decide to change the channel to find a more interesting program that's worth their time. However, imagine what this couple would have done if they had went out to see this specific romantic comedy at the movie theatre. In line with the sunk-cost fallacy, it is suggested that most individuals would stay and watch the movie despite the fact that it's underwhelming. Why? Because they paid to see it, they made an investment in it. People often rationalize the latter situation because they feel that they would be losing their money if they

didn't finish watching the lackluster movie. In actuality, this is an irrational decision because the money has already been spent and cannot be recovered.

Previous research has suggested, "sunk-cost decisions are motivated by loss avoidance" (Frisch, 1993 as cited in Strough et al., 2008). In processing information, considering that older adults tend to give more weight to positive information than their younger counterparts who give more weight to negative information (Carstensen & Mikels, 2005), it is posited that older adults are less likely to focus on loss aversion. In support of this, Strough et al. (2008) found that in comparison to younger adults, older adults were less likely to commit the sunk-cost fallacy. In fact, older adults were more likely to make more consistent, normatively correct decisions across the investment (i.e., "You paid \$10.95...") and no-investment conditions.

These results have further been supported by similar findings from Carpenter, Peters, Västfjäll, and Isen's (2013) card task in which participants were told that the goal of the task was to win the most money while avoiding losses. To win the most money, participants had to learn from feedback as to which decks of cards would produce the greatest gains and which decks would produce losses. They found that inducing a positive state in older adults, who were previously shown to be more motivated by a desire to avoid regret and maximize satisfaction (Carstensen & Mikels, 2005), improved their learning and subsequent decision performance, making better choices than younger adults to maximize their earnings (Carpenter et al., 2013).

Executive functioning

Related back to the conclusions made by Samanez-Larkin (2013), these more consistent, effective decisions have also been said to result from the maintenance of crystallized intelligence, the ability to use well-learned knowledge and experience, across one's lifespan. It is theorized that the preservation of crystallized intelligence may compensate for a decline in fluid intelligence, the ability to solve new problems with more recent learning. The working and long-term memory deficits associated with aging further justifies this decline in fluid intelligence, and adds support for an overall decline in executive functioning abilities. Executive function plays a role in almost all cognitive processes "that are involved in the planning, organization, coordination, implementation, and evaluation of," generally, nonroutine, novel tasks (Glisky, 2007). Therefore, with a reduction in executive control, specifically working memory processes, and thus a decline in fluid intelligence, older adults have been found to seek out less information than younger adults when faced with novel tasks that require decisions to be made or problems to be solved (Yoon et al., 2005).

For example, in a study conducted by Streufert et al. (1990, as cited in Yoon et al., 2005), older mid-level managers and younger mid-level managers were put into groups and asked to make various group decisions in an all-day decision simulation. Streufert et al. (1990 as cited in Yoon et al., 2005) observed that the group of older managers requested less information throughout the day than the younger group of managers. Similar to Samanez-Larkin (2013), these findings suggest that older individuals may compensate for their decline in fluid intelligence with their crystallized intelligence, using "their experience to choose and process relevant

information more effectively” (Meyer et al., 1995 as cited in Yoon et al., 2005). Therefore, this decline in information search with aging may not necessarily be problematic in some situations, as this behavior—a reliance on top-down processing—is similar to that used by experts to arrive at equivalent decisions (Meyer, Russo & Talbot, 1995 as cited in Peters, 2010).

Various studies presented in Kennedy and Mather (2007), however, suggest that the tendency for older adults to “generate fewer options, deliberate for less time and seek out and review less information—particularly negative information,” can lead to detrimental errors in everyday decision making. For example, a study found that errors in adherence to medication were most frequently due to older individuals not reading all of the information provided by a doctor (Willis, Dolan, & Bertrand, 1999 as cited in Kennedy & Mather, 2007). Additionally, when faced with a problem or a difficult decision, older adults tend to use avoidant strategies, which include either postponing a decision or delegating the responsibility to someone else. By circumventing the difficult decision, they are thereby able to avoid negative feelings (Yoon et al., 2005).

As aforementioned, older adults are more persuaded by information that is in line with their emotional goals (Fung & Carstensen, 2003) and are more motivated by a desire to avoid regret and maximize satisfaction (Carstensen & Mikels, 2005). To ensure positive feelings, older adults tend to rely on “the affective heuristic” when making decisions. The affective heuristic decision is an intuitive, nonanalytical decision, requiring minimal effort, which is heavily influenced by an individual’s current emotion (Kennedy & Mather, 2007). Kennedy and Mather (2007) describe: “when people have a strong

positive or negative affective reaction” to a certain stimulus, “the increased salience of emotion coupled with the declining memory,” and slower information-processing speed experienced in old age, “may lead to a greater reliance on the affective heuristic.” Furthermore, when an older adult feels pressured by time, he or she may also have a tendency to rely on the affective heuristic, which has been found to lead to poorer decision-making (Kennedy & Mather, 2007).

Age-related cognitive changes in relation to fraud

Contemplating the effects of natural cognitive changes on the decision-making processes of older adults aids in the investigation and potential explanation for why the aging population is so vulnerable to financial scams. The findings of the previous studies demonstrate that as the brain ages, there is a decreased dependence on the brain’s deliberative systems and an increased reliance on the brain’s affective processes. Although this cognitive shift does not always yield negative decisions, and, in fact, oftentimes leads to more consistent, beneficial judgments, this shift is also what forms the basis for many financial scams that feed off of this population’s new found reliance on emotion. One such scam, the “grandparent scam” (Sneed, 2015; CBS News, 2014), is a form of identity theft in which a scammer manipulates or evokes certain emotions in an older individual in order to trick him or her into making a hasty, unfavorable decision—which often results in deleterious emotional and financial consequences. In this scam scenario, the scammer targets an older person, generally over the age of 65, and pretends to be a close relative, such as a grandchild, that is in need of immediate monetary help. In a report by CBS News, Carter Evans, a former scammer describes a typical call:

You just say, “Hey, how are you, hi grandma, hi grandpa... I'm in a little bit of trouble right now. If I tell you, just keep it between us, I'm on vacation, but I got into a little accident, and I was arrested for a DUI.” You tell them, “Things got out of control, and I need you to send me the money” (2014).

Carter Evans later revealed that this type of scam knows no profession, no education level, because “once you get them [the older individual] emotionally involved, then they’ll do anything for you, basically” (CBS News, 2014). It works off of the idea, discussed in aforementioned studies (i.e., Fung & Carstensen, 2003), that older adults are more likely to attend to emotionally relevant information. In addition, there is an added time pressure involved in the “grandparent scam.” With a slower information-processing speed in the aging brain, this time pressure leads to a reliance on the affective heuristic, which causes the individual to make an impulsive, emotional decision without seeking additional information (Kennedy & Mather, 2007).

A grandmother who fell victim to this financial scam recounted, “You [the victim] are blinded by emotion. Totally blinded. You don’t think rationally when this happens. You know, your family comes first” (CBS News, 2014). Furthermore, she went on to say that falling prey to a scam such as this made her feel embarrassed, “stupid and gullible.” Therefore, to those who fall victim, these scams are about more than just losing money; it evokes mistrust and fear, not only, in the victim, but also, in his or her family, and may lead to decreased or revoked independence. Moreover, these scams have the ability to, and commonly, disrupt family ties. A granddaughter of a victim reported that, since the day her grandfather found

out he was scammed, she and he have only spoken twice over the phone. Now, every time she calls, he hangs up in fear that he is being deceived again (Sneed, 2015).

Unfortunately, many of these cases go unreported, and even if they are, they prove difficult to prosecute (National Council on Aging); once the money is gone, it is gone. Other financial scams that the elderly most commonly fall victim to, include: Medicare/health insurance fraud, counterfeit prescription drug scams, funeral and cemetery scams, fraudulent anti-aging product scams, various fraudulent telemarketing tactics, internet fraud, investment schemes, homeowner/reverse mortgage scams, and sweepstakes and lottery scams (NCOA). However, although these scams are becoming increasingly common, the National Council on Aging (NCOA) reports that over 90% of older adults who experience financial abuse, are being taken advantage of by those who they trust most. This adds more complexity to the situation, and can result in even greater emotional damage. In these cases, individuals not only have to grapple with this betrayal of trust, but they also must decide whether or not to report their loved ones and risk getting them in trouble with the law. Moreover, scam situations so close to home can also compel individuals to withdraw from society and reevaluate their relationships with the one's they love.

Taken together, although currently only one in seven Americans are age 65 and older, this group disproportionately holds 34% of the nation's wealth, making them prime targets for scammers (AoA, ACL, & HHS, 2014; Browning, 2013). Additionally, the normal cognitive changes experienced as a result of natural aging affects decision-making and judgment processes, which may lead to decreased financial capacity in certain cases.

