



Samford University

Samford Undergraduate
Research Journal

2016 - 2017

Table of Contents

| | |
|--|-----|
| Classical Birmingham: <i>Centurion</i> and the <i>Dying Gaul</i> <i>Emily London</i> | 4 |
| The Poor in Cicero’s <i>Pro Murena</i> <i>Stone Hendrickson</i> | 14 |
| After the Fact: Livy, Galileo, and a Historical Understanding of Truth <i>Madeline Perkins</i> | 26 |
| Natural Slaves as a Product of the Civil State: Analysis of the Natural State and Human <i>Telos</i> in the Political Writings of Aristotle and Rousseau <i>Casey Cunningham</i> | 35 |
| Freedom and Vengeance: How Late Period Egypt and the Late Classical Greek City-States Mutually Contributed to the Decline and Fall of the Persian Empire <i>Stephen Boutwell</i> | 44 |
| Predicting Conflict in Sino-British Relations in the Time of the Opium Wars <i>Lillie Mermoud</i> | 67 |
| A Comparison of Certain Beliefs and Attitudes of Men Who Pay for Sex Vs. Men Who Do Not <i>Meghan Shuffett</i> | 77 |
| Depression and Anxiety: The Influence of Stress, Coping, and General Health <i>Emma Paine, Judy Latta, Kristen Fowler, Elizabeth Batson, and Sandra Willis Ph.D</i> ...90 | |
| NASCAR Fans’ Recognition of Driver Sponsorship <i>Olivia Williams</i> | 111 |
| The Effects of Educational Attainment on Enrollment in Government Assistance Programs <i>Alex Matthews</i> | 138 |

Fiber Optic Shape-sensing for Woven Materials
Jack Kawell, Drew Hackney, Guodong Guo, Luca Girolamo, Kara Peters, Mark Pankow.....159

Analysis of Riverbank Restoration Initiatives in a Riparian Ecosystem
Jacqueline Taylor.....171

Classical Birmingham: *Centurion* and the *Dying Gaul*

Emily London

Abstract

From towering monuments like the iconic Vulcan statue of the Birmingham skyline, to the Wedgwood collection's Portland Vase copies in the Birmingham Museum of Art, glimpses of the classical world can be found all around the Birmingham area. One such manifestation of the classical influence in the city is *Centurion: Jus Fides Libertatum* a statue erected to honor Jefferson County law enforcement officers killed in the line of duty. This paper explores both the classical and modern contexts of *Centurion* as well as the implications that spring forth from the reception of the piece and its inspiration.

The clear inspiration for the physical form of *Centurion* is the *Dying Gaul* of the Capitoline Museum in Rome. The use of the Gaul's image bears significant meaning, as the ancient sculpture portrays Pergamon's and Rome's understanding of Gauls as their savage enemies—their “barbaric other.” *Centurion*, however, also possesses obvious inspiration from the inherently noble spirit of police officers who lose their life in the line of duty. That such a police memorial resides in Birmingham, Alabama, where public memory of police forces is not so favorable when reflecting on the events during the Civil Rights Movement that pushed the city in the nation's attention complicates audience response and interpretation of this piece.

Drawing from classical reception scholarship, I argue that in light of the ancient and modern contexts and the public response to *Centurion*, the artist's intended portrayal of nobility is subverted to convey a synthesis of dignity and barbarism.

Introduction: Interpretation in Light of Reception Theory

The ability of art to communicate meaning across time is an inexplicable and wonderful phenomenon; this is one of many reasons why the subject so intrigues a wide spectrum of audiences, from seasoned scholars to schoolchildren. Stranger and greater still is the notion that audience members, young or old, perceptive or oblivious, educated or novice, participate with artists in understanding and interpreting pieces. This process can span centuries, as in the case of the famous *Dying Gaul* of antiquity and *Centurion*, a law enforcement memorial located in Birmingham, Alabama. In *Centurion*, we see a poignant example of classical reception, an area of scholarship seeking to discover what modern reinterpretation of classical works reveals about their ancient counterparts. As a sculpture heavily inspired by the image of *Dying Gaul*, *Centurion* speaks a great deal about both its own modern context and its Hellenistic and Roman roots.

As a reimagining of *Dying Gaul*, *Centurion* bears a rich background in antiquity. Its classical inspiration heavily influences the meaning it conveys. The study of a rich classical tradition, as reception theorist Shane Butler describes, “is not just what happens to the past after the past, but an extension of the question of why the past... continues to compel our attention.”¹ And not only do those works compel human attention, but that attention and reaction interacts with art to form new meaning.² An ancient Gaul representing a police officer, then, discloses something about both the ancient and the modern figures and their audiences who find the image so compelling. As antiquity invades and influences modernity, as it does in *Centurion*, renewed understanding of both past and present becomes available.³ The ancient and modern context of barbarism for *Dying Gaul* and *Centurion* informs and subverts the artist’s intended nobility in the figure. Likewise, *Centurion* reflects the sympathy downplayed but nevertheless present in *Dying Gaul*.

A Background in Antiquity

The original bronze sculpture, of which now only Roman copies remain, is thought by art historian Arnold Schober to belong to a monument erected by Attalos I in Pergamon to commemorate his victory over the Gauls⁴ in the third century BCE.⁵ Roman historian Pliny the Elder refers to the Hellenistic sculptor of the Attalid victories against the Gauls, Epigonos,

¹ Shane Butler, introduction of *Deep Classics: Rethinking Classical Reception* (London: Bloomsbury Publishing, 2016), 15.

² Hans Robert Jauss, “History of Art and Pragmatic History,” in *Toward an Aesthetic of Reception*, trans. Timothy Bahti (Minneapolis: University of Minnesota Press, 1982), 46-75. Jauss postulates, “the history of art is distinguished from other spheres of historical reality by the fact that in it the formation of the immortal is not only visibly carried out through the production of works, but also through reception, by its constant reenactment of the enduring features of works that long since have been committed to the past.”

³ Miriam Leonard, “Reception,” in *The Oxford Handbook of Hellenic Studies*, ed. George Boys-Stones et al. (Oxford: Oxford University Press, 2009), 835-845.

⁴ Though the Greek transliteration would be “Galatian,” “Gaul” is used throughout this paper for consistency as well as clarity that, despite the complexities of the exact people group(s) ancient historians referred to, the overall cultural prejudice represented in “Dying Gaul” spans a range of peoples who were generalized, both then and now, by the terms used for them.

⁵ Christine M. Havelock, “Victory Monument of Attalos I” in *Hellenistic Art*, by Christine M. Havelock. (Greenwich: New York Graphic Society, 1970), 145-147.

creating a statue of a dying trumpeter.⁶ Epigonos' name is inscribed on two of the bases of the Attalid victory monuments,⁷ so if this statue he references is the original to the Capitoline Gaul, we can realistically place the original sculpture in the third century.⁸ In the same vein of Medenica's use of this image, the Romans immortalized the now-lost sculpture in their own copies, using them to depict their own defeat of their "barbarian" opponents.⁹

As one of these copies, *Dying Gaul* speaks much about the self-definition of the Greeks and Romans as well as their attitude toward their defeated opponents—the Gauls. Art historian J.J. Pollitt proposes that "the force that gave Pergamon its first important push toward greatness in the Hellenistic world was. . . not Greek culture but rather a kind of barbarism."¹⁰ He explains that through these victories, the Attalids began to view themselves "as the patrons and preservers of Greek culture in general and of Athenian culture in particular."¹¹ Ancient Greek historian Polybius describes Attalos' legacy of excellence as founded in "having conquered the Gauls, then the most formidable and warlike nation of Asia Minor" and characterizes such work as "fighting for the liberties of Greece."¹² Attalos gained status as a savior-figure of both Pergamon, and, as is seen in Polybius' comments, all of the Greek cities in the region of Asia Minor.¹³ The clearest demonstration of this mindset is the Pergamene acropolis, with its large monuments to the victories over the Gauls. Scholars (following Schober's example) almost universally agree that these monuments likely included the original Gaul statue that inspired the Roman copy, Capitoline Gaul.¹⁴ Through its form, the Gauls are depicted so as to "provoke in its viewers an imaginary experience, to inspire them to live through, in their imaginations, the events which the monument commemorated" in order to instill greater awe in the audience for the glory of Pergamon.¹⁵

Dying Gaul's physical attributes instill that awe through establishing the subject as "other" to the conventional view of "self" in the audience. The torque around the figure's neck is

⁶ J. J. Pollitt, "The sculpture of Pergamon" in *Art in the Hellenistic Age*, by J.J. Pollitt. (Cambridge: Cambridge University Press, 1990), 79-97.

⁷ Ibid.

⁸ Pliny the Elder, *Natural History*, trans. H. Rackham (Cambridge, MA: Harvard University Press, 1961), 34.84-88.

⁹ Paul Zanker, "The Barbarian and the Amphitheater" in *Roman Art*, trans. Henry Heitmann-Gordon. (Los Angeles: Getty Publications, 2010), 84-87.

Zanker emphasizes that in imperial art, "the barbarian embodies the threatening and uncivilized Other that has to be not only overcome but utterly eradicated." Though this depiction of the Gauls as utterly barbarian is wholly an exaggerated and overly simplified picture of a complex people, he also reminds us that "people within the Empire need a clear-cut, stereotyped enemy," so even though it was not accurate, it served its purpose powerfully through exaggeration.

¹⁰ Pollitt, *Art in the Hellenistic Age*, 79.

¹¹ Ibid., 81.

¹² Polybius, *The Histories*, trans. W. R. Paton (Cambridge, MA: Harvard University Press, 1960), 18.41.

¹³ Brigitte Kahl, *Galatians Re-Imagined* (Minneapolis: Fortress Press, 2010), 62-63.

¹⁴ Pollitt, *Art in the Hellenistic Age*, 89. Which base the sculpture corresponds to, however, is more highly contested. Pollitt discusses in greater detail Schober's thesis on this, as well as possible problems with his reconstruction.

¹⁵ Ibid., 92.

characteristic of Hellenistic depiction of Gauls, the hallmark of the ethnicity in ancient art and literature.¹⁶ The thick and matted hair of the Gaul is consistent with ancient Greek historian Diodorus Siculus' well-known description as well.¹⁷ Interestingly, the most striking element of the sculpture is the great emotion and agony ascribed to the Gaul. Pollitt describes it first as “a kind of animal ferocity [which] has been infused with an unusual sort of dignity. . . of the fanatical and fearless opponent whom one both fears and respects” and later that “empathy for his physical anguish as his strength ebbs away is stimulated not only by the lines of pain in his face by also by a taut and tortured feeling in the musculature.”¹⁸ This figure in crisis, though an enemy, evokes respect and unexpected sympathy that continue to strike audiences as the foremost features.

In addition to its physical attributes reinforcing awe for Greek power, the sculpture drew upon the already widespread hatred toward the Gauls, as seen in other Hellenistic art. Classicist Brigitte Kahl describes this cultural prejudice in great detail in *Galatians Re-Imagined*.¹⁹ She notes “in one way or another [the depictions all] deal with the Gauls in the context of war and victory” with the Gauls as the defeated foe.²⁰ *Dying Gaul* is a foremost example of what classicist John Marszal calls “victory dedications without victors,” but much debate exists as to whether such a depiction serves to be “a treatment of the defeated that is strangely sympathetic,” or if modern audiences impose sympathy.²¹ With or without that pity, a paradox of respect in the midst of distaste for the Gauls is nevertheless conveyed in the image. Art historian H.W. Janson interprets the basic sentiment as “they knew how to die, barbarians though they were,” reflecting that dual nature that so characterized the Hellenistic view of the Gauls.²²

This art drew from a widespread attitude toward the Gauls that is also poignantly represented in ancient historical accounts. In his *Bibliotheca Historica* (“Historical Library”), ancient Greek historian Diodorus of Sicily constructs the Gauls as barbarians because they are decidedly “other” to conventions familiar to him. To indicate how strange he finds them, he compares them to animals: “that their appearance is like that of Satyrs and Pans, since the treatment of their hair makes it so heavy and coarse that it differs in no respect from the mane of horses.”²³ Diodorus allows that some of the Gauls' customs concerning battle seem agreeable to him. He describes them as “a barbarous sort of greatness of soul” because they refuse to sell the

¹⁶ H.W. Janson, “Hellenistic Sculpture” in *History of Art* by H.W. Janson (New York: Harry N. Abrams, Inc., 1991), 196.

¹⁷ Pollitt, *Art in the Hellenistic Age*, 85.

¹⁸ *Ibid.*, 86.

¹⁹ Brigitte Kahl, *Galatians Re-Imagined*. Though Kahl is primarily concerned with how Hellenistic and Roman representations of Gauls inform reading of the apostle Paul's letter to the Galatians, her book contains an excellent overview of the nuances of ancient thought about the Gauls, including a concentrated effort to eliminate ambiguity in which of the many different overlapping groups the use of the term “Gaul” or “Galatian” could refer to.

²⁰ *Ibid.*, 31.

²¹ John R. Marszal, “Ubiquitous Barbarians: Representation of the Gauls at Pergamon and Elsewhere” in *From Pergamon to Sperlonga: Sculpture and Context*, ed. Nancy T. de Grummond and Brunilde S. Ridgway, (Berkeley: University of California Press, 2000), 191-234.

²² Janson, *History of Art*, 196.

²³ Diodorus Siculus, *Bibliotheca Historica*, trans. C.H. Oldfather (Cambridge, MA: Harvard University Press, 1961), 5.28.

heads of their greatest foes, but he also remarks “but to continue to fight against one of our own race, after he is dead is to descend to the level of beasts.”²⁴ Kahl notes that in most of these criticisms, their barbarism is characterized most succinctly as that which is distinctly “un-Greek.”²⁵ Classicist Edith Hall similarly observes that “the notion of the barbarian [is] the universal anti-Greek against whom Hellenic—especially Athenian—culture was defined.”²⁶ This, then, is the Hellenistic attitude toward the subject of *Dying Gaul*, which must inform interpretation of the piece.

A significant amount of Roman imperial art and thought also contributed to the sentiments reflected in *Dying Gaul*, as a piece preserved only as a Roman copy. Kahl observes that similar to their role in Greek self-identity, the Gauls “played a constitutive role in a colonial/imperial discourse centered on war and victory, order and counterorder, Self and Other in the battle of civilization versus barbarism.”²⁷ Livy consistently characterizes the Gauls as greedy and lawless. For example, he describes the Gauls’ departure after their conquest of Rome in 387 BCE as “shamelessly cheating the Romans by using false measures in weighing the thousand-pound ransom.”²⁸ And later, during one of the historically greatest threats to Rome’s existence in the conquests of Hannibal of Carthage, Livy records that Romans offered a Gaul man and woman as human sacrifice.²⁹ This instance demonstrates, to Kahl, that “the Galatians had grown in the Roman imagination into an enemy of mythological stature who represented the primeval threat to Rome.”³⁰ Later, Caesar would lay the very foundation for his rule and the Roman Empire as he waged war in Gaul, gained enough power, and notoriously crossed the Rubicon. It is with this consideration of Hellenistic and Roman attitudes that we now turn to the modern context of *Centurion*.

The Modern Context of *Centurion*

In 1991, The Fraternal Order of Police (FOP) Birmingham Lodge No. 1 commissioned Branko Medenica to create a statue to memorialize law enforcement officers slain in the line of duty. It was erected in the courtyard in front of the Mel Bailey Criminal Justice Center, located in downtown Birmingham.³¹ An annual memorial service is held to commemorate the honored officers. New names are added to the base, where a list of those honored is engraved.³² Medenica chose to use *Dying Gaul* as inspiration in his vision for *Centurion* once the FOP commissioned him to create the memorial because he found its presentation of a fallen warrior striking.³³

Just as Pergamon informs interpretation of *Dying Gaul*, the city of Birmingham, Alabama, also provides context for *Centurion*. Most pertinent, perhaps, is the city’s complicated

²⁴ Diodorus Siculus, *Bibliotheca Historica*, 5.29.

²⁵ Kahl, *Galatians Re-Imagined*, 78.

²⁶ Edith Hall, *Inventing the Barbarian: Greek Self-Definition through Tragedy* (New York: Oxford University Press, 1989), 5. Hall’s thesis is that Athenian tragedy greatly contributed to this collective self-definition through its portrayals of barbarism.

²⁷ Kahl, *Galatians Re-Imagined*, 51.

²⁸ Livy, qtd in *Galatians Re-Imagined*, 54.

²⁹ Cassius Dio, qtd in *Galatians Re-Imagined*, 55.

³⁰ Kahl, *Galatians Re-Imagined*, 55.

³¹ Medenica, interview by author. Oral history, digital recording. Birmingham, AL, 21 July 2016.

³² Ibid.

³³ Ibid.

history with the subject of the memorial: police officers. Richard Arrington, the first African-American mayor of Birmingham notes in his memoir that “Birmingham is internationally known for its oppressive police department, led by its infamous racist police commissioner, Eugene ‘Bull’ Connor, in the 1960s.”³⁴ He also cites the 1977 report from the National Police Foundation *Police Use of Deadly Force* which called the city “‘a most noteworthy exception’ to cities its size for the use of deadly force by its officers.”³⁵ As communications scholar Meg Spratt observes, influential photographs from Birmingham of officers loosing police dogs and turning fire hoses on civil rights protestors in the city circulated in major news sources like *The New York Times* and *Time* magazine. These pictures created a public mythology surrounding the Civil Rights Movement’s relationship to the police as “good versus bad,” or “us versus them,” which could also be described as “self versus Other.” Such images shaped a public view that demonstrates patterns of thought similar to the cultural attitudes that influenced *Dying Gaul*. By the time of the commissioning of *Centurion*, the police department’s reputation had transitioned from this notorious abuse into more just and lawful conduct.³⁶ In many viewers’ minds, however, this image of a police officer in the middle of downtown Birmingham is bound to recall the national attention the city received for its police officers. An interesting comparison to note is the correlation between instances of police brutality in Birmingham with racial prejudice; Medenica’s inspiration unconsciously connects the notorious racial issues in the city with those in antiquity both in taking inspiration from *Dying Gaul* and in the inherent reference to the city’s history of policing.

Centurion also fits in the culture of the city because of its consistency with the artistic identity of Birmingham. The iconic *Vulcan* statue of the skyline over Birmingham is an essential figure in the public image of the city. This statue was created for the 1904 Louisiana Purchase Centennial Exposition and was “intended to symbolize the spirit of [the] city” because of Birmingham’s significant success in the industry of iron production.³⁷ Public historian Matthew Kierstead notes that *Vulcan* is an example of allegorical sculpture of the nineteenth century even as such public art was about to transition in the twentieth century from symbolic representation to modern, literal imagery.³⁸ Even so, *Vulcan* successfully captured the hearts of not only those onlookers at the exposition, but also the citizens of Birmingham, for it eventually found a permanent home atop Red Mountain overlooking the city.³⁹ *Vulcan* is the foremost example of how classically-inspired art has infiltrated the identity of Birmingham as it remains an emblem of civic pride, heritage, and strength. *Centurion*, too, can be seen to reflect the identity of the city in a similar manner: both use allegory, but the former reflects a point of pride, whilst *Centurion* draws attention to a less ideal reason for Birmingham’s fame.

Reception in Birmingham

³⁴ Richard Arrington, *There’s Hope for the World: The Memoir of Birmingham, Alabama’s First African-American Mayor* (Tuscaloosa: The University of Alabama Press, 2008), 65.

³⁵ Ibid.

³⁶ Ibid., 80.

³⁷ Matthew A. Kierstead, “Vulcan: Birmingham’s Industrial Colossus” in *The Journal of the Society for Industrial Archaeology* 28 (2002): 59-74.

³⁸ Ibid.

³⁹ Ibid.

Public reception of the piece demonstrates the influence *Centurion* has had and how modern audiences have interpreted it. In his oral history, Medenica recalls several examples of how the public reacted to his piece. He describes,

There was a deputy-sheriff in North Alabama, or north of Birmingham, somewhere up there, who had gotten into some trouble. . . and they were going to indict him and all that, and so anyways, he came down here one night I think and got in front of the statue there, and killed himself.⁴⁰

That an officer would commit suicide right in front this statue rather than being indicted speaks to the power of the image of the fallen officer in *Centurion*.⁴¹ No one can ascertain that officer's reasoning, but a possible connection is that the officer saw himself as a different sort of fallen warrior, one "fallen from grace," despite the nobility of a memorial. His choice to make that sculpture the location of his death adds another layer to how audiences perceive *Centurion*.

Medenica recalls another instance that sheds more light on reception of the sculpture, one that illustrates the modern sympathy toward *Dying Gaul*. In this case, city officials had *Centurion* moved from its original place in the center of the plaza in front of the Mel Bailey Criminal Justice Center, to the side.⁴² This was done because judges deemed it "too influential on jurors" and felt that jurors entering the courtrooms in the building would become inclined to be more sympathetic toward police officers. Though it is highly debated whether the ancient piece is meant to inspire sympathy or not, it is evident that in *Centurion* this effect is a reality, intentional or unintentional.

Medenica's Intention in Light of *Centurion's* Context

Medenica's aesthetic changes to *Dying Gaul's* image affect the identity and interpretation of *Centurion*. Despite the background of police brutality in the city, Medenica says that in all his thinking about *Centurion* the individual identity and history of the police in Birmingham did not play a major role in his portrayal of them.⁴³ He chose to widen his scope to generally "portray the police as being strong figures in the community."⁴⁴ He intentionally portrays the subject as more sympathetic even as some of his changes also ennoble the figure in a way he did not expect. In his accounts of creating the sculpture, Medenica emphasizes more pronounced musculature in his piece in order to more strongly convey the figure as a fallen warrior.⁴⁵ This, in his mind, contributes to the sympathy felt upon viewing the memorial, conveying that a hero has fallen.⁴⁶ He also, for primarily aesthetic reasons, decided to rebrand the figure as a Roman centurion rather than a Gaul, choosing to do so mainly so that the figure's lower half could be covered and the sculpture's genitals would not be displayed in the middle of the city.⁴⁷ In doing so, Medenica unintentionally made the sculpture a nobler image, by eliminating its "barbarian" identity and replacing it with the centurion. The goriness of the sculpture is also an intentional

⁴⁰ Medenica, interview by author. Oral history, digital recording. Birmingham, AL, 21 July 2016.

⁴¹ Even more fascinating in connection to this story is the ancient possible companion piece to *Dying Gaul* which shows a Gaul who has just slain his wife and is in the midst of turning his sword on himself.

⁴² Medenica, interview by author. Oral history, digital recording. Birmingham, AL, 21 July 2016.

⁴³ Ibid.

⁴⁴ Ibid.

⁴⁵ Ibid.

⁴⁶ Ibid.

⁴⁷ Ibid.

change he made that contributes to this nobility. While droplets of blood flow from the warrior's wound in the original piece, *Centurion* retains the wound but lacks the dripping.⁴⁸ Medenica also tamed the hair of the figure to appear less matted and greasy.⁴⁹ Finally, for purely aesthetic reasons, Medenica moved the hand of the centurion from his knee to the ground.⁵⁰ However small the aesthetic change, or large the shift in identity, these elements emphasize the nobility and sympathy suggested by the Gaul's appearance in *Dying Gaul*.

In examining these two pieces across time and cultures, each informs the meaning of the other by emphasizing the aspects downplayed in the opposite piece. Through seeing *Centurion* through the lens of *Dying Gaul*, as well as the city of Birmingham, the sculpture reflects the city's police department's complex history with brutality and racism. It simultaneously reminds viewers of the great agony and nobility of sacrifice inherent in the lives of officers slain in the line of duty. In realizing the sympathetic portrayal of the noble warrior in *Centurion*, audiences can look back upon *Dying Gaul*, both acknowledging it as one of the most famous portrayals of the barbaric "Other" in art history, as well as seeing its nobility with fresh eyes.

⁴⁸ Medenica, interview by author. Oral history, digital recording. Birmingham, AL, 21 July 2016.

⁴⁹ Ibid.

⁵⁰ Ibid.

Bibliography

- Arrington, Richard. *There's Hope for the World: The Memoir of Birmingham, Alabama's First African-American Mayor*. Tuscaloosa: The University of Alabama Press, 2008. 65-80.
- Boardman, John. *Greek Art*. New York: Praeger Publishers, Inc., 1973.
- Butler, Shane. Introduction to *Deep Classics: Rethinking Classical Reception* by Shane Butler. London: Bloomsbury Publishing, 2016. 15.
- Campbell, Brian and Lawrence A. Tritle, eds. *The Oxford Handbook of Warfare in the Classical World*. New York: Oxford University Press, 2013.
- Diodorus Siculus. *Bibliotheca Historica*, trans. C.H. Oldfather. Cambridge, MA: Harvard University Press, 1961. 5.28-29.
- Goldhill, Simon. *Love, Sex, and Tragedy: How the Ancient World Shapes Our Lives*. Chicago: University of Chicago Press, 2004.
- Hall, Edith. *Inventing the Barbarian: Greek Self-Definition through Tragedy*. New York: Oxford University Press, 1989. 5.
- Havelock, Christine M. "Victory Monument of Attalos I" in *Hellenistic Art*, by Christine M. Havelock. Greenwich: New York Graphic Society, 1970. 145-147.
- Highet, Gilbert. *The Classical Tradition: Greek and Roman Influences on Western Literature*. Oxford: Oxford University Press, 1967.
- Janson, H.W. "Hellenistic Sculpture" in *History of Art* by H.W. Janson. New York: Harry N. Abrams, Inc., 1991. 196.
- Jauss, Hans Robert. "History of Art and Pragmatic History," in *Toward an Aesthetic of Reception*, trans. Timothy Bahti. Minneapolis: University of Minnesota Press, 1982. 46-75.
- Kahl, Brigitte. *Galatians Re-Imagined*. Minneapolis: Fortress Press, 2010. 62-63.
- Kierstead, Matthew A. "Vulcan: Birmingham's Industrial Colossus" in *The Journal of the Society for Industrial Archaeology* 28 (2002). 59-74.
- Koester, Helmut, ed. *Pergamon: Citadel of the Gods*. Harrisburg: Trinity Press International, 1998.
- Leonard, Miriam. "Reception," in *The Oxford Handbook of Hellenic Studies*, ed. George Boys-Stones et al. Oxford: Oxford University Press, 2009. 835-845.
- Marszal, John R. "Ubiquitous Barbarians: Representation of the Gauls at Pergamon and Elsewhere" in *From Pergamon to Sperlonga: Sculpture and Context*, ed. Nancy T. de Grummond and Brunilde S. Ridgway. Berkeley: University of California Press, 2000. 191-234.
- Medenica, Branko. Interview by author. Oral history, digital recording. Birmingham, AL, 21 July 2016.
- Moscatti, Sabatino et al., eds. *The Celts*. New York: Rizzoli, 1998.
- Mumford, Lewis. "What is a City?" *Architectural Record* 82 (1937): 58-62.
- Pliny the Elder, *Natural History*, trans. H. Rackham. Cambridge, MA: Harvard University Press, 1961. 34.84-88.
- Pollitt, J.J. "The sculpture of Pergamon" in *Art in the Hellenistic Age*, by J.J. Pollitt. Cambridge: Cambridge University Press, 1990. 79-97.
- Polybius, *The Histories*, trans. W. R. Paton. Cambridge, MA: Harvard University Press, 1960. 18.41.
- Ridgway, Brunilde Sismondo. *Hellenistic Sculpture I: The Styles of ca. 331-200 BC*. Madison: University of Wisconsin Press, 1990.

- Schaps, David. M. *Handbook for Classical Research*. New York: Routledge, 2011.
- Stecker, Robert. *Aesthetics and the Philosophy of Art: An Introduction*. Lanham: Rowman & Littlefield Publishers, 2010. eBook.
- Zanker, Paul. "The Barbarian and the Amphitheater" in *Roman Art*, trans. Henry Heitmann-Gordon. Los Angeles: Getty Publications, 2010. 84-87.

The Poor in Cicero's *Pro Murena*

Stone Hendrickson

Abstract

This paper seeks to contribute to our understanding of the poor in antiquity. Specifically, the paper focuses on a legal defense speech, the *Pro Murena*, given by Cicero in 63 BCE. In this speech, Cicero directly and indirectly describes the role that the poorer classes played in Roman consular elections. While the speech provides virtually no clear demographic information regarding the poorer classes, Cicero's references to the poor can nuance our understanding of the poor in Rome when read closely in the context of scholarship on the late Roman republic. In two passages especially, Cicero contrasts the roles of the Roman poor and the Roman elite in consular elections. Ultimately, a close analysis of these passages indicates that Cicero perceives "the poor" as those who participate in a certain type of reciprocal relationship with the elite. Although Cicero does not provide clear distinctions between classes, his perspective does offer valuable, albeit limited, insight into the nature of the poor in the society of the late Roman republic.

On September 17, 2011 protesters took up a two-month long residence in Zuccotti Park in New York City that sparked what became known as the “Occupy movement.” Out of this movement was coined the now familiar phrase “We are the 99%.” According to the Occupy movement, there exists a dangerous economic disparity between the wealthiest 1% of the United States population and the rest of the country. Anger at this inequality fueled support for Bernie Sanders’ unsuccessful 2016 bid for the Democratic presidential nomination.¹ The voices of the 99% have impacted American politics. Politics in the ancient Roman Republic, conversely, is widely considered to have been the prerogative of the elite.² David Laitin suggests that one of the problems regarding income inequality in democratically governed societies (and Rome was technically democratic),³ is understanding why the governments of such societies continue to produce policies that favor the wealthiest few.⁴ While people in lower economic classes presumably desire economic equality with the wealthy, the perpetuation of economic inequality in democracies continues from ancient times to now.

In trying to understand the poor in the ancient Roman Republic, one consistent challenge is that the literary record consists almost entirely of the words of the elite. As a consequence, the perspective of the Roman 99% remains hidden. Scholars are trying to coax some conclusions from the evidence⁵ and this paper will consider how one elite source constructed the poor in his speaking and writing. A close reading of Cicero’s *Pro Murena* can provide insight into the nature and role of the poor in the context of consular elections in the late Roman republic. An examination of the evidence in the speech must necessarily wrestle with the lack of precise demographic information that modern scholarship typically pursues. Cicero offers virtually no

¹ Adam Gabbatt, “Former Occupy Wall Street protesters rally around Bernie Sanders campaign,” *The Guardian*, last modified September 17, 2015, <https://www.theguardian.com/us-news/2015/sep/17/occupy-wall-street-protesters-bernie-sanders>; Gregory Krieg, “Occupy Wall Street rises up for Sanders,” *CNN*, last modified April 13, 2016, <http://www.cnn.com/2016/04/13/politics/occupy-wall-street-bernie-sanders-new-york-primary>.

² Fergus Millar, *The Crowd in Rome in the Late Republic* (Ann Arbor, Michigan: University of Michigan Press, 1998); Alexander Yakobson, *Elections and Electioneering in Rome: A study in the political system of the late republic* (Stuttgart, Germany: Franz Steiner Verlag, 1999); Henrik Mouritsen, *Plebs and Politics in the Late Roman Republic* (Cambridge: Cambridge University Press, 2001); Jeffrey Tatum, “Review: Elections in Rome,” *The Classical Journal* 99, no. 2 (December 2003-January 2004). Millar and, later, Yakobson disagree sharply with the traditional consensus that late Republican politics were dominated by the powerful elite and by concrete patron-client relationships. Both insist that the Roman plebs played a much more significant role than previously believed. Mouritsen and Tatum have offered critiques of Millar and Yakobson’s arguments and have emphasized evidence of the limited role played by the plebs in the politics of Cicero’s era.

³ W. Jeffrey Tatum, “Roman Democracy?” in *A Companion to Greek and Roman Political Thought*, ed. Ryan K. Balot (Oxford: Blackwell Publishing, 2009).

⁴ David Laitin, “Political Remedies to Economic Inequality,” in *Occupy the Future*, ed. David B. Grusky, Doug McAdam, Rob Reich, and Debra Satz (Boston: MIT Press, 2013), 137-154. Laitin examines the issue of why democratic political institutions continue to have policies that serve the rich.

⁵ Margaret Atkins and Robin Osborne, eds., *Poverty in the Roman World* (Cambridge: Cambridge University Press, 2006).

precise references to the income or wealth of those who qualify as poor or, conversely, as elite. Although Cicero provides very little evidence regarding precise demographic measurements for various classes in Rome, his speech does provide insight into his perception of the poor in Rome. In the *Pro Murena*, Cicero perceives “the poor” as those who participate in a certain type of reciprocal relationship with the elite, not necessarily those who possess a certain level of income or wealth.

Cicero delivered his speech in defense of Murena in November of 63 BCE⁶ near the end of his annual term as one of two elected consuls.⁷ As one of the consuls of the year 63, Cicero presided over the elections for the next year's consuls. L. Licinius Murena was one of two candidates elected to be consul for 62. Upon losing the election, S. Sulpicius Rufus, one of the other candidates in 63, charged Murena with the crime of *ambitus* (i.e. electoral bribery). Although Cicero supported Sulpicius in the campaign, he chose to defend Murena in court. As he argues in the speech, Cicero defended Murena in order to ensure that Rome had two consuls in place in 62 to protect the republic from Catiline and his threats.⁸ Cicero begins the speech by arguing why it is appropriate for him to defend Murena.⁹ He continues to briefly defend Murena's character and military record.¹⁰ Cicero quickly transitions to the central part of his speech, the *contentio dignitatis*.¹¹ From Cicero's arguments, it would appear that the prosecution had made the argument that Sulpicius was much more likely to win the consular election. Thus, the only way that Murena could have won his victory was through the use of bribery. Cicero responds in the *contentio dignitatis* by arguing that Murena's career as a general won him much more favor with the electorate than Sulpicius' career as a lawyer. Cicero also discusses Murena's and Sulpicius' past elections to the quaestorship and praetorship. Finally, he concludes the *contentio dignitatis* by arguing that Sulpicius lost support during the campaign by openly preparing his prosecution of Murena (an action equivalent to admitting defeat) and by calling for an additional law against *ambitus*, the *lex Tullia*. Cicero then transitions to the third main section of his speech, the *de ambitus criminibus*, where he finally addresses the charge of *ambitus* directly.¹² In this section, Cicero offers explanations for the crowds who followed Murena around during his campaign. Cicero concludes the speech with a brief but intense challenge to the jury to acquit Murena for the sake of Rome's safety and out of pity for Murena.¹³

Cicero makes references to the lower classes in Rome. While he uses various words and expressions to do so, he consistently refers to lower classes as *tenues*. The word literally means “thin” or “slender.” Cicero regularly applies that meaning to the economic substance (or lack thereof) of a class of people. When Cicero uses the word in this sense, it can be roughly

⁶ All dates are BCE unless otherwise noted.

⁷ A. D. Leeman, “The Technique of Persuasion in Cicero's *Pro Murena*,” in *Eloquence et rhétorique chez Cicéron* (Geneva: Fondation Hardt, 1982), 193-228. Leeman provides one of the best readings of the speech as a whole in English.

⁸ Cicero, *Pro Murena* 4. Unless otherwise noted, all primary sources are cited from the *Oxford Classical Texts*. All translations are my own.

⁹ Cic. *Mur.* 1-10

¹⁰ Cic. *Mur.* 11-14

¹¹ Cic. *Mur.* 15-53

¹² Cic. *Mur.* 54-83

¹³ Cic. *Mur.* 83-90

equivalent to the English phrase “the poor.” Of course, a phrase such as “the poor” tends to suppress socio-economic distinctions and compress people into a single category.