Particularly, as an older individual begins to perceive time as increasingly limited, there is the cognitive shift that occurs that augments the older adult's reliance on his or her affective processes. It is this increased attention to emotionally salient information, coupled with deficits in working memory, long-term memory, and, overall, executive functioning that contribute to an older individual's heightened probability to fall prey to financial abuse. Moreover, with this growing population, the National Institute of Health estimates that "the number of people living with dementia could double in the next 40 years with an increase in the number of Americans who are age 65 or older—from 40 million today to more than 88 million in 2050" (NINDS). The additional cognitive deficits that come along with the development of various kinds of dementia amplify this financial capacity issue and the likelihood of an older individual to be taken advantage of financially.

Future directions

With financial abuse of the elderly becoming the "crime of the 21st century" (NCOA), it is critical that future research be directed toward 1) determining who is most vulnerable within the elderly population (i.e., males or females, those with more or less education, etc.) and 2) identifying additional factors that may contribute to a person's susceptibility (i.e., well-being, financial literacy, stereotypes, etc.). However, how can one begin to tackle these types of questions without being able to put people in unsuspecting situations in which they may fall victim to a mock scam? This is possibly the most challenging obstacle that researchers must circumvent in attempting to accurately, directly, and ethically measure a person's likelihood to fall prey to a scam and in determining what factors contribute to his or her susceptibility. For this reason, little research on the topic has been conducted.

Further, the empirical research that has attempted to tackle these obstacles presents a variety of limitations.

For example, data that is collected from police records to determine the characteristics of the common fraud victim is thought to be inaccurate due to the elderly population's known likelihood to underreport victimization (James, Boyle, & Bennett, 2014). In addition, collecting information from known fraud victims (i.e., by sampling from fraud hotlines) also has its limitations as the findings "could reflect reporting errors or biases," and may only "reflect the make-up of the sampled population" (James et al., 2014).

Aside from these empirical limitations, research on this topic is further limited due to the common, generally undisputed, perceptions that society perpetuates about 'typical' scam victims. Namely, it is stereotypically thought that older adults are more vulnerable to scams due to the age-related cognitive declines mentioned in this paper. Further, society often pins older adults to be increasingly vulnerable due to their age-related declines in physical ability, which lend to more sedentary lifestyles that make them more likely to be "at home to receive telemarketing phone calls" (James et al., 2014). Additionally, older adults are commonly thought of as "naively trusting, socially isolated, and unsophisticated about financial matters" (Friedman, 1992, and Jackson & Hafemeister, 2011 as cited in James et al., 2014).

However, what society neglects to remember is that although the above may be true for some older adults, this is not true for all—everyone ages differently mentally, physically, and socially. In fact, in 1996, AARP conducted a telephone poll and found that many of the prevalent perceptions mentioned above were not in line

with the characteristics of actual older persons who reportedly fell victim to telemarketing scams that year: males who were married, “relatively well-educated, informed and active, and affluent” (as cited in James et al., 2014).

This being said, to date, no research has explored how stereotypical perceptions of elderly fraud victims affect the elderly when it comes to making financial decisions and their concern about/susceptibility to financial scams. Research has suggested, however, that stereotypes about aging affect older individual's cognitive processes and performance on related tasks. In a study conducted by Barber and Mather (2013), for example, older adults were found to do worse on a memory test if they were told prior to the test that older people have poorer memory capacities than younger adults. However, Barber and Mather (2013) noted that although the negative stereotype decreased the *quantity* of items that older adults recalled, the threat also increased the *quality* of the items that they recalled—answering the questions they did respond to, correctly. Barber and Mather (2013) then concluded that negative age-based stereotypes induce more conservative responses that may aid in more accurate recall and recognition abilities.

Studies such as this have been explained using the motivation-based, regulatory focus model of stereotype threat, in which the type of threat that is induced changes how people pursue their goals (Seibt & Forster, 2004). Negative stereotypes are said to induce a “prevention focus” in which an individual becomes more risk-averse and conservative in his responses in order to avoid loss; this focus should foster thoughtfulness and accuracy (Barber & Mather, 2013; Seibt & Forster, 2004). Conversely, positive stereotypes have been associated

with inducing a “promotion focus” in which an individual becomes more risk-taking and liberal in his responses in order to maximize gains; this focus is associated with eagerness and has been shown to foster speed (Forster, Higgins, & Bianco, 2003; Seibt & Forster, 2004).

Given that older individuals are more likely to attend to and remember positive information (Fung & Carstensen, 2003), and that they generally exhibit slower informational processing speeds and tend to review less information in making everyday decisions (Kennedy & Mather, 2007), I suggest that inducing positive stereotypes about aging (i.e., older adults are more wise and experienced, make more rational decisions than younger adults, etc.) could be more damaging than perpetuating negative stereotypes about aging when considering an older adult’s susceptibility to scams. Thus, I propose that negative age-based stereotypes (i.e., older adults have flawed memory abilities, slower information processing speeds than younger adults, etc.) may actually make older adults more cautious in their decision-making processes and, in effect, may make them less susceptible to financial abuse. Extending upon fundamental studies and conducting novel, intradisciplinary research such as this—which integrates phenomena/findings from varying disciplines within the psychological sciences—is vital in coming years in order to safeguard the future well-being and financial security of the growing aging population.

References

- Administration of Aging (AoA), Administration for Community Living (ACL), U.S. Department of Health and Human Services (HHS). "A profile of older Americans: 2014" 2014. http://www.aoa.acl.gov/Aging_Statistics/Profile/2014/docs/2014-Profile.pdf
- Barber, Sarah J., and Mara Mather. "Stereotype threat can reduce older adults' memory errors." *The Quarterly Journal of Experimental Psychology*, vol. 66, no. 10, 2013, pp. 1888-1895.
- Browning, E.S. "Financial scammers increasingly target elderly Americans." *The Wall Street Journal*, 23 Dec.2013, <http://www.wsj.com/articles/SB10001424052702303330204579248292834035108>
- Carpenter, Stephanie M., et al. "Positive feelings facilitate working memory and complex decision making among older adults." *Cognition & Emotion*, vol. 27, no. 1, 2013, pp. 184-192.
- Carstensen, Laura L., and Joseph A. Mikels. "At the intersection of emotion and cognition aging and the positivity effect." *Current Directions in Psychological Science*, vol. 14, no. 3, 2005, pp. 117-121.
- CBS News. "Grandparent scam" explained: What you need to know." 17 Apr. 2014. <http://www.cbsnews.com/news/grandparent-scam-explained-by-former-scammer-what-you-need-to-know/>

- Eisenberg, R. "Why elder financial abuse is such a slippery slope." *Forbes*, 13 Feb. 2015, <http://www.forbes.com/sites/nextavenue/2015/02/13/why-elder-financial-abuse-is-such-a-slippery-crime/>
- Förster, Jens, E. Tory Higgins, and Amy Taylor Bianco. "Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns?" *Organizational Behavior and Human Decision Processes*, vol. 90, no. 1, 2003, pp. 148-164.
- Fung, Helene H., and Laura L. Carstensen. "Sending memorable messages to the old: age differences in preferences and memory for advertisements." *Journal of Personality and Social Psychology*, vol. 85, no. 1, 2003, pp. 163-178.
- Glisky, Elizabeth L. "Changes in cognitive function in human aging." *Brain Aging: Models, Methods, and Mechanisms*, 2007, pp. 3-20.
- James, Bryan D., Patricia A. Boyle, and David A. Bennett. "Correlates of susceptibility to scams in older adults without dementia." *Journal of Elder Abuse & Neglect*, vol. 26, no. 2, 2014, pp. 107-122.
- Kennedy, Quinn, and Mara Mather. "Aging, affect and decision making." *Do Emotions Help or Hurt Decision Making? A Hedgfoxian Perspective*, 2007, pp. 245-265.
- Löckenhoff, Corinna E., and Laura L. Carstensen. "Aging, emotion, and health-related decision strategies:

motivational manipulations can reduce age differences.” *Psychology and Aging*, vol. 22, no. 1, 2007, pp. 134-146.

Mather, Mara. “A review of decision-making processes: Weighing the risks and benefits of aging.” *Committee on Aging Frontiers in Social Psychology, Personality, and Adult Development Psychology*, vol. 64, no. 145, 2006, pp. 145-173.

National Council on Aging (NCOA). “Top 10 financial scams targeting seniors.” <https://www.ncoa.org/economic-security/money-management/scams-security/top-10-scams-targeting-seniors/>

National Institute of Neurological Disorders and Stroke (NINDS). “Dementia: Hope Through Research.” 11 July 2014. http://www.ninds.nih.gov/disorders/dementias/detail_dementia.htm

Peters, Ellen. “Aging-related changes in decision making.” *The aging consumer: Perspectives from psychology and economics*, 2010, pp. 75-101.