Despite Cicero's lack of precision in his use of the term, we can gather a general sense of its meaning from his corpus. The instance most clearly relating *tenues* to lack of wealth occurs in the *De Inventione*. While discussing different types of men in Roman society, Cicero lists a series of parallel contrasts: “*servus sit an liber, pecuniosus an tenuis, privatus an cum potestate*” (“if he is slave or free, wealthy or poor, a private man or a man with [governmental] power”).¹⁴ The binaries clearly indicate opposed descriptions. “*Pecuniosus*” unambiguously means “wealthy;” therefore, it follows that Cicero uses “*tenuis*” here to mean simply “poor” in terms of wealth or income. Admittedly, a youthful Cicero wrote the *De Inventione*, but the later appearances of the term confirm this usage. Cicero's speech *In Verrem*, given only about seven years before the *Pro Murena*, includes multiple uses of the word *tenues*. Cicero uses the word to describe a small income: “*nos ita vivere in pecunia tenui*” (“that we live in this way in meager wealth”).¹⁵ Interestingly, he also uses the term in connection with another class description: “*Homines tenues, obscuro loco nati*” (“Poor men from a humble birth”).¹⁶ Here, Cicero associates lack of wealth with lack of nobility. While this is not the only way in which Cicero uses the word *tenues*, he does use the word in a consistent manner when describing the poor.

Cicero's use of the word *tenues* in the *Pro Murena* conforms to this pattern of describing the meager wealth of the lower classes. While his use of it in the *Pro Murena* paints a similarly broad picture of the poor, the speech does yield some particles of information that can allow us to further understand Cicero's perception of those designated as *tenues*. Although we may not arrive at an exact understanding of the poor in late republican Rome, we can better understand how Cicero perceives “the poor” in the Rome of his age.

In *Pro Murena* 67, Cicero provides a useful outline for approaching his use of the term *tenues* and his understanding of the lower classes in Rome. In this section, Cicero responds to the primary accusations of Cato (one of the prosecutors) regarding *ambitus*. According to Cicero's paraphrase (although we have no secondary verification of Cicero's accuracy), Cato had asserted that Murena violated the three parts of the *lex Calpurnia*.

Ambitum accusas; non defendo. Me reprehendis, quod idem defendam quod lege punierim. Punivi ambitum, non innocentiam; ambitum vero ipsum vel tecum accusabo, si voles. Dixisti senatus consultum me referente esse factum, si mercede obviam candidatis issent, si conducti sectarentur, si gladiatoribus volgo locus tributim et item prandia si volgo essent data, contra legem Calpurniam factum videri.

You [Cato] charge *ambitus*; I do not defend against it. You censure me, because I will defend the same which I should have punished by law. I punished *ambitus*, not innocence; in truth I will charge *ambitus* itself even with you, if you will desire. You said that the decree of the senate should be made with me bringing it forward, if they should go out to meet a candidate with a bribe, if they should follow as ones led, if they should be given places at the games and even feasts indiscriminately and by tribe, against the Calpurnian law it should seem to be done.

¹⁴ Cic. *De Inventione* 1.35.25

¹⁵ Cic. *In Verrem* 2.3.9

¹⁶ Cic. *Verr.* 2.5.167

Cicero proceeds to respond to Cato's charges as Cicero had formulated them: (1) bribing people to go out to meet a candidate, (2) bribing people to follow a candidate around, and (3) giving seats at games and providing feasts by tribe and at large.

In his response to the first part of Cato's charges, Cicero does not actually use the word *tenues*. However, he does provide an important description of different social groups involved in campaigns. Cicero's argument contributes to our understanding of how he later uses the word *tenues* and other ways in which he describes those who are less wealthy in Roman society. In *Pro Murena* 68, Cicero "quotes" Cato's complaint: "*Multi obviam prodierunt de provincia decedenti*" ("Many went out to the one returning from his province"). Apparently, when Murena returned from his province, a large crowd came out to meet him. Cato seems to have asserted that such a crowd would only have done so if bribed. Cicero responds to this charge by arguing that this welcome by a large crowd was in accord with Roman custom and tradition. Further, in *Pro Murena* 69, Cicero reminds the jury of who went to out to meet Murena.

Quid? si omnes societates venerunt quarum ex numero multi sedent iudices; quid? si multi homines nostri ordinis honestissimi; quid? Si illa officiosissima quae neminem patitur non honeste in urbem introire tota natio candidatorum, si denique ipse accusator noster Postumus obviam cum bene magna caterva sua venit, quid habet ista multitudo admirationis? Omitto clientis, vicinos, tribulis, exercitum totum Luculli qui ad triumphum per eos dies venerat; hoc dico, frequentiam in isto officio gratuitam non modo dignitati nullius umquam sed ne voluntati quidem defuisse.

What then? If all the companies, from the number of which many of the jurors sit, came; what then? If many men from our most honorable order came; what then? If that entire most officious group of candidates which allows no one to enter into the city without honor, if finally that Postumus, our accuser, came with his enormous band, what is unusual about that large crowd? I omit clients, neighbors, fellow-tribesmen, and the entire army of Lucullus who had come to the triumph during his days; I say this, that in that service a grateful crowd was never lacking, not only on account of a man's worth but even if he wished for a crowd.

The first part of this argument describes several groups, all of whom can be reasonably classified in the upper stratum of Roman society. The *societates* likely refers to the *societates publicanorum* which were companies of tax agents who operated collectively. According to C. MacDonald, the leaders of these companies were *equites*, a class just below the Senators.¹⁷

Cicero includes the men from his own Senatorial order as well as all of the candidates for other, probably minor, offices who would have been eager to seek the support of a major candidate for the consulship.¹⁸ Finally, one of the prosecutors, Postumus, certainly of the Senatorial order, was also there with his associates. Given that these groups include the peers of Cicero and Cato and those who have strong interests in the Senate, Cato cannot name them as the objects of bribery.

Cicero then mentions the presence of several other groups whom he forgoes to describe in detail. These are groups not above suspicion on account of their class or wealth, but because of their obvious and pre-existing connections to the candidate. His clients owe him favors, his neighbors and fellow-tribesmen have a natural interest in supporting one of their own, and

¹⁷ C. MacDonald, notes on Cicero, *Pro Murena* (Bristol, England: Bristol Classical Press, 1982), 126.

¹⁸ *Ibid.*, 127.

Lucullus' army continues to bear great loyalty toward its former commander.¹⁹ This list is helpful for our interests because it names groups of people who appear to be distinct from the wealthier groups described first. And while the tribesmen and neighbors may include individuals of various income levels, the soldiers are certainly among the lower levels of income. The clients, of course, are also beneath Murena and his peers.

This contrast does not necessarily give us precise information about the poor in Rome in general. It does, however, provide important insight into Cicero's perception of the poor. Cicero defends the presence of the elite simply by noting their status. Their wealth and position in society are enough to dispel concerns of bribery. But for anyone lacking wealth and status, Cicero must provide an ulterior explanation for their presence: pre-existing allegiances. This distinction suggests that Cicero largely sees two groups of people in Roman society: those with wealth and connections and those without. Cicero offers various descriptions for the different people in both halves of *Pro Murena* 69, but clearly divides them into two exclusive groups. The first half includes contrasting categories: politicians and businessmen, those seeking election and those already elected, individuals and members of associations. Despite their distinguishing characteristics, they all control significant power and wealth in Rome or are closely connected to those who do. In contrast, Cicero describes the second group in terms that have virtually no direct relationship to power in Rome. Rather, they assert various forms of relationships to Murena that are either ambiguous in regard to wealth (*vicinos* and *tribulis*) or clearly subservient to Murena and those of his class (*clientis* and *exercitum*). By grouping these distinct descriptions into two separate levels, Cicero suggests that the most fundamental classification—wealth—creates only two groups in Roman society: the rich and the poor.

In *Pro Murena* 70, Cicero transitions to the second part of Cato's charges: "*At sectabantur multi*" ("But many were following"). "[*S*]ectabantur" refers to *adsectatio* which was the service people could perform for a candidate by following him around in the Forum while the candidate conducted his public business. Those who did this (*sectatores* or *adsectatores*) provided a valuable service to a candidate by increasing his public prestige.²⁰ Cicero answers Cato's charge by arguing that there is a certain kind of relationship between the upper class and the lower class. The practice of *adsectatio* fulfilled an essential need in Roman social relations by facilitating political interaction between the two groups. Cicero assumes a dichotomy between rich and poor that is bridged by a reciprocally beneficial relationship.

Homines tenues unum habent in nostrum ordinem aut promerendi aut referendi benefici locum, hanc in nostris petitionibus operam atque adsectationem. Neque enim fieri potest neque postulandum est a nobis aut ab equitibus Romanis ut suos necessarios candidatos adsectentur totos dies; a quibus si domus nostra celebratur, si interdum ad forum deducimur, si uno basilicae spatio honestamur, diligenter observari videmur et coli; tenuiorum amicorum et non occupatorum est ista adsiduitas, quorum copia bonis viris et beneficiis deesse non solet.

Men of meager substance have one opportunity of earning and repaying a benefit: this labor and service of constant attendance during our campaigns. For it is not able to be

¹⁹ Cic. *Mur.* 20. In the *contentio dignitatis*, Cicero had been describing Murena's impressive service as a *legatus* to the *imperator* Lucullus.

²⁰ W. Jeffrey Tatum, *The Patrician Tribune* (Chapel Hill, North Carolina: University of North Carolina Press, 1999), 27. Tatum provides a helpful overview of this and other important campaign practices.

done nor must it be asked from us or from the Roman *equites* that they should follow the candidates they support for entire days; if our house is filled by these [Senators or *equites*], if sometimes we are led to the forum, if we are honored with one turn around a basilica, then we seem to be carefully paid attention to and honored; that constant attendance is the role of friendly and unoccupied men of meager substance, plenty of whom are not accustomed to be lacking for good and beneficent men.

Cicero specifically mentions three social classes in Rome: Senators, *equites*, and *tenues*. He groups these three classes according to their roles in political campaigns. The Senators and *equites* comprise all those who run for the offices or may lend weighty support to other candidates of their class. The *tenues* consist of those who have any role in supporting the candidates. He mentions no other group of people who might be involved in making public appearances with a candidate. *Pro Murena* 70 establishes the basic principle that, in some meaningful sense, Cicero can create a binary opposition by describing potentially all of Roman society as either in the extremely wealthy top two classes or in the class of *tenues*. However, Cicero is not necessarily asserting that the two ends of the binary are complete opposites (i.e. a Roman is either extremely wealthy or extremely destitute). Rather, we can infer that in Cicero's mind, there is enough difference between the Senators/*equites* and everyone else that they can be divided into two distinct groups. All who are not Senators/*equites* are similar enough to each other to be sufficiently described with a single term denoting lack of economic substance.

In addition to asserting a binary opposition, Cicero indicates that the performance of *adsectatio* fulfills a necessary reciprocal relationship between the elite and the *tenues*. The *tenues* offer the one useful service they can provide: constant attendance. In return, they gain the satisfaction of having meaningfully contributed to the success of a member of the elite. In *Pro Murena* 71, Cicero further emphasizes this reciprocation and provides another element.

Noli igitur eripere hunc inferiori generi hominum fructum officii, Cato; sine eos qui omnia a nobis sperant habere ipsos quoque aliquid quod nobis tribuere possint. Si nihil erit praeter ipsorum suffragium, tenues, etsi suffragantur, nil valent gratia.

Therefore, do not desire to seize this fruit of a service rendered from the lower class of men, Cato; allow those who hope for everything from us to have those things and also anything which they are able to grant to us. If there will be nothing except their vote, even if they do vote, the *tenues* have no influence.²¹

His description of the "fruit of a service" emphasizes the organic and healthy nature of this exchange. More specifically, it provides a very real benefit to the *tenues* by giving them an opportunity to contribute significantly to the electoral process.

If we assume that Cicero is not distorting reality here (to be rhetorically effective, his assertions must seem plausible to the jurors), then we can glean important information from the

²¹ Joachim Adamietz, notes on Cicero, *Pro Murena* (Darmstadt: Wissenschaftliche Buchgesellschaft, 1996), 219-220. Adamietz notes that *suffragantur* usually refers to "campaigning support" in Cicero's usage, as opposed to the more specific "vote," as I have translated it, but he also notes that both translations are viable. He concludes that it is ultimately unclear which translation is best. The alternative translation ("campaigning support") does not seem consistent with my later argument concerning the *comitia centuriata*.

fact that the votes of the *tenues* lack significant influence.²² The electoral assembly known as the *comitia centuriata* decided consular elections. This assembly consisted of 193 centuries into which all six classes of Roman voters were organized according to their relative wealth. Generally, the wealthiest classes were distributed across many of the centuries while the lower classes were concentrated into far fewer centuries, giving the votes of the wealthiest citizens an extremely disproportionate weight. For example, Henrik Mouritsen notes that Senators, *equites*, and others of the highest class composed 88 centuries while the poorest in Rome, the *proletarii*, made up only one century.²³ Voting proceeded one century at a time until two candidates achieved a majority at which point voting immediately ceased. As a result, the lower centuries consisting of the poorest citizens would often not get the chance to vote at all.²⁴

Research concerning the economic status of the different centuries can help qualify our understanding of the poorer centuries. According to Jeffrey Tatum's analysis of the *comitia centuriata*, even if the upper classes voted in unison and elected two consular candidates as quickly as possible, voting would still have to extend into the second class. However, according to Yakobson, evidence of intensely contested elections suggests that this unanimity did not always exist. The greater the contest, the more centuries would have to be called upon to decide the vote. Yakobson points out the ultimate irony of this system: "Paradoxically, the system of voting according to property-classes meant that even a small number of poor voters representing the lowest class which was called to vote might sometimes decide the fate of an election."²⁵ Thus, while numerically we can know that voting in the *comitia centuriata* extended at least into the second class, it could, and likely often did, involve even lower classes.

As we apply this evidence to the *Pro Murena*, it is important to note that Cicero never refers to distinct Roman classes in his speech. Instead, he refers only to the Senators, *equites*, and the *tenues*. While significant voting power extended into at least the second class, Cicero does not seem to acknowledge any people who are classed below Senators and *equites* but above the *tenues*. Regardless of where members of the second class figure in Cicero's perception, it is unlikely that he includes them with the *tenues*. Instead of expanding our understanding of Cicero's conception of the *tenues*, this evidence only limits the classes that the term could include. If Cicero is accurate when he asserts that the votes of the *tenues* simply do not matter, they must consist of classes low enough to be infrequently called to vote even when fairly typical divisions occurred in the votes of the upper classes.

However, this does not necessarily mean that the *tenues* were destitute. While the relationship between the classes and concrete figures of income is extremely difficult to establish, Walter Scheidel suggests a basic model. Following the work of Dominic Rathbone, Scheidel offers the following levels of income for the different classes: "[...] for the last century of the Republic, thresholds of HS100,000 for the first class, HS75,000 for the second, HS50,000, for the third, HS20,000 for the fourth, and as little as HS375 for the fifth."²⁶ Focusing on the

²² Cicero, *De Republica* 2.39. In his treatise on the Roman constitution, Cicero notes that the voting is organized "ne plurimum valeant plurimi" ("lest the many should be strong with much power").

²³ Mouritsen, *Plebs and Politics*, 94.

²⁴ Tatum, "Review: Elections in Rome," 207. Tatum provides a concise overview of the *comitia centuriata*.

²⁵ Yakobson, *Elections and Electioneering*, 48-9.

²⁶ HS is an abbreviation for sestertius, a basic unit in Roman currency.

threshold for the fourth class and using average costs of grain and median incomes for the period, Scheidel argues, "In other words, the lower census limit for the fourth class appears to have approximated the income threshold for a reasonably secure commoner household." He concludes, "Most members of the fourth class [...] would have been reasonably well cushioned against chronic want."²⁷ Even if the *tenues* extended no higher than the fourth class, they were not necessarily starving.

Finally, Cicero proceeds to the third point of Cato's charges regarding Murena providing seats at gladiatorial games and feasts in *Pro Murena* 72.

At spectacula sunt tributim data et ad prandium volgo vocati. Etsi hoc factum a Murena omnino, iudices, non est ab eius amicis autem more et modo factum est, tamen admonitus re ipsa recordor quantum hae conquestiones in senatu habitae punctorum nobis, Servi, detraxerint. Quod enim tempus fuit aut nostra aut patrum nostrorum memoria quo haec sive ambitio est sive liberalitas non fuerit ut locus et in circo et in foro daretur amicis et tribulibus?

But seats were given by tribes and men were called to a feast indiscriminately. Even if this was not done by Murena at all, judges, but by his friends it was done moderately and according to custom, nevertheless I recall that I admonished you, Servius, in this matter that as much as these complaints were pronounced in the senate they detracted votes from us. For what time was there either in our memory or in the memory of our fathers in which this (whether it is canvassing or generosity) was not so that a place should be given both in the circus and in the forum to friends and fellow-tribesmen?

Per Cicero's usual pattern, the first line appears to be a paraphrase of Cato's objection. His concern pinpoints the problematic aspect of the reciprocation between candidates and their supporters from the lower classes. In this case, Murena was allegedly using bribery to win over entire tribes. Andrew Lintott suggests that bribery was disruptive to the established political system because it involved "subversive raiding by means of *largitiones* of votes which had belonged to others," in this case, the votes of other candidates' tribes.²⁸ Additionally, Jeffrey Tatum explains how *ambitus* disrupts Roman social relations when he argues that bribery allowed a candidate to cheat the traditional necessity of earning *gratia* and *virtus* and simply purchase votes instead.²⁹ Such a shift would free a candidate from the system of favors and expectations that tied Roman society together. Lintott also argues that from the perspective of the voters, large-scale bribery "was not only profitable but liberating, as it created the assumption

²⁷ Walter Scheidel, "Stratification, deprivation and quality of life," in *Poverty in the Roman World*, ed. Margaret Atkins and Robin Osborne (Cambridge: Cambridge University Press, 2006), 49.