Peters, Ellen, et al. “Adult age differences in dual information processes: Implications for the role of affective and deliberative processes in older adults’ decision making.” *Perspectives on Psychological Science*, vol. 2, no. 1, 2007, pp. 1-23.

Rosenblatt, Carolyn. “Preventing financial abuse and exploitation of elderly parents.” *AgingCare.com*, 2010.

- Samanez-Larkin, Gregory R. "Financial decision making and the aging brain." *APS Observer*, vol. 26, no. 5, 2013, pp. 30-33.
- Salthouse, Timothy A. "The aging of working memory." *Neuropsychology*, vol. 8, no. 4, 1994, pp. 535-543.
- Seibt, Beate, and Jens Förster. "Stereotype threat and performance: how self-stereotypes influence processing by inducing regulatory foci." *Journal of Personality and Social Psychology*, vol. 87, no. 1, 2004, p. 38.
- Sneed, C. "The grandparent scam." *The New York Times*, 9 Apr. 2015.
- Strough, JoNell, et al. "Are older adults less subject to the sunk-cost fallacy than younger adults?" *Psychological Science*, vol. 19, no. 7, 2008, pp. 650-652.
- Tentori, Katya, et al. "Wisdom and aging: Irrational preferences in college students but not older adults." 2001.
- Yoon, Carolyn, et al. "Cognition, persuasion and decision making in older consumers." *Marketing Letters*, vol. 16, no. 3, 2005, pp. 429-441.

**Barbara Strozzi's Feminine Influence on the Cantata
in 17th-Century Venice
Rachel Rubin (CLA 2017)**

Abstract

In this paper, I will be discussing how Venetian composer Barbara Strozzi (1619-1677) achieved success in the seventeenth century through publication. She is recognized today as one of the earliest composers to write and publish cantatas in such high quantities, though her accomplishments were not met without difficulty. Attaining recognition in the male-dominated musical culture of the Baroque era was a massive undertaking for a woman and required Strozzi to manipulate resources granted to her by her father's social status. However, it was her own relentless ambition that ultimately propelled her to fame. As such, she paved the way for female composers and continues to inspire with her legacy. Analysis of her piece, *Lagrime mie* from *Diporti di Euterpe*, published in 1659, reveals a highly refined use of expressive techniques to convey the powers of song and sorrow. Her affinity for emotional expression combined with her sophisticated understanding of the human voice create a powerful yet sensitive tone in this piece. Given the physical evidence, Strozzi's success through publication is undeniable, though it was her bold pursuit for recognition that allowed her to break gender stereotypes, making her feats all the more impressive.

Introduction

Barbara Strozzi (1619-1677) was both a highly praised composer and vocal performer of the seventeenth century. She made significant contributions to Venetian arias and ariettas but is most widely recognized as one of the earliest composers, male or female, to write cantatas in such high numbers (Timms). From 1644 to 1664, Strozzi published eight volumes of music containing over 100 works, with particular emphasis on the soprano voice. From the seventh collection, *Diporti di Euterpe* (“Euterpe’s Recreations”), also known as *Amusements of Euterpe* or *Cantatas and Ariettas for Solo Voice*, published in 1659, comes *Lagrima mie*, or *My tears* (Robin 347). Analysis of this piece reveals a highly cultivated understanding of expression that she explored through the subject of unrequited love, homing in on the power of both song and sorrow. Strozzi’s music as presented in this piece has the unique capacity to “shock the reader...with a raw evocation of emotion” (Beer 68-69). While her successes as both vocalist and composer are appropriately acknowledged today, one must consider that her gender, given the time and place, makes these feats all the more impressive. Not to mention, it was through her exploration of the secular cantata in a musical culture primarily dominated by opera that she achieved her success. She left a lasting impact on composers and musicians, particularly female, in the generations that followed, leaving her a distinguished legacy as “one of the first secular female composers in Western Europe” (Robin 346).

Strozzi’s childhood and privileged upbringing

Strozzi’s success was inarguably catalyzed by her father, renowned poet and librettist Giulio Strozzi. He is recognized for having provided libretti (text for extended vocal works) for noteworthy figures like Claudio

Monteverdi and Francesco Cavalli, esteemed composers of the Renaissance and Baroque eras. Giulio began to cultivate Barbara's vocal talent when she was very young, arranging for her to study under Cavalli, who composed forty-one operas in his lifetime (Jezic 26). It appears, however, that Barbara never explored opera through composition or performance, which is peculiar given her relationship to Cavalli and her father's specialization in libretti for opera (Pendle 105). Having attained the title "virtuossima cantatrice" [most virtuosic singer] by the age of 17, it is certain that she had the proper training and vocal agility to succeed in a genre as demanding as opera (Rosand 242).

Strozzi was exposed to intellectual elites primarily through her father's Accademia degli Unisoni (Academy of the Like-Minded), which was created as a subset of the Accademia degli Incogniti (Academy of the Unknowns) (Pendle 106). Until this point, Strozzi had primarily performed and studied in private, as was typical for female musicians in the Baroque era. As her skills became more advanced, however, her father sought to create an appropriate environment in which to share her talent.

It appears that Giulio Strozzi had a progressive view of women and their societal role in general. During the sixteenth century, women were expected to be proficient in instrumental and vocal music, though they were discouraged from performing in the presence of others (with the exception of nuns and those born to professional musical families) (Pappano). By the following century, it was much more common for women to perform and compose, though to become a *published* female composer was less likely. Giulio's "forward-thinking" stance on women and his efforts to further his daughter's success likely played a part in her bold pursuit

of compositional fame. It should be noted early on, however, that without Barbara's own spirit and wit, she may have never reached such great heights, even with her father's support.

In 1637, Giulio "designed at least in part, to exhibit her talents to a wider audience," the Accademia degli Unisoni—a unique group of male intellectuals who gathered to discuss relevant topics (Rosand 244). The members of the academy were important figures in Venice, including distinguished writers, poets, and philosophers (Jezic). Discussions usually focused on politics, academic and social discourse, history, philosophy, poetry and literature. Strozzi's access to the academy granted her significant exposure when her presence was permitted (Rosand 245). Music was often a point of reference during these meetings, or was simply a form of entertainment. In these cases, musicians like Strozzi were invited to perform, though this opportunity was still a great honor. Women were not usually permitted into the academy, but given her father's status, exceptions were made for her.

It also should be mentioned that her impressive vocal talent and intellectual contributions helped her gain respect among the members, which certainly contributed to the events that followed. Eventually, she became a respected source of insight and wisdom, though she never became an official member of the academy (Rosand 242). While this was an important stepping stone for Strozzi and a comparatively large feat for a woman, it still limited her to a very small group of people in the public sphere (Jezic). Her biggest outreach by far was accomplished through the mass publication of her compositions.

Since her childhood was handled by her father in what appears to be a highly structured manner, Strozzi's upbringing afforded her the resources that ultimately led to her success. It was due to his "high stature among the Venetian intellectual elite that assured Barbara connections and opportunities for patronage" (Robin 346). This gave Strozzi an edge over her female peers, who had a significantly harder time making the same headway.

During Strozzi's time with the academy, she was busy composing her own works. She was aware of the strides she had already made, but was not nearly finished. She was determined to take the next step and publish her works despite her "feminine weaknesses" (Rosand 256). Strozzi knew that she could take advantage of her father's connections and resources and began to perform her own compositions in the academy. This is perhaps why the genre in which she composed was so limited, because the pieces needed to suit the "elite in a seventeenth-century Venetian drawing room" (Jezic). This is also arguably why so many of her pieces are well suited for the soprano voice: she was composing them for herself to sing. These opportunities, though slightly manipulated, allowed her to move towards a professional career as a published composer, about which there were few female success stories at the time.

Strozzi's forge into new territory: publication

In Strozzi's early works, such as opuses one through three, Giulio Strozzi provided the libretti for Barbara's compositions as a tool for enhancing their worth in the realm of music publishing. With his name on them, the compositions were likelier to be appreciated. However, this deal did not last long and seemed to have little to no effect on the outcome of Strozzi's success in publishing.

She quickly developed her own style of expression by utilizing text with topics centering around love and sorrow from other poets. The seven collections composed from 1651 to 1664, of which three used Giulio's text before his death in 1652, "went far beyond the achievements of her opus 1" (Beer 65). These compositions were more vocally challenging, impressive in their virtuosity, and highly dramatic, illustrating Strozzi's "ever-growing ambition" (Beer 65).

Most of Strozzi's surviving works are ariettas, arias and cantatas for solo voice (mainly soprano) and continuo (Glixon). While she was afforded significant opportunity for exposure, she was still limited to smaller works, unable to make strides in the realm of opera (though it is unclear if she ever tried). Strozzi was a notable contributor to the cantata, a comparatively neglected genre at the time in which poetic texts and lyric monologues are explored (Timms). The cantata in particular was seen in Italian vocal chamber music from the early seventeenth-late eighteenth centuries, but was largely outshined by Venetian opera (Timms). Over time, the parameters of the seventeenth-century cantata changed, ranging from strophic arias of pre-1620s Venice to the secular cantata of the mid-1600s that were divided into sections. Characteristically, Strozzi's works consisted of a type of refrain, sections of recitative, arioso, and aria, and long melismatic sections with repetition, which makes sense for the time of their conception relative to the evolution of the cantata. All of these traits were especially well suited for the lyric soprano. Strozzi took advantage of her own success as a vocalist by featuring the soprano voice in her work, though she did "[experiment] with musical forms" (Robin 347). This allowed her to perform her own compositions within the academy and other elite groups where her

talent was sought after, granting her works a higher rate of success. The long, melismatic passages in the upper register also lent themselves to the highly emotional themes that Strozzi was exploring.

In general, Strozzi was a proponent of *seconda prattica*, in which the devotion to expressive text took precedence over other contrapuntal standards, as displayed in Strozzi's work by "her faithful adherence to her texts" (Rosand 273). The term *seconda prattica* or "second practice" was developed by Monteverdi in the late 1500s in response to criticisms of the rigid standards of *prima prattica* or "first practice," in which the full potential of the text could not be realized due to the restrictions of traditional contrapuntal rules. "Second practice" is usually homophonic secular music.

Origins of the lament

One piece that falls into this category comes from Barbara Strozzi's Opus 7, *Diporti di Euterpe* or *Pleasures of Euterpe*. Among the collection, the secular cantata, *Lagrime mie*, borrows text from the Venetian nobleman Pietro Dolfin. It tells the story of a man lamenting his lover, Lidia, who has been imprisoned by her father. The text informs of the main figure's unrelenting anguish, which prevents his tears from flowing as he repeatedly asks, "My tears, why do you hold back?" These rhetorical questions are followed by recollections of his beloved Lidia, and he begs for death if he cannot have her. This piece epitomizes Strozzi's fascination with the power of emotion through song by using an evocative poem combined with expressive techniques. It also hints at the prominence of the Italian dramatic lament in the seventeenth century. The lament, meaning "a passionate expression of grief and sorrow" as defined by Oxford University Press Dictionaries, was quite a popular genre

in seventeenth-century Italy. It was frequently explored through poetry and song (Pappano). Whereas Strozzi was near obsessed with exploring anguish and suffering in her work, she was not alone in this pursuit, following in the footsteps of great composers such as Peri, Monteverdi, Caccini and others.

Many of these notable composers emulated ancient Greek and Roman ideals, particularly those of Aristotle. His theory of catharsis, “the purification of the emotions through art,” later emerged in the Italian lament of the seventeenth century (Pappano). Art’s role was then to move its audience in some extreme emotional way, either to pity or to sorrow (Pappano). An especially popular character of the Italian lament was the “madwoman,” who may have been wronged in love and sings of her fate. These moments in the plot are often accentuated by dissonant harmonies (Pappano). This craft can be referred to as text painting, a common feature of “second practice” in which musical techniques are used to enhance the meaning of the text, sometimes at the expense of traditional musical standards. *Lagrime mie* is full of examples of this.

Mastery in practice

At the beginning of *Lagrime mie*, “the singer lets rip an astonishingly doleful wail, from the very top of her vocal register, falling down and down, over a stationary harmony” (Beer 68). The tone is evident immediately, with this drawn-out chromatic phrase down the scale on the word “lagrime” (tears), including trills that suggest the falling of tears.



Figure 1.

After this opening, the music shifts to a slightly faster tempo, as the translation follows with “why do you hold back?” This entire phrase repeats towards the middle of the piece, with the same emotional intensity and stylistic traits.

As discussed earlier, Strozzi was lucky to have been granted access to great intellectual minds of the seventeenth century through her father’s academy. Although she was not an official member, she was still encouraged to participate in meaningful discussion. It could be argued, then, that one such debate that “concerned whether tears or song be the more potent weapon in love” inspired the fusion of the two in this piece (Rosand 245). Strozzi’s opinion resonated in this argument when the decision was made that song was clearly the most effective. In response to this, Strozzi said “I do not question your decision, gentlemen, in favor of song; for well I know that I would not have received the honor of your presence at our last session had I invited you to see me cry and not to hear me sing” (Rosand 279). Despite her mature and witty response to this debate, Strozzi still “borrow[s] the power of tears” to strengthen her music’s potency (Rosand 279). She believed she could accomplish something magnificent in *Lagrime mie* by combining the powers of song and sorrow, which are so

persuasive individually. These emotions were a source of inspiration in many of her works.

Although the emotional intensity of the poetry can stand alone, the content is heightened by Strozzi's use of expressive techniques. As is typical of second practice, Strozzi uses text painting to emphasize emotion, creating an intensity that is not exclusive to the words, but seen in the use of bold techniques, like chromaticism, syncopation, and leaps in the melody. The speech-like declamation and homophonic texture throughout the piece elucidate the text's meaning and encourages text painting; for example, the "scalar descent representing downward-flowing tears" (OAWM 321). On negative words like "suffering," "harsh," and "torment," Strozzi uses unsettling harmonies to create a pained listening experience. She cleverly inserts an eighth rest into the middle of the word "respiro," or breath, to stress the meaning of the line, "wash away the pain that takes my breath."



Figure 2.

She conveys "the superior pnce of the naked human instrument" through a series of melismatic passages,

much like the one featured in the opening phrase (Rosand 274).

Performance practice

Achieving a historically informed performance of this piece requires an understanding of Baroque period music (1600-1750s) in combination with textual evidence left for its performers in written scores. The musical score, or the physical document on which the music is notated, reveals information to be interpreted by the performer. The amount of information provided varies depending on the time period. During the Baroque era, the performer's interpretation was deemed as important as the written score, if not more so. The result is a relatively sparse notation of musical practice or nuanced techniques, with little indication of dynamics, tempo, and articulation, leaving a significant amount of creative freedom to the performer (Cyr 23).

While the performer's interpretation of Baroque music results in numerous versions of one piece of music, a few traits remain constant as evidenced by *Lagrime mie*. These qualities include the text's significance as an expressive tool and the use of basso continuo, a form of accompaniment that provides "a foundation or harmonic support" with a prominent bassline and improvised upper harmonies (Cyr 24). The basso continuo can be played by a number of instruments, sometimes indicated in the score but frequently chosen by the performer. Instruments that might be appropriate for the time period are the harpsichord, organ, lute, harp, and viola da gamba (Cyr 24). All of these instruments are capable of providing a strong bassline, which is notated in the score, and improvised chords, which are not notated. Many of these instruments also have modern counterparts, such as the harpsichord, whose keyboard is similar to that of a

piano, or viola da gamba, a moderately sized string instrument that can be plucked or bowed, like the cello. Modern performances of *Lagrima mie* are, as expected, varied in their instrumental accompaniment, though they maintain a distinctly expressive quality. The sparseness of the notation indicates how crucial the performer's interpretation was in conveying the music's meaning. This is particularly relevant to Strozzi, who performed her own pieces and took advantages of opportunities for heightened emotional expression in *Lagrima mie*'s text painting.

Strozzi's Legacy

This piece in particular is a perfect representation of Strozzi's heroic impact. It encompasses everything she stood for, from her ambitious pursuit for success to her affinity for emotional expression. It represents the wisdom and experience she accumulated while performing in the academy, as well as her understanding of the soprano voice, which she displayed in her performances and compositions. While she made significant contributions to the genre of the cantata, especially to the topic of unrequited love, her greatest achievement is arguably the "precedent [she set] for women's careers as published composers" (Rosand 266). She indeed was aided by her father's connections, but her talent was a huge asset to her success. Nevertheless, she managed to create a highly individualized niche for herself in a time when breaking out of the operatic mold was difficult, countering every threat to her progress with relentless determination. *Lagrima mie* epitomizes all of these factors, including Strozzi's grasp of the cantata, the soprano voice, *seconda prattica*, and secular music. Although there were many female musicians at the time, few were able to achieve to the same extent as Strozzi,

making her one of the first women to gain such success in seventeenth-century Venice.

Strozzi may have died nearly 350 years ago, but her legacy certainly has not. She is an inspiration for women battling for equal representation in male dominated fields – an issue that transcends music. Certainly, evidence supports the technicality of her success in the sheer number of her published compositions. More heartening, however, is the fortitude with which she achieved her success and the example she boldly set for women cross-generationally, from the seventeenth century to today.