²⁸ Andrew Lintott, "Electoral Bribery in the Roman Republic," *The Journal of Roman Studies* 80 (1990): 14.

²⁹ W. Jeffrey Tatum, "Campaign Rhetoric," in *Community and Communication: Oratory and Politics in Republican Rome*, ed. Catherine Steel and Henriette van der Blom (Oxford: Oxford University Press, 2013), 134.

that their votes were on the open market.”³⁰ While *ambitus* may have had a liberating effect from the perspective of some, Cato and others no doubt saw the effect as disruptive and threatening.³¹

Cicero responds to the concerns behind Cato's charges by arguing that Murena's activities upheld Roman societal traditions and expectations. In fact, according to Cicero, this practice was so normal that Sulpicius lost votes because he tried to use it as evidence against Murena. He asserts that providing *spectacula* for races in the circus and games in the Forum was a practice common in living memory. *Pro Murena* 72 returns to the principle of reciprocity between the *tenues* and the elite.

*Haec homines tenuiores praemia commodaque a suis tribulibus vetere instituto adsequebantur*****

Poor men gain these rewards and advantages from their fellow-tribesmen according to established custom****³²

Cicero adds to the nature of the reciprocal relationship between *tenues* and elite that he established in *Pro Murena* 70 and 71. Not only do the *tenues* provide *adsectatio* in return for the opportunity to have influence in the electoral process, but they also receive various “*praemia commodaque*” for their support of a candidate. In *Pro Murena* 73, Cicero cites some specific examples of Murena's supporters giving up their seats in the games to Murena's other potential supporters. His conclusion indicates that the recipients of these seats were *tenues*.

Omnia haec sunt officia necessariorum, commoda tenuiorum, munia candidatorum.

All these things are the services of close supporters, the advantages of the poor, the gifts of candidates.

With this statement, Cicero concludes his point-by-point response to Cato's charges and returns to the principle of reciprocity. The candidate offers gifts to the *tenues* who return the favors by supporting the candidate by their presence and (to a lesser but potentially still important extent) votes. However, “*commoda tenuiorum*” does not necessarily mean that the lowest classes in Rome were actually the beneficiaries of these gifts. As J. P. V. D. Balsdon argues, during the Augustan era, the Circus Maximus could only hold about 150,000 people (about 15% of Rome's population at the time). The total occupancy for the theaters allowed only about 1-3% of Rome's population to be present at that entertainment.³³ During Cicero's time, such venues would only have been more limited in their seating capacity. As such, there would simply not have been enough room for large numbers of Rome's poor to enjoy such games. Especially in the context of *Pro Murena* 72-73, Cicero's use of the word *tenues* may encompass not only the poorest classes in Rome but also those only slightly poorer than the elite.

The *Pro Murena* resists application to specific demographic questions and evidence. While the evidence of Tatum and Yakobson regarding the *comitia centuriata* suggests that Cicero uses *tenues* to refer only to the lowest classes, Balsdon's evidence indicates that the term may not refer to Rome's poorest. Although this tension is significant and frustrating, Cicero does

³⁰ Lintott, “Electoral Bribery,” 14.

³¹ Cato was such a staunch opponent of electoral “corruption” that he declared he would prosecute the winners of the election in 63 regardless of who they were because of the rampant bribery (Plut. *Cat. Min.* 21.2-3; Cic. *Mur.* 62, 64).

³² Although the line ends with a textual *lacuna*, and we have lost the rest of the sentence, the substance of the quoted portion is sufficient for my argument.

³³ J. P. V. D. Balsdon, “Panem et Circenses,” in *Hommages à Marcel Renard*. ed. Jacqueline Bibauw (Brussels: Latomus, 1969), 59-60.

use the term *tenues* consistently. Throughout the speech, Cicero emphasizes the reciprocal relationship between the “the poor” (whoever they may be in precise demographic terms) and the elite. Indeed, this is the essential concern in the charge of *ambitus*. As Andrew Riggsby argues, both *ambitus* and other legitimate forms of campaigning relied upon this principle of exchange. Riggsby concludes that the difference is that “*ambitus* is campaigning which too clearly reveals this reciprocal relationship between the candidate and the voter.”³⁴ Cicero did not need to deny that Murena ever provided benefits to the poor or the rest of the electorate during the campaign. Rather, Cicero strove only to demonstrate that Murena’s reciprocal relationship to the *tenues* conformed to the basic principles and traditions of Roman society. This relationship may provide one reason why economic inequality existed in a technically democratic state. Murena’s campaign exemplified a way in which the Roman elite accommodated economic inequality by creating a mutually interdependent relationship between themselves and the poor. At least, that is how Cicero and his fellow elites perceived their society.

³⁴ Andrew Riggsby, *Crime and Community in Ciceronian Rome* (Austin, Texas: University of Texas Press, 1999), 152.

Bibliography

- Adamietz, Joachim. Notes on Cicero, *Pro Murena*. Darmstadt: Wissenschaftliche Buchgesellschaft, 1996.
- Atkins, Margaret and Robin Osborne, eds. *Poverty in the Roman World*. Cambridge: Cambridge University Press, 2006.
- Balsdon, J. P. V. D. "Panem et Circenses." In *Hommages à Marcel Renard*, edited by Jacqueline Bibauw, 57-60. Brussels: Latomus, 1969.
- Gabbatt, Adam. "Former Occupy Wall Street protesters rally around Bernie Sanders campaign." *The Guardian*. Last modified September 17, 2015. <https://www.theguardian.com/us-news/2015/sep/17/occupy-wall-street-protesters-bernie-sanders>.
- Krieg, Gregory. "Occupy Wall Street rises up for Sanders." *CNN*. Last modified April 13, 2016. <http://www.cnn.com/2016/04/13/politics/occupy-wall-street-bernie-sanders-new-york-primary>.
- Laitin, David. "Political Remedies to Economic Inequality." In *Occupy the Future*, edited by David B. Grusky, Doug McAdam, Rob Reich, and Debra Satz, 137-154. Boston: MIT Press, 2013.
- Leeman, A. D. "The Technique of Persuasion in Cicero's *Pro Murena*." In *Eloquence et rhétorique chez Cicéron*, 193-228. Geneva: Fondation Hardt, 1982.
- Lintott, Andrew. "Electoral Bribery in the Roman Republic." *The Journal of Roman Studies* 80 (1990): 1-16.
- MacDonald, C. Notes on Cicero, *Pro Murena* 50-139. Bristol, England: Bristol Classical Press, 1982.
- Millar, Fergus. *The Crowd in Rome in the Late Republic*. Ann Arbor, Michigan: University of Michigan Press, 1998.
- Mouritsen, Henrik. *Plebs and Politics in the Late Roman Republic*. Cambridge: Cambridge University Press, 2001.
- Riggsby, Andrew. *Crime and Community in Ciceronian Rome*. Austin, Texas: University of Texas Press, 1999.
- Scheidel, Walter. "Stratification, deprivation and quality of life." In *Poverty in the Roman World*, edited by Margaret Atkins and Robin Osborne, 40-59. New York: Cambridge University Press, 2006.
- Tatum, Jeffrey. "Review: Elections in Rome," *The Classical Journal* 99, no. 2 (December 2003-January 2004): 203-216.
- . "Campaign Rhetoric." In *Community and Communication: Oratory and Politics in Republican Rome*, edited by Catherine Steel and Henriette van der Blom, 133-150. Oxford: Oxford University Press, 2013.
- . "Roman Democracy?" In *A Companion to Greek and Roman Political Thought*, edited by Ryan K. Balot, page-range. Oxford: Blackwell Publishing Ltd, 2009.
- . *The Patrician Tribune*. Chapel Hill, North Carolina: University of North Carolina Press, 1999.
- Yakobson, Alexander. *Elections and Electioneering in Rome: A study in the political system of the late republic*. Stuttgart, Germany: Franz Steiner Verlag, 1999.

After the Fact: Livy, Galileo, and a Historical Understanding of Truth

Madeline Perkins

Abstract

Livy, an ancient Roman historian, communicated the events of his history in a way that was not strictly factual. Though this offends the modern sense of a factual, historical narrative, Livy was not lying. He constructed speeches to communicate patriotism, changed facts to strengthen characterization, and altered story lines to communicate moral lessons. Hundreds of years later, Galileo Galilei entered the same dialogue, discussing the relationship between fact and truth. In his letter to the Grand Duchess Christina, which sought to defend the heliocentric theory to the church, he offers multiple perspectives on the relationship between fact and truth. He explains his hermeneutic for dealing with these difficulties and concludes in the end that scientific fact and scriptural truth never truly disagree.

Both authors successfully accomplish the purpose of promoting truth. Because teaching truth was Livy's object, when he distorted fact, he did not obscure this object. Galileo argued throughout his letter that Scripture's object is truth, and the presence of scientific discoveries do not negate this truth, but rather may entail a new interpretation of it in light of the facts.

My favorite books are a set of ten novels called the *Betsy-Tacy* series. These books by Maud Hart Lovelace are set in the early 1900's and narrate the story of two best friends. They start when Betsy and Tacy are five, and end when the girls are grown up and married. When the girls are young, they have all sorts of adventures: they pretend to be beggars, sell colored sand in jars, and make Everything Pudding.¹ The readers experience when Betsy writes stories modeled after dime novels and Tacy runs away from kindergarten because she's scared. As the girls grow up, the stories grow in maturity. Betsy tries to win the school-wide essay competition each year of high school, Tacy becomes braver about singing in front of people, and the girls navigate parties, classes, and the difficulties of not having naturally curly hair. These books have interesting stories, an enchanting timeframe, and endearing characters, yet one attribute of these stories makes them uniquely captivating: The main character Betsy is modeled after Maud Hart Lovelace herself. It is remarkable that after reading about Betsy's dream of becoming a writer throughout each book, the manifestation of this dream literally rests in the hands of the reader. The readers know that the stories blend factual interactions with delightful imagination, and the end picture is the truth about Betsy's life.

As she wrote these stories, Lovelace necessarily made some alterations to the factual occurrences of her own life. She introduced the character who becomes her husband in high school, though in her own life she did not meet him until later. She fused a childhood friend and a high school friend into one character to introduce continuity.² She changed ages and names, and the resulting picture is beautiful. Lovelace's blend of fact and imagination is not new, in fact, this style of writing was present even back in the Roman empire.

Livy wrote a history of the Roman people. A cursory reading of his work, however, would cause any critical reader to doubt the veracity of its status as a history, for it does not resemble a modern history. It contains individuals giving speeches that Livy made up or embellished.³ If Livy inserted imagination into his history text, then his sources ask to be called into question.

Livy's use of sources does not satisfy the modern historian. One major issue is his failure to employ primary sources. Robert Ogilvie, author of an extensive commentary on Livy's work, contends that Livy did not reference the *Annals*, a history published by one of Rome's priests around 123 B. C. E., or the original tablets on which these Annals were based, though Ogilvie believes that the contents of these were a viable and cogent compilation of fragments of stories.⁴ Ogilvie and Walsh both point out that Livy also did not reference the *libra lintei*, a record of Roman magistrates believed to be a possible alternative to that of the annals.⁵ Livy also did not even verify details of places that he could have visited.⁶ Ogilvie sympathizes with Livy's failure to do his own research, yet he points out that Livy did not even utilize information that had been compiled for him by someone else.⁷ Another weakness in Livy's use of sources is in his habit of relaying an entire story from one perspective. Walsh says that his mention of alternate sources "would be forgivable if Livy had consulted the secondary authorities . . . as to amend or modify .

¹ Lovelace. *The Betsy-Tacy Treasury*. 77ff.

² Whalen. *The Betsy-Tacy Companion*.

³ Kraus. *Livy Ab Vrbe Conditia*. 12.

⁴ Ogilvie. *A Commentary on Livy*. 6.

⁵ *Ibid.*, 6; Walsh. *Livy*. 111-2.

⁶ Ogilvie. *A Commentary on Livy*. 6.

⁷ *Ibid.*, 6.

. . . [i]nstead, he allows his main narrative to rest on the factual information of a single authority.”⁸ Ogilvie agrees, saying that Livy “accepted and adapted a single version for the main narrative” but did nothing more with other sources besides mention them.⁹ Additional proof of this deficiency is that the the great majority of the information in his first five books relies on only two sources.¹⁰ A final issue with Livy’s use of sources is that even when the original text of speeches was available, Livy would write his own.¹¹ This intentional alteration of sources, so abhorrent to the modern historian, comes from Livy’s desire to make his history moral.

One way Livy instructs the reader in morals is by characterization—both indirect and direct—even when it distorts historical facts. One example of Livy altering details to strengthen moral characterization is in the story of Veturia. Veturia’s son Coriolanus was preparing to march on Rome, so the women of Rome persuaded her to go to him and beg him to make peace. When Coriolanus attempted to greet her arrival with an embrace, she rejected him and delivered an impassioned, patriotic speech that derogated him for his disloyal behavior. Though his wife and children wept and embraced him, Veturia did not engage in this behavior.¹² A different author portrayed Veturia as weeping and did not mention the rejected embrace, though this author used the same reference source as Livy. Because Livy sought to make Veturia the ideal Roman mother, he changed these details to highlight her virtue and dedication toward the state.¹³ Veturia is not the only example of Livy’s characterization of Roman women.

Another example of Livy characterizing Roman morality is the story of Lucretia. Livy tells her story: her husband proved her to be the most virtuous wife among the princes, for when the men rode home unannounced, she was working even though it was late at night. When one of the men came back a few days later and attempted to seduce her, Lucretia only yielded to his advancements when he threatened to kill her and make her dead body look like she had slept with a slave.¹⁴ Livy comments that “[e]ven the most resolute chastity could not have stood against this dreadful threat.”¹⁵ After the transgressor had departed, Lucretia confessed the violation to her father and husband, and, though understanding that she was innocent by intention, she killed herself, saying that she would not “provide a precedent to unchaste women to escape what they deserve.”¹⁶ Each of Lucretia’s actions narrated by Livy paints an ideal Roman woman: Lucretia shows beauty, diligence, and chastity.¹⁷ Livy does not allow for Lucretia to make mistakes or act in contradiction to this ideal; she is not a person with flaws, but a character with instructive attributes.

Almost every character in the History could also be analyzed in this way. In the first volumes, Livy regularly retains characterization of familial lines, only changing them to further sharpen the characterization.¹⁸ Later in the work, he reshapes entire facets of certain characters in

⁸ Walsh. *Livy*. 141.

⁹ Ogilvie. *A Commentary on Livy*. 5.

¹⁰ *Ibid.*, 7.

¹¹ Kraus. *Livy Ab Vrbe Condita*. 12.

¹² Livy. *The Early History of Rome*. 156-7.

¹³ Walsh. *Livy*. 91.

¹⁴ Livy. *The Early History of Rome*. 100-1.

¹⁵ *Ibid.*, 101.

¹⁶ *Ibid.*, 102.

¹⁷ Walsh. *Livy*. 13, 214.

¹⁸ *Ibid.*, 92-3.

order to fashion them into his ideal character. An example of this is when Livy changed the religious stance, political ambition, and military decisions of Scipio Africanus, which he did to make Scipio the closest character in the work to the ideal Roman. Walsh summarizes Livy's treatment of Scipio by saying that he "concentrated our attention especially on his moral qualities, and he has not hesitated to suppress or modify his less desirable traits in order to present an edifying portrait."¹⁹ One example of this is that Livy reported that Scipio portrayed himself as divinely inspired, but Livy also showed his disbelief in this superstition.²⁰ Livy alters historical facts to fit characterization many more times, such as when he characterizes Flaminius and Marcellus as selfless and patriotic toward Rome and eliminates facets of their personalities that are harsh or unethical.²¹ Walsh aptly elucidates the purpose of these alterations, saying that "[f]rom these examples of characterization of individuals, it is clear that Livy seeks to bring out their predominantly moral qualities. . . ."²²

Another way Livy changed facts was by changing story lines. He would force stories from their natural chronology and structure into moral episodes.²³ One specific instance of this, though examples permeate the book, is when he eliminated two consuls and made two campaigns seem like they were one.²⁴ Livy would trim the historical facts into a cohesive story, which, Ogilvie says, "[resulted in] historical nonsense but tense reading."²⁵ Though this tactic changed the facts, it kept the stories tight, easy to read, and morally instructive.

Much of Livy's scholarship merely reflects the moral emphasis of ancient historiography. Walsh says that Roman historians added, removed, and altered details because they were "motivated by an indiscriminating patriotism . . . This complex of patriotic and gentile partiality became the 'authorised version' of earlier Roman history."²⁶ The Romans adopted the tradition of using history to teach moral lessons from the Greeks.²⁷ He claims that "Livy's writing is also conspicuously affected by the view that history is a medium for moral instruction . . . [t]oday most historians in theory attempt to eschew moral judgements."²⁸ Christina Kraus, editor of a commentary on Livy, argues that Livy's writing shifts from confidence and explanation to doubt and questioning in order to engage the reader in the process of reading history. He seeks this reader participation in order to accomplish his purpose of teaching a moral lesson.²⁹

As evidenced above, Livy's book is not strictly factual, nor does it seek to be. It not only deviates from sources in ways that were acceptable in ancient historiography. As Foster writes in his introduction to Livy, "[Livy] declines to vouch for [the stories'] authenticity, though he means to set them down as he finds them; and he apparently regards them as possessing a certain symbolic truth, at least."³⁰ Livy's work—though not factually accurate by a modern

¹⁹ Walsh. *Livy*. 95-101.

²⁰ *Ibid.*, 94.

²¹ *Ibid.*, 101-2.

²² *Ibid.*, 107.