References

- Beer, Anna. *Sounds and Sweet Airs: The Forgotten Women of Classical Music*. London: Oneworld Publications, 2016. Print.
- Cyr, Mary. *Performing Baroque Music*. Portland, Oregon: Amadeus Press, 1992.
- Glixon, Beth L. and Ellen Rosand. "Strozzi, Barbara." *Grove Music Online*. *Oxford Music Online*. Oxford University Press. Web. 20 Sep. 2016.
- Jezic, Diane, and Elizabeth Wood. *Women Composers: The Lost Tradition Found*. New York: Feminist at the City U of New York, 1994. Print.
- Pappano, Annalisa. "The Italian Dramatic Lament." *NAXOS*. Naxos Digital Services LTD, 15 Nov. 2005. http://www.naxos.com/catalogue/item.asp?item_code=8.557538
- Pendle, Karin Anna. *Women and Music: A History*. 2nd ed. N.p.: Indiana UP, 2001. Print.
- Robin, Diane Maury, Anne R. Larsen, and Carole Levin. *Encyclopedia of Women in the Renaissance: Italy, France, and England*. Santa Barbara: ABC-CLIO, 2007. Print.
- Rosand, Ellen. "Barbara Strozzi, Virtuosissima Cantatrice: The Composer's Voice." *Journal of the American Musicological Society* 31.2 (1985): 241-81. *JSTOR* [*JSTOR*].

Timms, Colin et al. "Cantata." *Grove Music Online*. Oxford Music Online. Oxford University Press. Web. 20 Sep. 2016.

On the Constructability of Numbers
Jeff Moorhead (CLA 2018)

Abstract

Before the advent of modern mathematics, mathematics was largely thought of in geometric terms, where numbers were constructed using a compass and straightedge. In this paper, I aim to critically analyze the mathematical implications of constructability, as well as the role field theory and polynomials play in the constructability of numbers using the properties of several number systems. I will also analyze the three classical problems posed by Ancient Greek mathematicians, followed by a detailed explanation and proof of each of these three problems. Finally, I will explain the relationship between field theory and the constructability of regular polygons and how abstract algebra can be used to show that certain polygons cannot be constructed with a compass and straightedge.

Introduction

Ancient Greece provided some of the most important contributions to mathematics in history. The earliest first-hand account of Greek mathematics comes from Euclid, whose most famous work, *Elements* (circa 300 BCE), describes many important geometrical propositions. The most important form of Greek mathematics was geometry (1). Many Greek geometers made discoveries that are still important today, including Archimedes who showed that the surface area of a sphere is equal to four times the area of the sphere's largest circle, i.e. $S = 4\pi r^2$ (11), and Pythagoras who most famously discovered that "the area of the square built upon the hypotenuse of a right triangle is equal to the sum of the areas of the squares upon the remaining sides," i.e. $a^2 + b^2 = c^2$ (Morris).

Constructible Numbers

In this paper we will discuss the constructability of numbers. The ancient Greeks were able to use their methods of geometry to "construct" some numbers using a straightedge and compass. Based on their constructions, the Greeks posed three famous questions:

1. Is it possible to construct a cube with twice the volume of a given cube?
2. Is it possible to construct a square with the same area as a given circle?
3. Is it possible in general to trisect a given angle θ ?

To understand these problems, we must first discuss what it means for a number to be constructible.

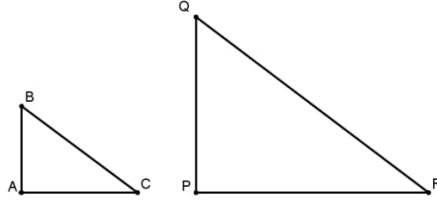


Figure 1. Similar triangles (“SAT”)

We can show that multiplication and division are given by considering a set of similar triangles. From Figure 1, we have triangles ABC and PQR . Let $AC = 1$. By the properties of similar triangles, we have $\frac{1}{BC} = \frac{PR}{QR}$. Cross-multiplying gives $QR = (BC)(PR)$. Similarly, if we let $AC = 1$, then $\frac{1}{BC} = \frac{PR}{QR}$ which gives $BC = \frac{QR}{PR}$. Since it is possible to construct a similar triangle if we are given just one side length of the second triangle, we can construct the product and quotient of any two lengths. In addition to the four operations above, we can also construct square roots.

Proposition 1.3. *Square roots are constructible.*

Proof. By Thales’ Theorem we know that “an inscribed angle in a semicircle is a right angle” (Weisstein, “Thales’ Theorem”). From Figure 2, we know that angle ACB is a right angle because AB is the diameter of circle O . We can draw a segment starting at C perpendicular to AB . Label the intersection of AB and our segment point D . Let DB have length 1. Because angle C is a right angle, we know that triangles ACD and BCD are similar (this follows from elementary properties of triangles). Thus, $\frac{1}{CD} = \frac{CD}{AD}$.

It follows that

$$AD = (CD)^2 \rightarrow CD = \sqrt{AD}.$$

Therefore, square roots are constructible (Harris 14 Oct).

□

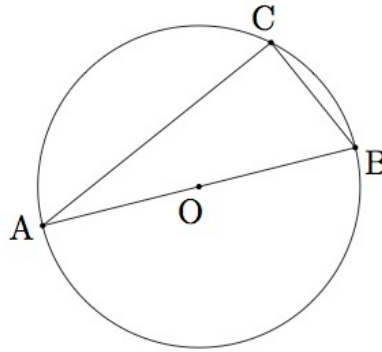


Figure 2. Constructability of square roots (“G-C Right Triangles”)

Field Theory and Constructability

Because compass and straightedge constructions give addition, subtraction, multiplication, and division in \mathbb{Z} , we can construct all the rationals \mathbb{Q} . Additionally, we can construct square roots. Using these five operations, we can already find examples of constructible numbers such as: $\sqrt{2}$, $\sqrt[4]{2}$, (i.e., $\sqrt{\sqrt{2}}$), $6\sqrt{2 + 5(\sqrt{2})}$, etc. We notice that because we can construct square roots, we can construct certain irrational numbers (but not all irrationals, as we will show). Therefore, the constructible numbers form a subfield of \mathbb{R} strictly larger than \mathbb{Q} (Dummit and Foote).

As shown above, we know that numbers are constructible in four ways: connecting two points by a line, intersecting two lines, intersecting a line and a circle, and intersecting two circles. If we consider a field F , the intersection of lines $a_1x + b_1y = c_1$ and $a_2x + b_2y = c_2$ with coefficients in F yields a system of equations with solutions x and y in F (Harris 19 Oct). The intersection of lines, therefore, will not produce new points that are not already in F . However, if we consider the intersection of a line and a circle, the problem is reduced to solving a system of two equations, $ax + by = c$ and $(x - h)^2 + (y - k)^2 = r^2$ where a, b, c, h, k, r are in F (Dummit and Foote). The intersection of two circles is equivalent to solving a system of two equations: $(x - h_1)^2 + (y - k_1)^2 = r_1^2$ and $(x - h_2)^2 + (y - k_2)^2 = r_2^2$. Solving either of these for x will give a quadratic equation with solutions in an *extension* of our field F (the quadratic equation gives solutions more precisely in a quadratic extension of F , which will be discussed below) (Dummit and Foote).

Definition 2.1. If K is a field containing the subfield F , then K is said to be an *extension field* (or simply an *extension* of F) denoted K/F . The field F is sometimes called the *base field* of the extension (Dummit and Foote).

An important consequence of extension fields is that for an extension K/F , K forms a vector space over F . The *degree* of the extension field K/F is equal to the dimension of the vector space of K over F , denoted $[K:F]$ (Dummit and Foote). Certain extension fields have a close connection to polynomials. We know that the equation $x^2 + 1 = 0$ has no solutions in \mathbb{R} . Pretending we have no knowledge of a field beyond \mathbb{R} , does there exist an extension of \mathbb{R} for which our polynomial has a root? Well, $i = \sqrt{-1}$ is a solution to our polynomial. So if we extend

\mathbb{R} by i , we have a field that contains a solution to $x^2 + 1 = 0$. We denote this extension $\mathbb{R}(i)$. This is the smallest extension of \mathbb{R} containing i . To generalize this example, it can be shown that any polynomial with coefficients in \mathbb{R} has at least one root, $z = x + iy$, in the complex numbers \mathbb{C} , which means that \mathbb{C} is an extension field of \mathbb{R} . Next, we will find the degree of \mathbb{C} over \mathbb{R} .

We know that $x^2 + 1 = 0$ has degree 2. We will show that the degree of an irreducible polynomial with roots in \mathbb{C} is the dimension of the vector space \mathbb{C} over \mathbb{R} .