²³ Ogilvie. *A Commentary on Livy*. 18.

²⁴ Livy. *The Early History of Rome*. 151.

²⁵ Ogilvie. *A Commentary on Livy*. 18.

²⁶ Walsh. *Livy*. 31.

²⁷ *Ibid.*, 26.

²⁸ *Ibid.*, 39.

²⁹ Kraus. *Livy Ab Vrbe Condita*. 13-4.

³⁰ Foster. *Livy*. xxi.

understanding—was able to communicate truth to ancient readers in its moral lessons. The modern reader is also able to learn about Roman virtue from Livy’s work, and because he or she is already skeptical of the work’s facts, he or she will be able to see the moral intent more clearly. This clarity of intent is what Galileo seeks to provide in his exposition on scientific and Scriptural truth.

In Galileo’s letter to the Grand Duchess Christina, which he wrote to defend the heliocentric theory in the eyes of the church, he offers multiple perspectives on the relationship between fact and truth. His final conclusion is that Scriptural truth and scientific fact will not disagree. This clear resolution is convoluted in practice, however, by humans: Scripture can be polluted by wrong interpretation, and scientific fact can be theorized without definitive evidence. Galileo explains his hermeneutic for dealing with these difficulties in the letter and is able to conclude in the end that Truth and Scripture never truly disagree.

Galileo is helped in his exposition on Scripture by Augustine’s *Literal Commentary on Genesis*. Augustine’s aim in this work was much the same as Galileo’s—defending Scripture’s truth when it apparently contradicted accepted scientific facts—and both his logic and his authority aid Galileo’s argument.³¹ The *Cambridge Companion to Galileo* explains different maxims offered by Augustine and utilized by Galileo in an attempt to reconcile two truths that appeared disparate. The first maxim is the principle of prudence, which urges the reader to remember that different interpretations of Biblical texts are possible, and therefore one reading should not be rashly defended over another without absolute justification. The second maxim concerns the priority of demonstration: If a scientific fact has been thoroughly demonstrated—despite any apparent contradiction to Scripture—the interpretation of the Scripture must be changed. The third maxim discusses the priority of Scripture, claiming that if a fact has not been thoroughly demonstrated, Scripture should continue to be read literally. Another principle concerns accommodation: this view holds that the words of Scripture have been adapted for its ancient audience. The final maxim is that of limitation, which claims that because the primary purpose of Scripture is salvation, Scripture does not report in a scientific manner about scientific facts.³² Galileo utilizes each of these principles in his argument in the letter.

Galileo opens his letter with the principle of prudence. He claims that men should not make absolute statements about things with potentially obscure meanings.³³ This principle will be helpful to his argument because it applies generally to people who disagree with him without reaching the specific and explosive issue of the heliocentric theory. He tells of the “great circumspection” used by early church fathers in treating points of Scripture that could conflict with physical demonstrations and points out the harm that could come to the church if these were believed to be accepted church doctrine.³⁴ Concerning this issue he referenced the same Augustine quotation twice: “[w]e ought not to believe anything inadvisedly on a dubious point, lest in favor to our error we conceive a prejudice against something that truth hereafter may reveal to be not contrary in any way to the sacred books.”³⁵ The second time he used it, he juxtaposed it with Augustine’s opinion that when experiences are proved, the Bible’s words must

³¹ McMullin. “Galileo on Science and Scripture.” 291.

³² Ibid., 292-8.

³³ Galilei. “A Letter to the Grand Duchess Christina.” 120.

³⁴ Ibid., 135.

³⁵ Ibid., 120, 131.

be shown, even if it required reinterpretation, to agree with these facts,³⁶ which leads into the next maxim.

The second maxim used by Galileo supports the priority of demonstration, which claims that if a scientific fact is proven against the apparent meaning of Scripture, then the interpretation of the Scripture must be changed. Galileo claims that “if one were always to confine oneself to the unadorned grammatical meaning, one might fall into error.”³⁷ This statement aligns well with the principle of prudence, and this alignment bolsters Galileo’s movement into the second maxim. He says that “it is first to be considered whether anything is demonstrated beyond doubt or known by sense-experience . . . if it is . . . it ought to be applied to find out the true senses of holy Scripture.”³⁸ This quotation agrees with the maxim offered by Augustine; the maxim is important because it allows for scientific knowledge and truth to coexist, which is necessary to the argument of Galileo’s letter. Elsewhere he echoes this idea and follows it with the principle that the two truths—inspired Scripture and demonstrated fact—cannot disagree.

The third maxim Galileo uses is the priority of Scripture. He claims that when a fact is conjecture and has not been thoroughly demonstrated, Scripture passages concerning it must be assumed to be literal.³⁹ He only uses this maxim once in the letter, likely because he attempts to treat the heliocentric theory as demonstrated fact, and this maxim addresses facts that have not been clearly demonstrated. In his book on Galileo, Jerome J. Langford argues that presenting this maxim is a misstep because Galileo cannot empirically prove the heliocentric theory. He claims that because Galileo had not been able to prove the theory, his presentation of the maxim logically defeated his argument, because it required that Scriptural texts about the sun must be treated as literal.⁴⁰ Though this was not the best rhetorical choice, because it appears to weaken Galileo’s argument, it does not entirely defeat the argument: When Galileo reinterprets the Scripture passages about the sun, he retains a literal interpretation. This, then, allows him to still place priority on a literal reading of Scripture in agreement with the third maxim.

Galileo also argues for the heliocentric theory using the maxim of accommodation. He claims that because the Bible was written for the common people, it presented its information in ways that would not confuse them. Because scientific issues are secondary to matters of faith, Biblical authors did not elucidate them with scientific precision if doing so would distract the common people from salvation.⁴¹ Galileo gives an example of Aquinas analyzing Scripture in this way: Aquinas explained that the Bible, to accommodate the common people, calls the area above earth “void” even though it is actually filled with air.⁴² He also supports this maxim by pointing out that the Bible discusses God as if he had human qualities, both emotional and physical, yet the Church understands that these are only used to provide clarity to the reader.⁴³ He draws a comparison between this and heliocentrism by claiming that in both instances, the

³⁶ Galilei. “A Letter to the Grand Duchess Christina.” 131.

³⁷ Ibid., 122-3.

³⁸ Ibid., 131.

³⁹ Galilei. “A Letter to the Grand Duchess Christina.” 130.

⁴⁰ Langford. *Galileo, Science, and the Church*. 73-4.

⁴¹ Galilei. “A Letter to the Grand Duchess Christina.” 131-2.

⁴² Ibid., 132.

⁴³ Ibid., 123.

authors of Scripture choose to “obscure some very important pronouncements . . . [in] matters infinitely beyond the comprehension of the common people.”⁴⁴

The final maxim used by Galileo is the principle of limitation. Galileo claims that “the primary purpose of the sacred writings . . . is the service of God and the salvation of souls.”⁴⁵ He claims that the texts did not intend to teach about scientific matters, for if they did, they would not have avoided the subject so completely.⁴⁶ He also claims that because heliocentrism does not have any bearing on salvation and is not clear in the text, one interpretation cannot be chosen over another as if it were an essential matter of faith.⁴⁷

Each of these maxims builds to form Galileo’s hermeneutic, which allows him to claim that the two truths—Scriptural and scientific—do not disagree. This alignment of truths is the central thrust of his argument. He states this multiple times throughout the letter: “two truths can never contradict each other.”⁴⁸ This fulfills the purpose of his work because he wants his theory to be accepted, which could only happen if the church believed that the theory agreed with Scriptural truth.

Truth is a similar theme in the works of Livy and Galileo. Both authors seek truth, yet the truths they seek are different. Livy’s truth is the morality he shows through his stories, while Galileo’s truth is the salvation presented in Scripture. Despite this difference in definition, fact and truth interact similarly in both works. Livy obscures fact in an attempt to enhance truth. Galileo points out that the relationship between fact and truth in Scripture is often ambiguous: Scripture does not always relate facts—well, factually. This, then, shows that the text of Livy and Galileo are similar: they both distort fact in seeking to place greater emphasis on truth. Their purpose in this distortion is to share their truth.

Both authors successfully accomplish the purpose of arguing for truth. In his introduction, Livy says, “I invite the reader’s attention to the . . . consideration of the kind of lives our ancestors lived.”⁴⁹ Because history was his medium for moral instruction, he manipulated this medium to best further truth. Because teaching truth was indeed his object, when he distorted fact, he did not obscure his purpose. Galileo argued throughout his letter that Scripture was intended to promote salvation. This argument, discussed at length in reference to his use of Augustine’s maxims, fulfills Galileo’s purpose because he defined truth as the contents of Scripture. Consequently, Galileo also was able to argue that Scripture successfully accomplished its purpose. Livy was not writing a modern history; as Walsh points out, he does not include social or economic influences that the modern is so accustomed to,⁵⁰ and Galileo says that the Bible is not trying to tell us about the specific formation of the cosmos.⁵¹ The purpose of both texts was to share truth, not fact. Their purposes, furthering truth, are deemed by the authors to be more important than the facts that were altered or misrepresented.

Despite the purpose behind it, moderns take issue with this distortion of facts because, to the modern mind, fact and truth are equivalent. This, however, has not always been the case.

⁴⁴ Galilei. “A Letter to the Grand Duchess Christina.” 123.

⁴⁵ Ibid., 123.

⁴⁶ Ibid., 124.

⁴⁷ Ibid., 125.

⁴⁸ Ibid., 130.

⁴⁹ Livy. *The Early History of Rome*. 30.

⁵⁰ Walsh. *Livy*. 34.

⁵¹ Galilei. “Letter to the Grand Duchess Christina.” 123.

Karen Armstrong discusses two ways the ancients arrived at truth in her book *The Battle for God*. There were two kinds of truth, *mythos* and *logos*. *Mythos* was considered primary, and it communicated meaning about what was meaningful and timeless. *Logos* was pragmatic and rational, and sought new knowledge. To the ancient mind, both were necessary to their conception of truth. An understanding of *mythos* has since been lost, so the modern mind has difficulty grasping the complimentary natures of the two truths.⁵² Armstrong says that ancient people categorized history as *mythos*; people were “less interested . . . in what actually happened, but more concerned with the meaning of an event.”⁵³ This disparity between the ancient and modern understanding of truth leads to confusion of modern people when considering the interplay of truth and fact. Livy lived in the time of *mythos* and *logos*, and Galileo was dealing with Scripture, which was also written during that time. Galileo defended the truth of scripture both as *mythos* and *logos* because he sought to argue that it was legitimate not only for the ancients, who would have accepted its *mythos*, but also for his contemporaries, who would have been more comfortable addressing the situation from a perspective that included *logos*.

Galileo argued that Scriptural truth supersedes fact in accordance with the principle of priority of Scripture. This is what leads to his agreement with Livy, as both men seek truth in their writing. Galileo also argues, however, that Scripture and fact will never disagree. Though this may seem to contradict both Livy and the opinion Galileo himself established earlier, it does not, because he still relegates truth and fact to their places, as seen by the principles of limitation and of priority of demonstration. This delicate agreement is embodied by his treatment of the Joshua text: Though he argues that it is permissible in certain cases to depart from a literal interpretation of Scripture, Galileo still explained how the Joshua text could align with a heliocentric view of the world.⁵⁴

Though the texts of Livy and Galileo are more different than similar, they can be juxtaposed and examined in light of how they interact with truth and fact. This interaction yields analysis and maxims, similarities and differences, and each of these gives an interesting perspective on truth. These two texts seek to communicate truth, whether the truth of a historian or God, and follow the western tradition of seeking after truth.

⁵² Armstrong. *The Battle for God*. xiii-xv.

⁵³ *Ibid.*, xiv.

⁵⁴ Galilei. “Letter to the Grand Duchess Christiana,” 137-8.

Bibliography

- Armstrong, Karen. *The Battle for God*. New York: Alfred A. Knopf, 2001.
- Foster, B. O., trans. *Livy*. Cambridge: Harvard University Press, 1961.
- Galilei, Galileo. "Letter to the Grand Dutchess Christiania." In *Reformation, Revolution, and Enlightenment*, ed. by Aaron B. Carter.
- Kraus, Christina Shuttleworth. *Livy Ad Vrbe Condita Book VI*. Cambridge: Cambridge University Press, 1998.
- Langford, Jerome J. *Galileo, Science, and the Church*. South Bend, IN: St. Augustine's Press, 1998.
- Livy. *The Early History of Rome*. London: Penguin Books, 2002.
- Lovelace, Maud Hart. *The Betsy-Tacy Treasury*. New York: Harper Perennial, 1979.
- McMullin, Ernan. "Galileo on Science and Scripture." In *The Cambridge Companion to Galileo*, ed. by Peter Machamer. Cambridge: Cambridge University Press, 1998.
- Ogilvie, R. M. *A Commentary on Livy*. Clarendon: Oxford University Press, 1965.
- Walsh, P. J. *Livy: His Historical Aims and Methods*. Cambridge: Cambridge University Press, 1961.
- Whalen, Sharla Scannell. *The Betsy-Tacy Companion: A Biography of Maud Hart Lovelace*. Portalington Press, 1995.

Natural Slaves as a Product of the Civil State: Analysis of the Natural State and Human *Telos* in the Political Writings of Aristotle and Rousseau

Casey Cunningham

Abstract

Although Aristotle and Rousseau seem to have opposite ideas about both society and slavery, analysis of *Politics* and *The Social Contract* reveals that both philosophers believe “natural” slavery is possible, but it is a product of society, not the inherent purpose of any human. Their differing views of human purpose, or *telos*, are the points that separate their ideologies. In both, natural slaves exist, but they are made by society, not born. Aristotle accepts slavery as an integral part of the natural *polis* and a necessary instrument with which citizens are enabled to reach their *telos*—a virtuous life. Rousseau, however, rejects the institution of slavery as a result of the disastrous effects of the unnatural inequality created by the exit from the natural state into civil society, which is directly opposed to the Rousseauian *telos*: freedom. By considering Rousseau and Aristotle together, one comes to understand that both thinkers desired the greatest amount of freedom and happiness within a political society, and their opposing views of slavery are not, as many would assume, because one philosopher was more humanitarian or egalitarian than the other, but because of their different understanding of the human purpose.

In modern political thought, the institution of slavery is unequivocally condemned as an unnatural and immoral violation of human rights. This outlook provokes negative reactions of many contemporary readers to Aristotle's *Politics*, in which it seems that some people are naturally born to be slaves of others. Scholars through the years have contributed a multitude of interpretations of the Aristotelian account of natural slavery in the state, or *polis*, attempting either to attribute all of Aristotle's ideas to his cultural bias, or to reveal him as an inconsistent and ethnocentric supremacist attempting to justify slavery in order to enhance his own social superiority¹. However, when one compares Aristotle's philosophy of natural slavery as a characteristic of the family unit with the political writings of French philosopher Jean-Jacques Rousseau, *On the Origins of Inequality* and *The Social Contract*, which firmly condemn the institution of slavery, another explanation becomes evident. Although Rousseau and Aristotle seem to have opposite ideas about both society and slavery, both philosophers believe "natural" slavery is possible, but it is a product of society, not the inherent purpose of any human. Their differing views of human purpose, or *telos*, are in fact the points that separate their ideologies. In both, natural slaves exist, but they are made by society, not born. Aristotle accepts slavery as an integral part of the natural *polis* and a necessary instrument with which citizens are enabled to reach their *telos*—a virtuous life. Rousseau, however, rejects the institution as a result of the disastrous effects of the unnatural inequality created by the exit from the natural state into a civil society that is directly opposed to the Rousseauian *telos*: freedom. By considering Rousseau and Aristotle together, one comes to understand that both thinkers desired the greatest amount of freedom and happiness within a political society, and their opposing views of slavery are not, as many would assume, because one philosopher was more humanitarian or egalitarian than the other, but because of their different understanding of the human purpose.

The Aristotelian "Natural" Slave

Slavery is an essential component of the family unit that forms the backbone of Aristotle's *polis*. He lists the relationship between master and slave even before that of man and wife in its centrality to the life of the political man.² The characteristics that qualify one as a "natural" slave, that is, a human who is innately and justifiably designated for a life of physical labor instead of one of rational, philosophical thought, however, may seem unclear to a modern reader. Although a simplistic reading of the first book of *Politics* may give the idea that by nature, slaves are subhuman and lack the capacity to live as citizens, close reading reveals that Aristotle actually views natural slaves as persons with the same level of humanity as free citizens but born with personalities or abilities less inclined to reasoning. In a description of the natural slave, he distinguishes: "Those whose condition is such that their function is the use of their bodies and nothing better can be expected of them, those, I say, are slaves by nature."³ This quote suggests that the institution of slavery is right and even necessary for certain members of the *polis* because their function is only physical and does not involve any higher, intellectual thought. The distinction between the slave and the free person, according to Aristotle, is not a distinction based upon social class or level of power. Instead, as Jill Frank explains, "There is

¹ Darrell Dobbs, "Natural Right and the Problem of Aristotle's Defense of Slavery," *The Journal of Politics* 56, no. 1 (1994): 71.

² Aristotle, *Politics*, trans. T. A. Sinclair, ed. Trevor J. Saunders (Harmondsworth: Penguin Classics, 1981), 1253b1.

³ *Ibid.*, 1254b16.

nothing immutable that singles out any particular person as a slave. Instead, slave identity, like citizen identity, is determined by activity. If this is right, then there is no . . . difference between slave and citizen.”⁴ Frank’s interpretation posits that any person whose actions do not use reason, which Aristotle prizes as that which separates man from animals,⁵ can be considered to have the nature of a slave. However, by this logic, not only do slaves have the ability to become free citizens through use of their “deliberative faculties,”⁶ free citizens could descend to slavery if they neglect their political duties. Though Aristotle does not explicitly describe the procedure for this descent or the extent to which one would have to ignore reason in order for it to occur, following his argument to its logical conclusion implies that irrational, unvirtuous masters could theoretically lose their status in the *polis*. Thus, although slavery is central to the existence of the Aristotelian society and occurs as a natural result of the tendencies of some to neglect rationality in pursuit of appetitive desires, slaves are not inherently subordinate to citizens, and the “naturalness” of slavery is simply due to its importance in the functioning of the *polis*.