Theorem 2.2 (Dummit and Foote 13.1.4). *Let $p(x) \in F[x]$ be an irreducible polynomial of degree n over the field F and let K be the field $F[x]/(p(x))$. Let $\theta = x \bmod (p(x)) \in K$. Then the elements $1, \theta, \theta^2, \dots, \theta^{n-1}$ are a basis for K as a vector space over F , so the degree of the extension is n , i.e., $[K:F] = n$. Hence,*

$$K = \{a_0 + a_1\theta + a_2\theta^2 + \dots + a_{n-1}\theta^{n-1} : a_0, a_1, \dots, a_{n-1} \in F\}$$

consists of all polynomials of degree less than n in θ .

The above theorem may appear daunting at first, so it might be helpful to relate the theorem to our previous example. In our example, $p(x) = x^2 + 1$, $F = \mathbb{R}$ and the extension $K = \mathbb{C}$. It is easy to see that the polynomial $x^2 + 1$ is irreducible over \mathbb{R} . If we let $\theta = i$, then we know that $1 + i^{2-1} = 1 + i$. The set $\{1, i\}$, then, should form a basis for \mathbb{C} by Theorem 2.2. Converting $\{1, i\}$ into column vector notation and solving

$$c_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + c_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

gives solutions $c_1 = a$ and $c_2 = b$ for $c_1, c_2 \in \mathbb{R}$, where a and b are coefficients of any linear combination of 1 and i . So $\{1, i\}$ spans \mathbb{C} and is therefore a basis for \mathbb{C} and the dimension of the vector space \mathbb{C} over \mathbb{R} ,

$$\dim_{\mathbb{R}}\mathbb{C} = \deg(x^2 + 1) = 2.$$

If we change our polynomial to $x^2 - 2 = 0$ (irreducible over \mathbb{Q}), we see that the extension $\mathbb{Q}(\sqrt{2})$ (which denotes the smallest field containing \mathbb{Q} and $\sqrt{2}$) has a root to our new polynomial, i.e. $\theta = \sqrt{2}$. We can confirm that $\{1, \sqrt{2}\}$ forms a basis for our new extension field $\mathbb{Q}(\sqrt{2})$, so by Theorem 2.2 we have $[\mathbb{Q}(\sqrt{2}) : \mathbb{Q}] = \deg(x^2 - 2) = 2$.

A final step in our development of the constructability of numbers is to define what it means for a number to be algebraic over a field F .

Definition 2.3. Let K be an extension field of F . An element $\alpha \in K$ is said to be *algebraic* over F if α is a root of some nonzero polynomial $f(x) \in F[x]$. If α is not algebraic over F (i.e., is not the root of any nonzero polynomial with coefficients in F), then α is said to be *transcendental* over F . The extension K/F is said to be algebraic if every element of K is algebraic over F (Dummit and Foote).

Consider the following:

1. The number i is algebraic over \mathbb{R} because it is a root of $x^2 + 1 = 0 \in \mathbb{R}[x]$.
2. The number π is transcendental over \mathbb{Q} because π is not a root of any polynomial with coefficients in \mathbb{Q} .

If we consider an element α , which is algebraic over, but not necessarily in our field F , there is a unique monic (has leading coefficient 1), irreducible polynomial with coefficients in F with α as a root. We will denote this polynomial as $m_\alpha(x)$. The degree of $m_\alpha(x)$ is called the *degree* of α (Dummit and Foote). It is important to note that if we consider the extension $F(\alpha)$ generated by α , then the degree of $m_\alpha(x)$ is equal to the dimension of $F(\alpha)$ as a vector space over F by Theorem 2.2. That is to say,

$$[F(\alpha):F] = \deg m_\alpha(x) \text{ (Dummit and Foote).}$$

We have already discussed this idea when we discussed the polynomial $x^2 - 2 = 0$. This polynomial is clearly monic. It can also be shown to be irreducible. Letting $\alpha = \sqrt{2}$, we have $m_{\sqrt{2}}(x) = x^2 - 2$, so $[\mathbb{Q}(\sqrt{2}):\mathbb{Q}] = \deg(x^2 - 2) = 2$ as shown above.

Quadratic extensions are an important class of field extensions connecting field theory to constructability. If we consider an element $\alpha = a + b\sqrt{c}$ in an extension K of a field F with $\alpha \notin F$ (but $a, b, c \in F$) and with $[K:F] = 2$, then our element α is the root of a polynomial of degree 2 with coefficients in F . This is a consequence of constructible numbers being the intersections of lines and circles, the equation of a circle having degree 2, and the equation of a line having degree 1. Quadratic extensions will have important consequences relating to the constructability of numbers, but first we must establish one more result relating to algebraic extensions.

Theorem 2.4 (Dummit and Foote 13.2.14). *Let $F \subseteq K \subseteq L$ be fields. Then*

$$[L:F] = [L:K][K:F]$$

i.e., extension degrees are multiplicative.

Let us consider an example. We know that $[\mathbb{Q}(\sqrt{2}):\mathbb{Q}] = 2$. If we extend \mathbb{Q} by a different number, such as $\sqrt[4]{2}$, we can ask, “What is $[\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}(\sqrt{2})]$?” We notice that $\sqrt[4]{2}$ is a solution to the irreducible polynomial $x^4 - 2 = 0$, so we know that $[\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}] = 4$ by Theorem 2.2. We have already established that $[\mathbb{Q}(\sqrt{2}):\mathbb{Q}] = 2$. Therefore, by Theorem 2.4, we have

$$[\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}] = [\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}(\sqrt{2})][\mathbb{Q}(\sqrt{2}):\mathbb{Q}],$$

or $4 = [\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}(\sqrt{2})]$, so $[\mathbb{Q}(\sqrt[4]{2}):\mathbb{Q}(\sqrt{2})] = 2$. We may confirm this result using the minimal polynomial method by noticing that $\sqrt[4]{2}$ is a solution to $x^2 - \sqrt{2} = 0$ (note that $x^2 - 2$ is an element of $\mathbb{Q}(\sqrt{2})[x]$).

Using the above theorems, we notice a property of constructible numbers. Because numbers can only be constructed by the intersection of lines and circles, we see that any new number we construct will be an element of at most a quadratic extension of a given field F . We may conclude from Theorems 2.2 and 2.4 that any new number in \mathbb{R} that we construct from a field of previously constructed numbers F has degree 2^n over F .

Theorem 2.5 *If a number α is constructed by any number of compass and straightedge constructions from a field F , then*

$$[F(\alpha):F] = 2^n, n \in \mathbb{Z}.$$

Proof. If we use points in F to construct a number α not in F , then α must be the intersection of a line and a circle or of two circles. In the case of two lines, we have a linear equation with coefficients in F that we know must have solutions in F itself, meaning $F(\alpha) = F$. Thus, no new point has been constructed. If α is constructed by a line and a circle, or by two circles, we have a quadratic equation with coefficients in F . The solution to the equation will be of the form $p + q\sqrt{r}$ with $p, q, r \in F$. We see that α is in $F(\sqrt{r})$, and so $[F(\alpha):F] = [F(\sqrt{r}):F]$. By Theorem 2.2, we have

$$[F(\alpha):F] = \deg m_\alpha(x) = 2.$$

Now let $\beta_1, \beta_2, \dots, \beta_{r-1}$, with $\beta_r = \alpha$, be all the numbers we construct on the way to α where each β_k for $1 \leq k \leq r$ generates a new quadratic extension of F . By Theorem 2.4 we have

$[F(\beta_1, \dots, \beta_r):F] = [F(\beta_1, \dots, \beta_r):F(\beta_1, \dots, \beta_{r-1})] \cdot [F(\beta_1, \dots, \beta_{r-1}):F(\beta_1, \dots, \beta_{r-2})] \cdots [F(\beta_1):F]$. Since each β_k generates a quadratic extension of F , we have $[F(\beta_1, \dots, \beta_r):F] \leq 2^r$. Therefore, $[F(\beta_1, \dots, \beta_r):F]$ is a power of two, 2^q for $q \leq r$. Because $\beta_r = \alpha$, we conclude that $[F(\alpha):F] = 2^n$ for $n \in \mathbb{Z}$, $n \geq 0$ (Harris 14 Oct).

□

Using the above results, we are able to draw an important conclusion about the constructability of a number α .

Theorem 2.6 *Every constructible number α has the following properties:*

1. α is algebraic over \mathbb{Q} .
2. The degree of the characteristic polynomial of α , $p(x) \in \mathbb{Q}[x]$ is a power of 2 (“Constructible Numbers”).