The *Polis* as the Natural State for the Human

The *polis*, or political state, is the fundamental premise of Aristotle’s political theory. It is not only the organization through which citizens associate with one another, but also the natural and therefore ideal state for all humans. As an institution which enables the *polis* to function, slavery becomes justifiable. In fact, Aristotle asserts, “The state belongs to the class of objects which exist by nature, and man is by nature a political animal. Any one who by his nature. . . has no state is. . . either subhuman or superhuman,”⁷ suggesting that participation in a *polis* is as essential for slaves as it is for citizens. According to Aristotle, those outside of the state are not really human; in order to participate in humanity, one must have a role in the city. Wayne Ambler explains, “the city is natural because it completes or perfects man. If it is true that cities cannot exist without human beings, it comes to appear that human beings cannot exist without cities.”⁸ Although many philosophers, including Rousseau, consider the city to be a departure from the natural state, Aristotle believes that the “political” nature of humans, their inherent sociability, creates natural community. Daniel Richter explains that this community is the defining factor of Aristotelian human nature because “Human beings, alone of mortal creatures, possess the capacity for reason and judgement—faculties which can only be exercised in the context of a community of similarly endowed creatures.”⁹ Outside of the state, a person is not a human but a “monster,” because the qualifying ability to reason is impossible.¹⁰ Richter’s interpretation seems to suggest that slaves are not human, because the factor that determines slavery is lack of reason, but Aristotle also says that a slave “participates in reason so far as to

⁴ Jill Frank, “Citizens, Slaves, and Foreigners: Aristotle on Human Nature,” *The American Political Science Review* 98, no. 1, (2004): 95

⁵ Aristotle, *Politics*, 1253a7.

⁶ *Ibid.*, 1259b32.

⁷ *Ibid.*, 1253a1.

⁸ Wayne H. Ambler, “Aristotle’s Understanding of the Naturalness of the City,” *The Review of Politics* 47, no. 2 (1985): 169-170.

⁹ Daniel S. Richter, *Cosmopolis: Imagining Community in Late Classical Athens and the Early Roman Empire*, (New York, Oxford University Press, 2011): 41.

¹⁰ Darrell Dobbs, 75.

recognize it but not so as to possess it.”¹¹ Because slaves are associated with their rational masters in the natural city, they are able to participate in humanity.

The association between slave and master is one of the two primary relationships which form the household, and a group of households form the *polis*.¹² Because the *polis* is the natural state for humans, it seems that the associations which combine in its composition, including slavery, must also be natural. However, Aristotle makes it clear that reason, the ability which separates people from animals and enables the formation of the *polis*, is the natural function of a human, of which slaves are mere observers, not participants. As humans who do not partake in the naturally human function, slaves are considered to be natural only because they are an integral part of the natural state. Darrell Dobbs states of this paradoxical relationship between the natural slave and the unnatural institution of slavery: “Slavery can exist in accordance with nature only because slavishness exists contrary to nature,”¹³ making it clear that Aristotle views slavery as natural within the hierarchy of the state but does not believe that the condition of slavery itself is natural. Dobbs concludes, “[The] natural slave derives from the fact that all human beings belong in a life partaking in the distinctively human *telos*.”¹⁴ Because it is better to be a part of the state than outside of it, slavery is a justifiable way to help those less capable of reason. Through this institution, slaves as well as their masters are able to fulfill their natural purpose.

Aristotelian *Telos* and the Role of Slavery in Achieving the Good Life

Slavery is mutually beneficial to both masters and slaves in the *polis* because it enables each to get closer to achieving their *telos*. *Telos* is the Greek word Aristotle uses to describe the highest end of human existence, the purpose humans exist to fulfill. In his philosophy, he identifies human *telos* as *eudaimonia*, a word commonly translated in English as “happiness,” which is more accurately described as “living well and doing well.”¹⁵ In *Nicomachean Ethics*, Aristotle explains that “the function of man is to live a certain kind of life, and this activity implies a rational principle [. . .] so happiness [*eudaimonia*] turns out to be an activity of the soul in accordance with virtue.”¹⁶ Thus a life lived to achieve the *telos* of *eudaimonia* is a life in which a human uses reason to become virtuous. Dobbs suggests that “Aristotle's defense of natural slavery must be understood as intrinsic to his teleological account of the human good,”¹⁷ and this makes sense when one considers that the relationship between slave and master allows each to become more virtuous.

For Aristotle, all actions and associations should help one towards *telos* by promoting virtue and reason. When explaining the function of slavery in the *polis*, he must show how slaves and masters alike grow in virtue through their partnership. However, when slaves do not participate in reason like citizens, it is difficult to explain how and why they attain virtue. Aristotle addresses this at the end of the first book of *Politics*:

¹¹ Aristotle, *Politics*, 1254b16.

¹² Ibid., 1252a34-1252b27.

¹³ Darrell Dobbs, 86.

¹⁴ Ibid., 87.

¹⁵ Caroline Janeway, “Aristotle’s *Nicomachean Ethics*,” Lecture for Virtues 101 at Samford University, Birmingham, AL, September 22, 2014.

¹⁶ Aristotle, *Nicomachean Ethics*, trans. Terence Irwin, (Indianapolis, Hackett Publishing Co., 1999): 1098a13.

¹⁷ Darrell Dobbs, 92.

About slaves the first question to be asked is whether in addition to their virtue as tools and servants they have another and more valuable one. Can they possess restraint, courage, justice, and every other condition of that kind, or have they in fact nothing but the serviceable quality of their persons? . . . If we say that slaves have these virtues, how then will they differ from free men? If we say that they have not, the position is anomalous, since they are human beings and share in reason.¹⁸

This predicament is solved by his conclusion that “both ruler and ruled must have a share in virtue, but that there are differences in virtue in each case.”¹⁹ The key difference in slaves and citizens is that slaves lack a “deliberative faculty.”²⁰ In their daily activities they require virtue, but “only enough to ensure that [they do] not neglect [their] work through intemperance or fecklessness.”²¹ Citizens, in contrast, must have all of the moral virtues and the time to use them well. Without slaves to take care of the labor of daily life, citizens would not be free to exercise reason and virtue to their full capacity, which is their purpose in life. The *polis* exists to assist in this purpose, so its institutions, including slavery, are meant to promote this end. Frank explains that slavery “is necessary to free masters from meeting their daily needs so that they can, as rulers or citizens, practice politics,”²² and Dobbs corroborates, suggesting that, “The natural despotic partnership is a mutually beneficial association wherein a master gains studious leisure by procuring in a noble way some of the necessities of life through his slave.”²³ By taking care of menial labor, slaves provide their masters with the time needed to participate in the life of the state and turn their thoughts to the higher ideas of virtue and reason. The slaves also benefit from this relationship because the Aristotelian slave “is in a sense a part of his master, a living and separate part of his body.”²⁴ Slaves, despite being “human beings possessed with reason,”²⁵ do not actively display the mental capacity for reason or moral virtue. This arrangement means that they can still partake in *telos* through the virtue of their masters. In this mutually beneficial relationship, slaves are not dehumanized, but instead become human through their enslavement by entering into relationship with, and, according to Aristotle, becoming part of their virtuous, rational master. When viewed in this way, the institution of slavery is essential to the proper functioning of the state, and even more importantly, to the ability of each person to live a good life in accordance with his or her inherent purpose.

Rousseau’s Condemnation of Natural Slavery

In many ways, Jean Jacques Rousseau rejects the Aristotelian ideas of slavery, the natural state, and the purpose of life. However, when he directly addresses Aristotle’s view of slavery in *The Social Contract*, he admits that “Aristotle, before all others, had also said. . . that some were born for slavery. . . . Aristotle was right, but he took the effect for the cause.”²⁶ Rousseau agrees

¹⁸ Aristotle, *Politics*, 1259b18.

¹⁹ Ibid., 1259b32.

²⁰ Ibid., 1259b32.

²¹ Ibid., 1260a14.

²² Frank, 92.

²³ Dobbs, 87.

²⁴ Aristotle, *Politics*, 1255b4.

²⁵ Ibid., 1259b18.

²⁶ Jean Jacques Rousseau, *Basic Political Writings*, ed. Donald A. Cress (Indianapolis: Hackett Publishing Company, 1987): 142-143.

with Aristotle that slaves can be “natural,” but while Aristotle attributes the naturalness of slavery to its position as a foundational association of the natural *polis*, Rousseau asserts, “If there are slaves by nature, it is because there have been slaves against nature. Force has produced the first slaves; their cowardice has perpetuated them.”²⁷ Slavery began as a result of inequality, when those with power oppressed the weak by force. As it continued, it became natural to the oppressed people. As a result of their change in nature, Rousseau suggests that slaves are no longer humans:

Renouncing one’s liberty is renouncing one’s dignity as a man, the rights of humanity and even its duties. There is no possible compensation for anyone who renounces everything. Such a renunciation is incompatible with the nature of man. Removing all liberty from his will is tantamount to removing all morality from his actions.²⁸

Unlike the Aristotelian slave, who maintains humanity and has a share in virtue, the natural slave of Rousseau’s philosophy rejects humanity and gives up morality and basic rights. Although the institution becomes natural when perpetuated in the civil state, it is by no means justifiable, and “from every point of view, the right of slavery is null, not simply because it is illegitimate, but because it is absurd and meaningless.”²⁹ Although slavery became natural to the civil state, Rousseau explains that the very nature of humanity was altered to legitimize it: “just as violence had to be done to nature in order to establish slavery, nature has to be changed in order to perpetuate this right. And the jurists, who have gravely pronounced that the child of a slave woman is born a slave, have decided, in other words, that a man is not born a man.”³⁰ For Rousseau, slavery is never acceptable because all humans are born equal and should have the same rights. Anything else is not only contrary to nature—it renders one less than human.

Contrasting the Unnatural Inequality of Civil Society with the Free State of Nature

Although slavery becomes natural through its propagation in society, according to Rousseau, the society itself is an unnatural state. In direct contrast to Aristotle’s “political animal” and idea of the natural community in the *polis*,³¹ Rousseau identifies a “state of nature” outside of and entirely separate from the civil state, in which humans are not bound by the inequalities of society. He looks back at previous philosophers, including Aristotle, and says that “all of them, speaking continually of need, avarice, oppression, desires, and pride, have transferred to the state of nature the ideas they acquired in society. They spoke about savage man, and it was civil man they depicted.”³² While Aristotle would say that anyone outside of the *polis* is an inhuman monster, Rousseau believes that the savage, primitive person who exists outside of society is more virtuous than the statesman. Also, because those outside of the state of nature are not in community with anyone, they are entirely independent and free to do as they want. Everyone in this state is equal and “restrained by natural pity from needlessly harming anyone himself.”³³ From his perspective, human nature, until corrupted by society, is opposed to inequality; humans are naturally content to take care of personal needs and have an aversion to

²⁷ Rousseau, 143.

²⁸ Ibid., 145.

²⁹ Ibid., 146.

³⁰ Ibid., 74.

³¹ Aristotle, *Politics*, 1253a1.

³² Rousseau, 38.

³³ Ibid., 64.

suffering. In contrast, one who enters into a civil state is subject “to his fellow men, whose slave in a sense he becomes even in becoming their master; rich, he needs their services; poor, he needs their help”³⁴ Entering into society forces one to sacrifice independence and freedom for unnatural inequality, leading to the development of slavery—the ultimate inequality.

Although the state of nature, outside of society, is the ideal and natural place for humankind, Rousseau does not think humans are able to return to that primitive age. He recognizes that people entered into civil states for a reason and asserts that this reason was the belief that participation in a state, where some rights are sacrificed for the common cause, would better protect their basic rights and freedoms.³⁵ However, although the civil state has the potential to correct “natural inequality” and actually promote equality among citizens,³⁶ Rousseau explains in his *Discourse on the Origins of Inequality* that the introduction of interdependence created vices and greed as people discovered the ability to profit from the subordination of others.³⁷ Thus, once the civil state began, humans were unable to maintain their pure natures, and inequality began. Laurence Cooper explains this shift: “What we cannot do is retrace our steps to an earlier, happier time. . . the felicity of those epochs disappeared irrevocably with the innocence that had been its protective shield. Human nature has been forever changed.”³⁸ With the exit from the state of nature, humans lost not only their independence but also the inherent goodness of their nature. Slavery, impossible before, becomes a reality, and Rousseau laments that “since the bonds of servitude are formed merely from the mutual dependence of men. . . it is impossible to enslave a man without having first put him in the position of being incapable of doing without another. . . a situation that did not exist in the state of nature”³⁹ Now, as unnatural products of a society which has abandoned nature, humans promote the very inequality they entered the state to escape, and the goal of equality and freedom has been forgotten in the greed and unnatural desire to control others.

Freedom as the *Telos* of Rousseau’s Political Philosophy

Like Aristotle, Rousseau believes that there is a highest good for humanity, but for Rousseau, this good is not a virtuous and happy life, but rather a life of freedom. Cooper compares their perspectives: “The good life. . . is for Rousseau, as it was for Aristotle, a life lived in accordance with nature. Or, rather, it is for Rousseau, as it was *not* for Aristotle, a life lived in accordance with nature,”⁴⁰ illustrating the difference in the Aristotelian concept of what is natural from Rousseau’s beliefs. The state of nature for Rousseau is one in which man is totally free from society. Aristotle’s natural state is the society, so his philosophy justifies slavery; the inequalities of civil associations are natural and helpful in achieving a life of virtue. Of course, Rousseau’s rejection of civil society prevents any acceptable form of slavery; such an institution cannot fit with his view of the human good.

Differing conceptions of *telos* allow the two philosophers to maintain opposite views of “natural” slavery. Although many modern readers assume that any philosophy in which slavery

³⁴ Ibid., 67.

³⁵ Ibid., 151-153.

³⁶ Ibid., 153.

³⁷ Ibid., 38.

³⁸ Laurence Cooper, *Rousseau, Nature, and the Problem of the Good Life*, (University Park: Pennsylvania State University Press, 1999): 17-18.

³⁹ Rousseau, 59.

⁴⁰ Cooper, 40.

is a natural condition for certain people must be oppressive and unjustifiable, examination of Aristotle's perspective of natural slavery in light of Rousseau's view of the natural slave reveals both the reasoning behind his justification of natural slavery and the way in which each philosopher's *telos* determines their understanding of slavery as helpful or hurtful for the slave. As Cooper explains, "Nature, as Rousseau conceives it, is not teleological. It [. . .] does not prescribe any particular way of life for human beings."⁴¹ A classical view of nature, such as that of Aristotle, proposes one correct way of living towards which all of humanity should strive. Rousseau rejects this *telos* when he identifies the "essential gifts of nature" as nothing more than "life and liberty."⁴² His idea of a good life for a human is any life which allows these gifts to rule. As Stanley Hoffmann explains, "Freedom in this world is what he seeks above anything else: freedom from opinion, from prejudice and dishonesty. Freedom is the most human of qualities, greater than reason."⁴³ For Aristotle, reason is the ability which separates humans from animals, through which humans achieve *telos*, but for Rousseau, freedom is the quality which gives the savage humanity. Slavery, because it removes freedom, also takes away humanity, and for this reason above all others, it can never be a legitimate association.

Aristotle and Rousseau, although in agreement that slavery can become a natural part of the civil state, express opposite opinions of the humanity of slaves and the legitimacy of the institution as a result of their different views of the natural state and purpose for humans. Aristotle's often-quoted statement, "Man is by nature a political animal,"⁴⁴ and Rousseau's famous introduction to *The Social Contract*, "Man is born free, and everywhere he is in chains,"⁴⁵ neatly summarize their contrasting views of slavery. Aristotle views humans as naturally social, so the community of the *polis* is the natural, ideal state. Slavery is a necessary association in the composition of this state, so it must also be natural and good for humans. All of humanity wants to participate in reason and virtue so as to achieve the *telos* of *eudaimonia*, so slaves and masters alike benefit from the institution. In contrast, Rousseau's human is born as one would be in the state of nature, already partaking in the *telos* of freedom, but the unnatural civil society forces the human to depend on others. Slavery, as the ultimate surrender of freedom, takes away one's very humanity. In this way, Rousseau and Aristotle both begin with an institution of slavery considered to be natural to civil society, but Aristotle ends by justifying slavery while Rousseau unequivocally condemns it. Modern readers of these philosophers are often quick to condemn Aristotle and agree with Rousseau, but comparison of their arguments for and against slavery reveals that both philosophers have the same goal: to allow individuals to participate in humanity and strive towards their *telos*.

⁴¹ Cooper, xi.

⁴² Rousseau, 74.

⁴³ Stanley Hoffmann, "The Social Contract, or the Mirage of the General Will," In *Rousseau and Freedom*, Comp. Stanley Hoffmann and Christie McDonald, (Cambridge, Cambridge University Press, 2010): 174.

⁴⁴ Aristotle, *Politics*, 1253a1.

⁴⁵ Rousseau, 141.