Classical Problems

We are now prepared to discuss the three problems posed by ancient Greek mathematicians. We will first discuss whether or not it is possible to “double the cube”, i.e., whether or not it is possible to create a cube with exactly twice the volume of a given cube. To begin, we have a cube with unit side lengths. The volume of this cube is $V = 1 \cdot 1 \cdot 1 = 1$. A cube of double this volume would have $V = 2$. So the question becomes: What side lengths will give a volume of 2 when cubed? Clearly $\sqrt[3]{2}$ works, so all we need to do is construct a cube with side lengths $\sqrt[3]{2}$. But we know $[\mathbb{Q}(\sqrt[3]{2}) : \mathbb{Q}] = \deg m_{\sqrt[3]{2}}(x) = \deg(x^3 - 2) = 3$, which is not a power of 2, so $\sqrt[3]{2}$ is not constructible by Theorem 2.5. Because $\sqrt[3]{2}$ is not constructible, we cannot possibly construct a cube with side lengths $\sqrt[3]{2}$, and so it is not possible to double the cube with side lengths one. Further, it is not possible to double *any* cube because that would require constructing a number α with degree 3 in \mathbb{Q} .

Next, we will show that it is not possible to “square the circle,” i.e. construct a square with precisely the same area as a given circle. If we consider a circle with radius 1, then $A_c = \pi(1)^2 = \pi$. So we must construct a square with area π . Such a square would have side lengths $\sqrt{\pi}$. But what is $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}]$? It is important to note that π was

proven transcendental over \mathbb{Q} by Ferdinand von Lindemann in 1882 (Weisstein, “Transcendental”). We will assume the transcendental of π without proof. Because π is transcendental, it follows that $\sqrt{\pi}$ is transcendental as well, so $\sqrt{\pi}$ does not have a minimal polynomial with coefficients in \mathbb{Q} . It follows that $[\mathbb{Q}(\sqrt{\pi}) : \mathbb{Q}] = \infty$, which is clearly not a power of 2. We therefore conclude that it is not possible to construct a square with side length $\sqrt{\pi}$, making it impossible to square the circle.

It is interesting to note that it is possible to come very close to squaring the circle. One of the most accurate compass and straightedge constructions which approximately squares a circle was given by Indian mathematician Srinivasa Ramanujan in 1913. In a paper published in *The Journal of the Indian Mathematical Society*, Ramanujan gives the approximation $\frac{355}{113}$ for π . This approximation is indeed very close to the true value of π , the difference between the two values being only about .00000027. At the end of his paper, Ramanujan notes that amazingly if the area of the constructed circle were 140,000 square miles, the side lengths of the square would be greater than the square’s true value by only about one inch (Ramanujan).

The third problem we will discuss is whether it is possible in general to trisect an angle. To begin, we see that a point with distance 1 from the origin has coordinates $(\cos(\theta), \sin(\theta))$. This means that our angle θ is constructible if and only if $\cos(\theta)$ and $\sin(\theta)$ are constructible (Dummit and Foote). It is possible to trisect certain angles. For example, an angle of 180° can be trisected because $\cos(60^\circ) = \frac{1}{2}$ and $\sin(60^\circ) = \frac{\sqrt{3}}{2}$, both of which are constructible. We can also trisect a 90° angle for

the same reason. However, it is not possible in general to trisect an angle ϑ . Notice that trisecting ϑ is equivalent to constructing $\cos(\frac{\vartheta}{3})$ and $\sin(\frac{\vartheta}{3})$. By the triple angle formula, which we assume without proof (see Dummit and Foote),

$$\cos(\theta) = 4\cos\left(\frac{\theta}{3}\right) - 3\cos\left(\frac{\theta}{3}\right).$$

If we consider the specific example $\theta = 60^\circ$, we see that the above equation is equivalent to $4\cos(20^\circ) - 3\cos(20^\circ) - \frac{1}{2} = 0$. If we let $\cos(20^\circ) = \alpha$, we have $4\alpha^3 - 3\alpha - \frac{1}{2} = 0$. Multiplying this equation through by 2 yields $8\alpha - 6\alpha - 1 = 0$. Letting $\beta = 2\alpha$ and substituting gives $\beta^3 - 3\beta - 1 = 0$. Let $\beta^3 - 3\beta - 1 = p(x)$. To trisect an angle of 60° , we must find a root to $p(x)$. We know that $p(x)$ has no roots in \mathbb{Q} by Dummit and Foote 9.4.11. So by Dummit and Foote 9.4.10 we know that $p(x)$ is irreducible. Therefore, $[\mathbb{Q}(\alpha):\mathbb{Q}] = 3$, which is not a power of 2. We conclude that $p(x)$ has no constructible roots, and so it is not possible to trisect an angle of 60° , proving that it is not possible in general to trisect an angle.

Dummit and Foote remark that in order to construct a trigonometric function with integer argument (in degrees), the argument of the function must be a multiple of 3, i.e., $\cos(\theta)$ and $\sin(\theta)$ are constructible if and only if $\theta = 3n$ for $n \geq 1$, $n \in \mathbb{Z}$. Further, note that it is also possible to bisect any constructible angle. If we construct an angle ϑ , we know that $\cos(\theta)$ and $\sin(\theta)$ are constructible and we know that ϑ must be a multiple of three. If we divide our angle ϑ by 2, the question becomes, “Are $\cos(\frac{\vartheta}{2})$ and $\sin(\frac{\vartheta}{2})$ constructible?” If we notice that for any constructible θ , $\theta = 3n$ for some $n \geq 1 \in \mathbb{Z}$, then

dividing by 2 gives us $\frac{\theta}{2} = 3 \left(\frac{n}{2}\right)$. Substituting $\frac{n}{2} = m$ gives us $\frac{\theta}{2} = 3m$. We conclude that $\frac{\theta}{2}$ is constructible and so θ can be bisected in general. Note that $\frac{\theta}{2}$ does not necessarily have to be an integer. For example, an angle of 3° can be bisected giving two angles of 1.5° .

Our final point about the three classical problems is that it is possible to solve two of the problems (doubling of the cube and trisecting an angle) using a ruler and compass, the distinction being that a ruler has measured marks while a straightedge does not (Dummit and Foote).

Below, Figure 3 shows a construction created by Archimedes using a ruler and compass in which the angle DAO is precisely one third of the angle COB . A proof showing that DAO trisects COB is beyond the scope of this paper and is left to the reader. It is also possible to double the cube with ruler and compass, which as discussed above, means that we can construct $\sqrt[3]{2}$ with a ruler and compass.

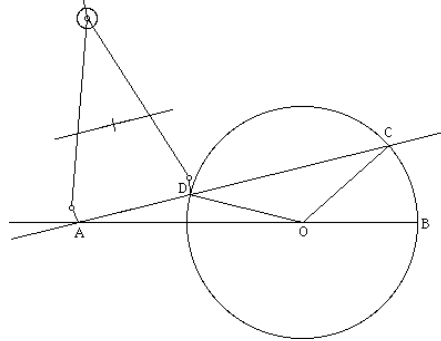


Figure 3. Archimedes' construction trisecting the angle ϑ (Wilson)

Constructability of Polygons

The ancient Greek geometers were also interested in discovering which regular polygons could be constructed using a straightedge and compass (Weisstein, “Constructible”). The Greeks knew how to construct many of the well-known polygons, including triangles, squares, and hexagons, but it was not until 1796 before a young Carl Friedrich Gauss showed definitively which regular polygons are constructible and which are not. The result provided proof that for a regular n -gon, if n is the product of a power of 2 and any number of distinct Fermat primes, then the n -gon is constructible (Weisstein, “Constructible”). Gauss also stated that the converse of the above is true, but did not have a proof for this claim (Weisstein, “Constructible”).

To begin our discussion on constructible n -gons, we consider the n th roots of unity. As we will show, the construction of a regular n -gon is essentially a question of constructing the n th roots of unity, which are all solutions to the equation, $x^n - 1 = 0$.

We denote the first n th root of unity ζ_n . Note also that by Euler’s identity, the n th roots of unity are given by

$$e^{\frac{2\pi ik}{n}} = \cos\left(\frac{2\pi k}{n}\right) + i\sin\left(\frac{2\pi k}{n}\right), 0 \leq k \leq n.$$

To construct a regular n -gon, we must be able to construct $\cos\left(\frac{2\pi k}{n}\right)$ and $\sin\left(\frac{2\pi k}{n}\right)$. We will now discuss an important theorem relating to the constructibility of polygons.

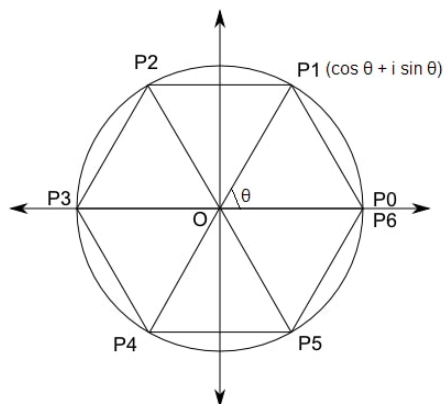


Figure 4. The sixth roots of unity. Note that the n th roots of unity always form a regular n -gon in the complex plane (Singh).