Bibliography

- Ambler, Wayne H. "Aristotle's Understanding of the Naturalness of the City." *The Review of Politics* 47, no. 2 (1985): 163-185.
<http://www.jstor.org.ezproxy.samford.edu/stable/1406669>.
- Aristotle. *The Politics*. Translated by T. A. Sinclair. Edited by Trevor J. Saunders. Harmondsworth: Penguin Classics, 1981.
- . *Nicomachean Ethics*. Translated by Terence Irwin. Indianapolis: Hackett Publishing Company, 1999.
- Cooper, Laurence. *Rousseau, Nature, and the Problem of the Good Life*. University Park: Pennsylvania State University Press, 1999.
- Dobbs, Darrell. "Natural Right and the Problem of Aristotle's Defense of Slavery." *The Journal of Politics* 56, no. 1 (1994): 68-94.
<http://www.jstor.org.ezproxy.samford.edu/stable/2132346>.
- Frank, Jill. "Citizens, Slaves, and Foreigners: Aristotle on Human Nature." *The American Political Science Review* 98, no. 1 (2004): 91-104.
<http://www.jstor.org.ezproxy.samford.edu/stable/4145299>.
- Hoffmann, Stanley. "The Social Contract, or the Mirage of the General Will." In *Rousseau and Freedom*. Compiled by Stanley Hoffmann and Christie McDonald. Cambridge, UK: Cambridge University Press, 2010.
- Janeway, Caroline. "Aristotle's *Nicomachean Ethics*." Lecture for Virtues 101 at Samford University, Birmingham, AL. September 22, 2014.
- Richter, Daniel S. *Cosmopolis: Imagining Community in Late Classical Athens and the Early Roman Empire*. New York: Oxford University Press, 2011.
- Rousseau, Jean-Jacques. *Basic Political Writings*. Edited by Donald A. Cress. Indianapolis: Hackett Publishing Company, 1987.

Freedom and Vengeance: How Late Period Egypt and the Late Classical Greek City-States Mutually Contributed to the Decline and Fall of the Persian Empire

Stephen Boutwell

Abstract

Due to rising difficulties both inside and outside its borders, the Persian Empire lost control of Egypt in the fourth century B.C. The Persians could not tolerate this. The resulting struggle spanned more than half a century. The Egyptians did not fight alone against the Persians. The Greeks, especially Athens and Sparta, were just as eager to damage the Persian Empire as the Egyptians were. In response to increasing threats from Egypt and Greece, the Persians enacted new administrative and defensive policies, but these policies backfired. The result was more chaos and political breakdown in certain regions of the Persian Empire affected by such administrative reforms. However, the ultimate victor would be none of these powers. Instead, a new power would emerge in Macedonia capable of consuming Greece, Persia, and Egypt. Its ability to do so was in part due to a world destabilized by Greco-Egyptian collusion against Persia. This paper will explore the cooperation that occurred between the worlds of ancient Egypt and Greece to undermine the Persian Empire and how the Kingdom of Macedonia took advantage of the circumstances created by this cooperation.

The purpose of this paper is to examine the political connections between Egypt and the Greek city-states, specifically Athens and Sparta, during the fourth century B.C. (400-300 B.C.) and the way the Persian Empire reacted to these relations. I hypothesize that Egypt's connections with Greece and vice versa were extremely strong during this time period. The goal for both the Greeks and the Egyptians involved in the establishment of such relations was to undermine the military capabilities of the Persian Empire. Furthermore, I suggest that the mutual actions of the Athenians, the Spartans, and the Egyptians so critically destabilized the Persian Empire that by 334 B.C., when Alexander the Great began to invade the Persian Empire, the Persians were ill-prepared to face their new enemy. The Persians were never able to fully respond to threats emerging either in Egypt or Greece due to the need to divide their forces between the two fronts. As a result, the Persians developed a policy of destabilizing Greece through the efforts of semi-autonomous satraps (governors) on their northwestern frontiers who possessed only limited forces. Meanwhile, the Great King of Persia himself would dedicate the majority of Persia's forces to crushing Egypt. However, limited Persian involvement in Greece created a power vacuum in that region. The Persians could not fill this power vacuum themselves because they were too devoted to subjugating Egypt. This problem helped Macedonia to ascend to power in Greece and, through Alexander the Great, in the Near East and Egypt. Additionally, new measures in Persian administration that were intended to make the Persian Empire more defensible only caused more chaos for the Persians, a problem that the Greeks, Egyptians, and even Macedonians were quick to capitalize on.

The fourth century B.C. was a time of increased contact between the Greek and Egyptian civilizations. Egypt began to show signs of cultural influence from abroad in ways that exceeded anything else ever seen (though such phenomena were not unprecedented). These signs of foreign influence will be discussed below. This increased Greco-Egyptian contact was precipitated by a mutual need to oppose Persian imperialism, but its results were far more damaging for the Persian Empire than the Greeks or Egyptians anticipated.

The Revival of Persian Power in Greece

The Persians, while having been defeated in Greece by a Spartan and Athenian led Greek coalition in the Greco-Persian Wars (490-449 B.C.), did not stop interfering in mainland Greek political affairs. In 412 B.C., the Persians took advantage of the catastrophic war occurring in Greece between Athens and Sparta known as the Peloponnesian War (432-404 B.C.). Athens and Sparta had been locked in a stalemate; the Persians under Great King Darius II (423-404 B.C.) entered the war on the side of Sparta in order to reestablish their control in western Anatolia (modern day Turkey), a region that had been contested between Athens and Persia since 479 B.C. Persian support allowed Sparta to emerge victorious in the Peloponnesian War, but the shadow of the Persian Empire now was once again cast over the Greek world.¹ The Persians did not make any new attempts to invade Greece in the fourth century B.C., but they did learn that by providing military and especially financial backing to one Greek state or another, they could destabilize the Greek world by encouraging its descent into a constant state of war. Persian support went to whatever city-state could best ensure the security of Persian influence in Anatolia and the Aegean Sea. However, the Greeks were far too divided by their own conflicts to seriously challenge Persian interference. Following the Peloponnesian War the Spartans behaved aggressively towards the other Greek states. From 401-399 B.C., they waged war

¹ Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York, New York: Cambridge University Press, 2014), 157-196.

against the Greek city of Elis and dismantled it. They also attempted to withhold the spoils of the Peloponnesian War from their allies.² As a result, hostility began to grow towards Sparta on the Greek mainland. Growing hostilities between Sparta and the rest of the Greek world ensured political division in Greece. However, if one city-state began to wage war against the Persians, the Great King of the Persian Empire simply aligned with that city-state's opponents on the Greek mainland. The political division of Greece created a great obstacle for any Greek state seeking to limit Persian ambitions.

The Birth of the Athenian-Egyptian Alliance

However, there was one serious obstacle to Persian ambitions. Egypt had revolted in the late 400's B.C. By 399 B.C., it was fully independent of the Persian Empire.³ This independence was achieved largely due to Persian problems elsewhere. In 401 B.C. the Persian Empire was rocked by the Spartan aided revolt of Cyrus the Younger.⁴ After this revolt was quelled, the Persians found themselves fighting in Anatolia against the Spartans for control of the Ionian Greek cities beginning around 400 B.C. However, the Spartans were forced to abandon Anatolia due to Persian-supported attacks from their enemies on the Greek mainland, including Athens.⁵ Thus began the Corinthian War (395-386 B.C.). Then the Persians supported Sparta beginning in 388 B.C. after the Spartans renounced their claim to Ionia. This alliance resulted in the Persian-Spartan seizure of the Hellespont around 387 B.C., which was Athens' link to the valuable grain producing Black Sea region. This action forced Athens to come to peace terms favorable to the Persians and Spartans. But the Persian victory in Anatolia came at a cost. Because they were focused on Cyrus the Younger and the Spartans, the Egyptians slipped away from imperial Persian control and quickly used Egypt's spectacular wealth to prepare for a war against the Persian Empire. The ensuing conflict between the Persians and Egyptians was a major event in the early to middle fourth century B.C. The conflict drew in additional powers.

Egypt at this time resumed its traditional pharaonic culture, including the concept of a semi-divine king who mediated between the gods and the people of Egypt that we remember today as the "pharaoh."⁶ However, one new concept became a major factor for Egypt during its period of independence: intensified relations with the Greek world.⁷ This association was motivated by political reasons in both Greece and Egypt. The Egyptians needed allies if they were to maintain their independence from the Persian Empire. The Persians could never accept Egyptian independence. The possibility of another major power arising in the fourth century B.C. Near East was simply too much of a threat. The independence of Egypt not only insulted Persian imperial prestige, but it also invited revolts from other peoples subject to the Persians. This presented the Egyptians with a serious dilemma—while Egypt could boast easily

² J. B. Bury and A. W. Picard, *The Ancient World: Egypt, Greece, and Persia in the 4th Century BC* (Cambridge: Ozymandias Press, 2016), Kindle Edition.

³ Karol Mysliwiec, *First Millennium B.C.E.: The Twilight of Ancient Egypt*. trans. David Lorton (Ithaca, New York: Cornell University Press, 2000), 158-159.

⁴ Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York, New York: Cambridge University Press, 2014), 177-180.

⁵ *Ibid.*, 184-191.

⁶ B. J. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, "The Late Period, 664-323 BC" in *Ancient Egypt: A Social History*, 2nd ed. (Cambridge, New York: Cambridge University Press, 2012), 279-348.

⁷ *Ibid.*, 342-343.

defendable borders due to geography and tremendous wealth, they could not hold out long term against the power of the Persian Empire. The Persians simply possessed far too many resources and could marshal armies that far outnumbered the Egyptian army at this time. Therefore, allies would be a necessity. The Egyptians developed a policy of creating alliances with other anti-Persian factions throughout the Near East and eastern Mediterranean. Their goal was to coordinate with these factions to destabilize the Persian Empire by supporting rebels and states opposed to Persia in order to ensure that the Persians were too occupied responding to new threats of revolt to be able to fully commit all of their armies to attacking Egypt.⁸ They found ready allies among the Greeks, especially Athens.

The Athenians also desired to create a Greco-Egyptian alliance. The Persians were ultimately victorious in defeating Spartan attempts to expand into Anatolia, but their victory was due largely to the Persian Empire's ability to finance Sparta's Greek enemies in Greece. Persia's policy of meddling with Greek affairs by providing support for whatever city-state might best serve Persian interests proved effective. Additionally, the Persian Empire's ability to coordinate its naval power with Sparta to cut off Athens' grain supply showed how vulnerable Athens could become if the Persians acted in such a way again. Therefore, the Athenians had two goals after 386 B.C. The first was to find an alternate source of grain. The Black Sea trade through the Hellespont was too vulnerable, and Sicily, another traditional source of grain, was left scarred by political and economic chaos in the beginning of the fourth century B.C.⁹ Egypt was the only other source of abundant grain that was accessible for Athens.¹⁰ In order to ensure Athens' ability to import grain from Egypt, it was vital that a pro-Athenian—and anti-Persian—government be in power there. The second goal was to diminish Persia's ability to contribute military forces to the Aegean if warfare was to break out again in Greece. Egypt could also serve this purpose. The threat of Egyptian attacks or Egyptian inspired revolts across the Persian Empire, as well as Persia's undying commitment to retaking Egypt, meant that the Persians would be drawn away from the Aegean as long as Egypt was a threat.

Of course, if Athens was to meet these goals Egypt had to stay independent. But a clearly visible alliance between Athens and Egypt might become dangerous. The Persians might simply coordinate with Athens' enemies in Greece again, such as Sparta, in order to force Athens' to cease its support for Egypt. Therefore, a different approach was taken. The Athenians dispatched commanders, and possibly troops, who would play the role of mercenaries.

⁸ Alan B. Lloyd, "The Late Period" in *The Oxford History of Ancient Egypt*, ed. Ian Shaw (New York: Oxford University Press, 2002), 385-390.

⁹ Josiah Ober, "Disorder and Growth, 403-340 BCE" in *The Rise and Fall of Classical Greece* (Princeton, New Jersey: Princeton University Press, 2015), Kindle Edition. Despite the overall successful-albeit tyrannical-career of Dionysius I of Syracuse in the early 300's B.C., Ober believes that warfare and internal political confusion between the Sicilian Greeks themselves and between the Greeks and the Carthaginians prevented the Sicilian Greek states from creating effective economies throughout much of the fourth century B.C.. He believes that Sicilian grain production and exportations particularly were affected by Sicily's chaotic state of affairs.

¹⁰ B. J. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, "The Late Period" in *Ancient Egypt: A Social History 2nd* ed. (Cambridge, New York: Cambridge University Press, 2012), 325-328.

The Egyptians extensively recruited Greek mercenaries at this time anyway.¹¹ Such a disguise for Athenian commanders would be easy to sell. These commanders would provide the Egyptians with Greek military skills and expertise. They likely helped oversee the training and command of Greek mercenaries, covert Athenian troops (or at least mercenaries hired by Athens to serve in Egypt under Athenian commanders), and even native Egyptian troops. Egypt had not possessed a formal army of its own since before 525 B.C. when it was conquered by Persia. Athenian commanders could help recreate a standing Egyptian army. Athenian commanders in Egypt during the fourth century B.C. perhaps can be compared to American Special Forces providing training for Iraqi troops fighting ISIS today.

The first of these Athenian commanders was a man named Chabrias.¹² Chabrias had made a name for himself during wars on the Greek mainland.¹³ He also had the important task of overseeing the construction of a series of fortifications on Egypt's northeastern border with the Persian Empire and possibly along Egypt's northern coastline.¹⁴

Two ancient writers, Diodorus Siculus and Plutarch, wrote that Chabrias actually was a mercenary himself.¹⁵ According to them, the Athenian government had no part in arranging his service in Egypt. However, a stele that mentions Chabrias' involvement with an official Athenian *theoria* to the oracle of the god Ammon at the Siwa Oasis in Libya suggest that Chabrias' involvement with Egypt was more complicated than Diodorus or Plutarch realized.¹⁶ The exact date of the *theoria* is uncertain, but it is known that it dates to somewhere in the fourth century B.C. because Chabrias was a member of it. A. M. Woodward supposes that this *theoria* to the oracle of Ammon should be dated to sometime before 360 B.C.¹⁷ Chabrias' visit could have occurred any time between 400 and 360 B.C., but Egyptian expansion into the Libyan oases began as early as 400 B.C. when Pharaoh Amyrtaeus (404-399 B.C.) took control of the Kharga Oasis.¹⁸ The Egyptians undoubtedly controlled Siwa itself by 370's or 360's B.C.¹⁹ Egyptian

¹¹ Diod. XV, 29.1-4 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 396.

¹² Ibid. 396.

¹³ *The Complete Works of Diodorus Siculus*, (Hastings, East Sussex: Delphi Publishing, 2014), Book XIV, 92. 1-2, Kindle Edition. Also, Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire*, (New York, New York: Oxford University Press, 2012), 79.

¹⁴ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 99-100

¹⁵ Diod. XV, 29. 1-4 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 396. Also, Plut. Ages. in *Plutarch: Lives of the Noble Grecians and Romans*, Kindle Edition

¹⁶ A. M. Woodward, "Athens and the Oracle of Ammon," *The Annual of the British School at Athens*, Vol. 57 (1962): <http://www.jstor.org/stable/30104495>. A *theoria* was a religious visit to a sacred precinct officially organized and sanctioned by an ancient state's government. *Theorias* to the Oracle of Delphi were very common. Woodward supposes that this *theoria* to the oracle of Ammon should be dated to the mid fourth century.

¹⁷ Ibid.

¹⁸ Olivier Perdu, "Saites and Persians" in *A Companion to Ancient Egypt* ed. Alan B. Lloyd (Malden, Massachusetts: Blackwell Publishing, 2010), 152-153.

¹⁹ Ibid., 155. Nectanebo I (380-362 B.C.) unquestionably brought Siwa under Egyptian control. Writings on the walls of temples at Siwa attest to Nectanebo I's influence in Siwa.

control of the Siwa Oasis around this time is important because any official Athenian delegation to Siwa likely would have been visiting Egyptian territory. Egypt's kings had exercised a policy of intense colonization and expansion into the Libyan oases during the fourth century B.C., and Siwa was a part of this policy.²⁰ The desire to control the oases of Libya was connected to economic reasons, as abundant trade coming north from the African interior passed through them. The Greek colony of Cyrene was also close to Siwa, and Egypt could use the oasis as a base for trade with Cyrene. Control of the oases also helped defend Egypt's western border from attacks by hostile Libyan tribes.

The Egyptian pharaoh Amasis II (570-526 B.C.) was responsible for building the first temple at the Siwa Oasis at a date preceding the Persian conquest of Egypt.²¹ The Egyptian ruler that Chabrias was said to have served under during the 380's B.C., Achoris (393-380 B.C.), may have restored Egyptian authority there as well.²² The oasis was undoubtedly in Egyptian hands again by 360 B.C. as Egyptian temple building there suggest.²³ Pharaoh Nectanebo II (360-342? B.C.), was the builder of another impressive temple at Siwa.²⁴ Therefore, there was an official Athenian visit to what was probably Egyptian territory sometime in the fourth century B.C. Even if Egypt did not yet control Siwa when this *theoria* was in Libya the Athenians still dispatched a delegation to a location that was culturally tied to Egypt and near its borders. One of the members of this delegation was none other than Chabrias,²⁵ the very individual who ancient writers tell us so often served under Egypt's kings in this same century.

There are other indications of political connections between Egypt and Athens at this time. It was during the fourth century B.C. that the Egyptian cults of Isis and Amun, the Greek "Ammon," became well known in Athens.²⁶ The spread of religious cults from one nation to another often coincided with political relations between two nations. The Thracian goddess, Bendis, became officially recognized in Athens around 430-429 B.C. when the Athenians were

²⁰ B. G. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, *Ancient Egypt: A Social History* 2nd ed. (Cambridge, New York: Cambridge University Press, 2012), 343-345.

²¹ *Ibid.*, 343-345.

²² Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 73-74.

²³ Olivier Perdu, "Saites and Persians" in *A Companion to Ancient Egypt* ed. Alan B. Lloyd (Malden, Massachusetts: Blackwell Publishing, 2010), 155. Pharaoh Nectanebo I (380-362 B.C.) solidified Egyptian control of Siwa. Interestingly, it is close to this time that Woodward suggests the Athenian visit to Siwa occurred.

²⁴ B. G. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, *Ancient Egypt: A Social History* 2nd ed. (Cambridge, New York: Cambridge University Press, 2012), 345.