Theorem 4.1. Consider $x^p - 1 = (x - 1)(x^{p-1} + x^{p-2} + x^{p-3} + \dots + x + 1)$. If p is prime, then $x^{p-1} + x^{p-2} + x^{p-3} + \dots + x + 1$ is irreducible.

Proof. Let $x = u + 1$ for some u . Substituting gives

$$(u + 1)^{p-1} + (u + 1)^{p-2} + \dots + (u + 1) + 1,$$

which can be rewritten as

$$\frac{(u + 1)^p - 1}{u} = \frac{\binom{p}{0} + \binom{p}{1}u + \binom{p}{2}u^2 + \dots + \binom{p}{p}u^p - 1}{u}$$

using the binomial formula. Because $\binom{p}{0} = 1$, and noticing that the u in the denominator will cancel with each u in the numerator, we are left with

$$\binom{p}{1} + \binom{p}{2}u + \binom{p}{3}u^2 + \dots + \binom{p}{p}u^{p-1}.$$

This polynomial is irreducible by the Eisenstein Criterion, which says that a polynomial is irreducible if for some prime p , p divides all coefficients except the leading one, p does not divide the leading coefficient, and p^2 does not divide the constant term (Dummit and Foote). We see that the polynomial above is monic because $\binom{p}{p} = 1$. Further, p divides $\binom{p}{i}$ for $1 \leq i \leq (p-2)$, p does not divide the leading coefficient $\binom{p}{p}$ because p is strictly greater than 1, and p^2 does not divide $\binom{p}{1}$ because this is equal to p and clearly p^2 cannot divide p . We conclude that $x^{p-1} + x^{p-2} + x^{p-3} + \dots + x + 1$ is irreducible (Harris 24 Oct). \square

The polynomial we proved irreducible is in fact the irreducible polynomial for the first p th root of unity, denoted ζ_p . We may notice by Theorem 2.2,

$$[\mathbb{Q}(\zeta_p) : \mathbb{Q}] = p - 1.$$

Using this fact and Theorem 2.6, we know that if ζ_p is constructible, then $p - 1 = 2^k$ for $k \geq 1 \in \mathbb{Z}$, which is equivalent to $p = 2^k + 1$.

Proposition 4.2. *If $2^k + 1$ is prime, then $k = 2^m$ for some m .*

Proof. Assume $2^k + 1$ is prime and $k = qr$ where $r \geq 3$ and r is odd. Then

$$2^{qr} + 1 = (2^q + 1)(2^{q(r-1)} - 2^{q(r-2)} + 2^{q(r-3)} - \dots - 2^q + 1).$$

From this equation we see that $2^q + 1$ divides $2^{qr} + 1$. But we assumed that $2^k + 1$ was prime, so the only prime

divisor of k is 2. Thus, $k = 2^m$ (Cooper, Harris 24 Oct).

□

From the proof above, we notice an important corollary: If the p th roots of unity are constructible, then $p = 2^{2^k} + 1$. Primes of this form are known as *Fermat primes* after Pierre de Fermat, who proposed in 1650 that all of the numbers of this form are in fact prime numbers. However, the only Fermat primes that have actually been shown to be prime are the first five, which are 3, 5, 17, 257, and 65,537 (Weisstein, "Fermat"). If a regular polygon has n sides for n odd, then it is constructible if and only if n is the product of a power of 2 and any number of distinct Fermat primes (Dummit and Foote). By this result we see that it is possible to construct a regular 15-gon ($15 = 3 \cdot 5$, which are both Fermat primes), but we cannot construct a regular heptagon. Although 7 is a prime, it is not a Fermat prime. We can generalize which polygons are not constructible by the theorem below.

Theorem 4.3. *For a regular n -gon, if $p^2 \mid n$ such that p is an odd prime, then the n -gon is not constructible.*

From this result, it suffices to show that the p^2 -gon is not constructible because if the n -gon were constructible, we may simply connect the vertices differently and construct the desired p^2 -gon (Harris 24 Oct).

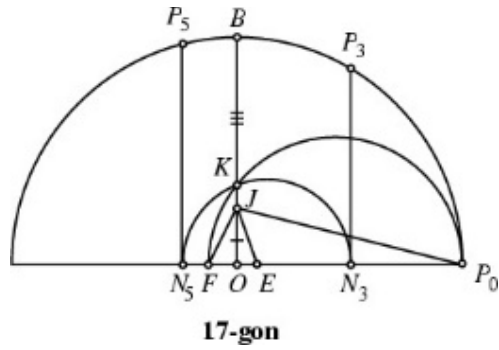


Figure 5. Gauss's construction of a regular heptadecagon (Weisstein, "Heptadecgon")

The constructability of numbers is an ancient topic of mathematics that is still of interest today. Many modern amateur and professional mathematicians find enjoyment in learning the compass and straightedge constructions discovered by the ancient Greeks. Many mathematicians, however, may find the theory behind why the compass and ruler constructions work more interesting. Field theory, and abstract algebra as a whole, have numerous applications in many disciplines, including chemistry, computer science, cryptography, and physics. Field theory may seem inaccessible to the mathematically uninitiated, but it comes in handy in many situations and plays an important part in our modern technological world. On the other hand, the specific application of constructing numbers and polygons with a compass and straightedge is simply an interesting and accessible way to explore the concepts of abstract algebra. Overall, the mental challenge of understanding how the math of constructability works can help understand problems in fields far outside of pure mathematics.

References

- “Constructible Numbers.” Math 3040. University of Utah, Salt Lake City, Utah. Web. 14 Oct. 2016.
- Cooper, Christopher. “Mersenne and Fermat Primes.” *Mathematics Notes*. Macquarie University, 2013. Web. 25 Oct. 2016.
- Dummit, David, and Richard Foote. *Abstract Algebra*. John Wiley & Sons, 2004. Print.
- “G-C Right triangles inscribed in circles I.” *Illustrative Mathematics*. N.p. Web. 16 Oct. 2016.
- Harris, Seth. Drew University, Madison, NJ. 14 Oct. 2016. Lecture.
- Harris, Seth. Drew University, Madison, NJ. 19 Oct. 2016. Lecture.
- Harris, Seth. Drew University, Madison, NJ. 21 Oct. 2016. Lecture.
- Harris, Seth. Drew University, Madison, NJ. 24 Oct. 2016. Lecture.
- Morris, Stephanie J. “The Pythagorean Theorem.” *The Pythagorean Theorem*. J. Wilson, n.d. Web. 15 Oct. 2016.
- O’Connor, J. J., and E. F. Robertson. “Archimedes of Syracuse.” *MacTutor History of Mathematics*. N.p., Jan. 1999. Web. 15 Oct. 2016.

- Ramanujan, Srinivasa. "Squaring the Circle". *Journal of the Indian Mathematical Society*, vol. 5, 1913, pp. 132. Web. 23 Oct. 2016.
- "SAT Math- Similar triangles." *Barron Test Prep*. N.p. Web. 16 Oct. 2016.
- Singh, Paramanand. "Gauss and Regular Polygons: Complex Numbers." Paramanand's Math Notes. N.p., 24 Dec. 2009. Web. 04 Nov. 2016.
- Weisstein, Eric W. "Constructible Polygon." *Mathworld*. Wolfram, 19 Oct. 2016. Web. 24 Oct. 2016.
- Weisstein, Eric W. "Fermat Prime." *Mathworld*. Wolfram, 19 Oct. 2016. Web. 24 Oct. 2016.
- Wesstein, Eric W. "Heptadecagon." Michigan State University. Web. 24 Oct 2016.
- Weisstein, Eric W. "Thales' Theorem." *MathWorld*. Wolfram, 19 Oct. 2016. Web. 16 Oct. 2016.
- Weisstein, Eric W. "Transcental Number." *Mathworld*. Wolfram, 19 Oct. 2016. Web. 23 Oct. 2016.
- Wilson, Steve. "Archimedes' Trisection of the Angle." *Math 150 Projects*. Sonoma University, California. N.d. Web. 25 Oct. 2016.

Acknowledgments

The Editors of *The Drew Review* would like to thank especially the Board of Trustees, President MaryAnn Baenninger, Dean Christopher Taylor, and our faculty advisors for their ongoing and generous support of the Review. While only one-fifth of all nominated papers are published, the Review reflects the efforts of the entire Drew University community. Without the participation of the fifty-one students who submitted their essays and the seventy faculty members who participated as nominators or outside readers, The Drew Review would not be able to maintain its high standards for publication. All nominated essays were recognized by their nominating professors as particularly outstanding examples of undergraduate research and writing; a nomination to the Review honors that achievement. We are grateful to those who made the Review possible by serving in all of these capacities.