²⁵ A. M. Woodward, "Athens and the Oracle of Ammon," *The Annual of the British School at Athens*, Vol. 57 (1962): <http://www.jstor.org/stable/30104495>.

²⁶ For Isis, *Greek Historical Inscriptions*, ed(s). P. J. Rhodes and Robin Osborne (Oxford, New York: Oxford University Press, 2003), 462-466 and also Josiah Ober, "Disorder and Growth, 403-340 BCE" in *The Rise and Fall of Classical Greece* (Princeton, New Jersey: Princeton University Press, 2015), Kindle Edition; for Amun, A. M. Woodward, "Athens and the Oracle of Ammon," *The Annual of the British School at Athens*, Vol. 57 (1962): <http://www.jstor.org/stable/30104495>.

engaged in diplomatic relations with the Thracian king Sitalces.²⁷ There are less direct indications of Athenian involvement in Egypt from within Egypt itself. Coinage is attested in the material record for Egypt beginning in the fourth century B.C. The minting of coins in Egypt began around the late 360's under an Egyptian pharaoh named Tachos (362-360 B.C.).²⁸ Interestingly, this same monarch was said to have taken much economic advice from, again, Chabrias.²⁹

Greek influence, though not necessarily Athenian, may be seen in the funerary art of a certain Egyptian high priest from Hermopolis named Petosiris. Petosiris' life dates to the mid to late fourth century B.C., but he initially served as priest during the latter stages of Egypt's period of independence. A relief in his tomb depicts a traditional scene of a wheat harvest and hieroglyphics are written above the relief. Despite these images of Egyptian culture, the way the people are depicted, both in clothing and pose, highly resembles Greek style imagery.³⁰ The poses of the people depicted are far more natural looking than traditional poses for humans in Egyptian art, and the clothing that men are wearing resembles clothing depicted in ancient Greek art. The texture and volume of the environment in the painting also may reflect inspiration from Greek art.

These indications of cultural exchange between Egypt and Greece only begin to appear in the material record beginning in the fourth century B.C. when Egypt was independent of Persian control.³¹ The Greeks, or at least Athens, were adopting Egyptian religious ideas while trading with Egypt in order to benefit from Egypt's spectacular agricultural wealth. On the other hand, the Egyptians were gaining economic and military expertise from the Athenians, both fields being areas where the Athenians excelled, as well as some degree of artistic influence.³² The political connections between Greece and Egypt resulted in remarkable examples of cultural influence in both directions. This cultural interaction was initiated by the political needs of both

²⁷ Deborah Boedeker, "Athenian Religion in the Age of Pericles," in *The Cambridge Companion to the Age of Pericles* ed. Loren J. Samons II (Cambridge, New York: Cambridge University Press, 2007), 49.

²⁸ G. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, *Ancient Egypt: A Social History* 2nd ed. (Cambridge, New York: Cambridge University Press, 2012), 328-329. Also, Karol Mysliwiec, *First Millennium B.C.E.: The Twilight of Ancient Egypt*. trans. David Lorton (Ithaca, New York: Cornell University Press, 2000), 169.

²⁹ Poly. *Strat.* III, 11 (1973): <http://www.attalus.org/translate/polyaenus3.html>

³⁰ Alan B. Lloyd, "The Late Period" in *The Oxford History of Ancient Egypt*, ed. Ian Shaw (New York: Oxford University Press, 2002), 392-393.

³¹ Alan B. Lloyd, "The Late Period" in *The Oxford History of Ancient Egypt*, ed. Ian Shaw (New York: Oxford University Press, 2002) and B. J. Trigger, B. J. Kemp, D. O'Conner, and Alan B. Lloyd, "The Late Period, 664-323 BC" in *Ancient Egypt: A Social History*, 2nd ed. (Cambridge, New York: Cambridge University Press, 2012)

³² Josiah Ober, "Disorder and Growth, 403-340 BCE" in *The Rise and Fall of Classical Greece* (Princeton, New Jersey: Princeton University Press, 2015), Kindle Edition. Athens took measures to expand its commercial interests throughout the Mediterranean world in the fourth century B.C. and open its markets to foreign traders. Furthermore, various Athenian magistrates possessed titles that stressed financial duties during the 300's B.C.. One example of financially oriented Athenian magistrates is the general in charge of "financing the fleet," but the Athenians took various measures to insure the integrity of coins passing through their markets from traders.

Egypt and Athens, especially their mutual desire to diminish Persian power. Such interactions do not appear to be occurring before the fourth century B.C.³³ On the other hand, while Egyptian officials serving under the kings of Persia in former times are attested as cultural and architectural advisors, very few if any Persian cultural traits seem to have assimilated into Egyptian culture.³⁴ Cambyses II (525-522 B.C.) and Darius I (521-486 B.C.) repaired or constructed some Egyptian style temples, but these actions were mainly intended to cater to the preexisting Egyptian culture. There was outspoken Egyptian hatred for Persia, as evident by the depiction of Xerxes I (486-464 B.C.) as a bound captive whenever he is referred to in Egyptian hieroglyphics!

How the Persians Reformed their Frontier Satrapies in Response to Athenian and Egyptian Relations

The policy that the Persians adopted in response to Athenian-Egyptian collusion was to change the way they managed some territories. This reform is the major factor that destabilized the Persian Empire as a result of Greek and Egyptian incursions. Sometime in the late 350's B.C., the Persians appointed an official named Mazaeus to a position that placed Phoenicia under more direct Persian control.³⁵ This was an important development. In former times, Phoenicia was ruled by the individual kings of the Phoenician city-states who acknowledged the Great King of Persia as their overlord. Phoenicia still was incorporated into the Persian satrapal system, but the Phoenician kings ruled with a degree of autonomy in their own cities. A Phoenician king of Sidon was even one of Xerxes' main naval commanders during his campaign against Greece in 480 B.C.³⁶ Sidon itself maintained a position as one of the most greatly favored cities by the Persians throughout most of the Persian Empire's history. However, beginning around the mid to late 360's B.C., Sidon became a frequent ally of Egypt. Contacts between Phoenicia and Egypt are known to have existed even as early as the 380's B.C. due to the discovery of altar bases attesting the name of Pharaoh Achoris (393-389 B.C.) in several Phoenician cities dating to that time.³⁷ Simultaneously, the Egyptians even had contested

³³ The presence of Greeks in Egypt is certainly attested before this time. However, Greek influence on Egyptian art, the existence of Egyptian coinage, and the presence of the cult of Isis in Athens are not present until the fourth century B.C. For a discussion of Greek settlers in Egypt before or during the fourth century B.C. see, Alan B. Lloyd, "The Late Period" in *The Oxford History of Ancient Egypt*, ed. Ian Shaw (New York: Oxford University Press, 2002)

³⁴ Karol Mysliwiec, *First Millennium B.C.E.: The Twilight of Ancient Egypt*. trans. David Lorton (Ithaca, New York: Cornell University Press, 2000), 125-158. It should be noted that a statue of Darius I at Susa shows Egyptian hieroglyphics. However, this is in addition to other languages. Additionally, the statue potentially is modeled after Egyptian style statuary, but the Egyptians do not seem to have shown as much interest in Persian culture.

³⁵ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 158-159. Ruzicka notes the presence of Persian coins minted at Sidon from the late 350's into the 340's, which indicates the position of authority Mazaeus held in Phoenicia.

³⁶ Barry Strauss, *The Battle of Salamis: the Naval Encounter that Saved Greece-and Western Civilization* (New York, New York: Simon and Schuster, 2004), 125-140.

³⁷ Olivier Perdu, "Saïtes and Persians" in *A Companion to Ancient Egypt* ed. Alan B. Lloyd (Malden, Massachusetts: Blackwell Publishing, 2010). 153-154.

control of Phoenicia on several occasions with the Persians.³⁸ Phoenicia's status as a major source of Persian naval strength meant that Persian demands on Phoenicia became greater as warfare became more frequent for the Persians in the fourth century B.C. Unrest coupled with Egyptian support caused many Phoenician cities to seek an end to their ties with Persia.

Persian war demands and Egyptian intervention also created havoc in Persian held Cilicia and Cyprus, two other sites of Persian naval power. Persian administration of the island of Cyprus was the same as in Phoenicia. Semi-autonomous kings of Cypriot city-states who acknowledged the king of Persia as their overlord ruled with a degree of local freedom.³⁹ The Athenian and Egyptian supported expansionism of Evagoras I of Salamis on Cyprus in the 380's B.C. caused immense disorder on Cyprus as well. Persian authority on the island was difficult to maintain.⁴⁰ The Ionian Greek cities, who also sporadically cultivated contacts with Sparta and possibly Egypt,⁴¹ also had no love for their Persian conquerors. The result was that the Persian Empire was losing support in the regions where its naval power could be bolstered and Persian attempts to tighten control of Phoenicia only increased opposition from the Phoenicians. This would be crucial when Alexander invaded the Persian Empire in 334 B.C. The Persian navy was the empire's main advantage against Alexander, but the peoples who comprised Persia's naval force—Egyptians, Cilicians, Cypriots, Phoenicians, and Ionian Greeks—were not at all lovers of the Persians. When Alexander marched into Phoenicia, the greatest source of Persian naval power, only Tyre attempted to resist. All of the other cities quickly submitted to Alexander the moment that Darius III retreated eastwards after Alexander defeated him at Issus in 333 B.C. As will be discussed below, Sidon even began to take action that would help pave Alexander's advance into Phoenicia after Issus, but this hostility towards Persia originated with political links between the Phoenicians and the Greek-supported Egyptian Kingdom.

Another way the Persians readjusted the administration of their empire was to allow a greater degree of autonomy to the satraps of many Anatolian satrapies.⁴² Caria was made an independent province under its own dynasty of rulers, the Hecatomnids, around 392 B.C. who acknowledged the Persian king as their overall sovereign. Ionia was made into its own satrapy

³⁸ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 83-99, 134-145, and 164-177. Sidon seems to have acted as an Egyptian ally on all of the occasions where Egyptian troops were present in Phoenicia.

³⁹ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 68-70.

⁴⁰ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 83-98.

⁴¹ Xen. *Ages.* in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington (Cambridge, Massachusetts: Harvard University Press, 1962), 95-99. Also, Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 136. Ruzicka suggests that the presence of Egyptian representatives with Agesilaus as he sailed away from Ionia indicates Egyptian contacts in Ionia.

⁴² Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 63-65. But also see Diod. XIV, 98. 1-5 in *The Complete Works of Diodorus Siculus* (Hastings, East Sussex: Delphi Publishing, 2014), Kindle Edition. In this passage, Hecatomnus, the Carian dynast/satrap, acts independently but under the direction of Artaxerxes II. This passage provides an example of how at least some satraps were capable of large-scale independent action so long as they had the Great King's approval.

as well. Hereditary rulers also commanded the satrapy of Hellespontine Phrygia, where a certain Pharnabazus' son, Ariobarzanes, was allowed to succeed his father. The satraps were also allowed to maintain some of their own military forces, many of which were Greek mercenaries. The satrapy/protectorate of Caria even maintained its own navy. The idea seems to have been for the Anatolian satrapies to be more efficient at responding to Greek threats and Greek supported uprisings while the Great King focused his own armies against Egypt and Egyptian supported rebellions throughout Palestine, especially at Sidon. However, the satraps of Anatolia proved more ambitious than the Persian kings could anticipate, and the Anatolian satrapies frequently revolted against their Persian overlords, oftentimes with support from Athens, Sparta, and possibly even Egypt.

Egypt continued to benefit from Greek military expertise through covert Athenian support and hired Greek mercenaries. Sparta also was aligned with Egypt by 361 B.C. Warfare in Greece had taken a different turn after Thebes became the dominant Greek city-state following its victory at Leuctra against the Spartans in 371 B.C.⁴³ The Thebans attempted to expand their power, but their ambitions were checked after an extremely difficult victory at Mantinea in 362 B.C. cost the city the life of its most important political and military mind, Epaminondas.⁴⁴ Nevertheless, the surge of Theban power resulted in the Persians shifting their financial support to Thebes, apparently feeling that the Thebans could best achieve Persian interests in Greece.⁴⁵ The Spartans were outraged and resolved to divert Persia's power away from Greece just as much as the Athenians had been doing.

Greco-Egyptian Cooperation in Action: Persian Defeats, Unrest in Anatolia, and the Spartan and Athenian supported Campaign of Pharaoh Tachos

The Persians launched three failed attacks on Egypt during the reigns of Great Kings Artaxerxes II (405-359 B.C) and Artaxerxes III Ochus (359-338 B.C) from 390-387 B.C.: in 373 B.C.: and in 351 B.C.⁴⁶ The Persians suffered heavy casualties on all three occasions and Persian failure was usually followed by Egyptian counter-attacks and Egyptian-inspired revolts. While Persian kings were focused on warfare with Egypt, the satraps of Anatolia became restless. After the failed Persian invasion of Egypt in 373 B.C., many satraps of Anatolia rose up against the Persian king somewhere in the late 360's B.C., but disorder had been brewing there since about as early as 370 B.C.⁴⁷ The increased autonomy of Persia's Anatolian satraps proved to have grave consequences. Sparta and Athens both supported these revolts and Athens seized Samos from Persian control around 366 B.C.⁴⁸ Simultaneously, Pharaoh Tachos (362-360 B.C.)

⁴³ J. B. Bury and A. W. Picard, *The Ancient World: Egypt, Greece, and Persia in the 4th Century BC* (Cambridge: Ozymandias Press, 2016), Kindle Edition.

⁴⁴ Ibid.

⁴⁵ Xen. *Hell.* VII, 1.33-40 in *The Landmark Xenophon's Hellenica* ed. Robert B. Strassler (New York: Pantheon Books, 2009), 279-281.

⁴⁶ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 66-77, 114-122, 154-164. The date of the first invasion is disputed. Ruzicka bases the 390 to 387 B.C. date off of Isocrates' *Panegyricus*.

⁴⁷ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 122-134. The earlier instability was in connection to the revolt of Datames, the satrap of Cappadocia.

⁴⁸ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 134-144 and also Josiah Ober, "Disorder and Growth, 403-340

of Egypt launched a massive invasion of the Persian territories of Palestine and Syria with covert Athenian aid through Chabrias and more explicit aid from Sparta under King Agesilaus II (400-360 B.C.).⁴⁹ Unlike the Athenians, the Spartans were not interested in trying to hide their “distaste” for the Persians by this date. The Persian Empire was in shambles at this point, but intrigue, assassination, and counter-revolts quelled the threats of the Egyptian pharaoh and the Anatolian satraps by 360 B.C.⁵⁰ Pharaoh Tachos himself was dethroned due to a military coup ultimately led by the future king of Egypt Nectanebo II. The Persians had dodged a bullet, but the simultaneous threats of the Egyptian king, the Greek states of Athens and Sparta, and the disloyal Anatolian satraps in the 360’s B.C. would leave a lasting impression on Persian foreign policy. From this point on, the Persians dedicated themselves to neutralizing the threat of Egypt at any cost and removing the threat of Greek intervention in their empire by inspiring political upheavals for Athens and Sparta. The Persians succeeded in checking Athenian and Spartan interference in Egypt or Anatolia. This was largely due to Persian financial support for Theban warfare against Athens, Sparta, and other Greek states.⁵¹ By helping Thebes, the main rival of Sparta and Athens on the Greek mainland by the 360’s B.C., to threaten the rest of Greece, the Persians insured that Athenian and Spartan attention and military forces could not interfere in Anatolia or Egypt. However, by weakening the major powers of the Greek world through the funding of Theban militarism, the Persians created a power vacuum in Greece. The Persians would not fill it themselves because they were too devoted to suppressing Egypt. The Persian Empire’s decision to invest their military forces in their war against Egypt and only their financial resources in Greece (by funding Thebes) created a power vacuum in Greece. This power vacuum formed because no Greek state possessed the power to become supreme in Greece. Therefore, Greece remained in a state of constant, inconclusive warfare that exhausted the major Greek city-states. Persian funding of this endless warfare resulted from Persian efforts to isolate Egypt from Athens and Sparta. However, the power vacuum created in Greece by Persian interference allowed Macedonia to rise to prominence beginning in 359 B.C.

Persian Reaction to Greek Intervention and Tachos’ Failed Invasion and Persian Efforts to Create a Power Vacuum in Greece

Anatolia remained difficult territory. Great King of Persia Artaxerxes III Ochus tried to restrain the autonomy of the satraps there by commanding the dismissal of their satrapal armies, but this triggered another revolt by Artabazus, the satrap of Hellespontine Phrygia, in the 350’s B.C.⁵² Athens and even the Greek state of Thebes supported this uprising, but through threats and diplomacy Artaxerxes III Ochus managed to isolate Artabazus from his Greek allies. This

BCE” in *The Rise and Fall of Classical Greece* (Princeton, New Jersey: Princeton University Press, 2015), Kindle Edition

⁴⁹ Plut. *Ages.* in *Plutarch: Lives of the Noble Grecians and Romans*, Kindle Edition and also Diod. XV, 92.2-5 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 401.

⁵⁰ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 122-145.

⁵¹ Diod. XVI, 40 in *The Complete Works of Diodorus Siculus* (Hastings, East Sussex: Delphi Publishing, 2014), Kindle Edition.

⁵² Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 154-158.

revolt was also quashed.⁵³ However, the renewed threat of Athenian intervention may have forced Artaxerxes III to permit the dynast/satrap of Caria, Mausolus, to maintain his military forces and semi-independent status. Despite Artaxerxes III's attempts to restrain the satraps, Caria certainly remained heavily militarized during and after the revolt of Artabazus. Other satraps might have retained their semi-autonomous status as well.

In the interest of Persia and Caria, Mausolus then provided aid for several Greek city-states rebelling against Athenian expansionism in the Social War (357-355 B.C.). Breaking Athenian power was essential to Persian plans to limit Greek involvement in their empire and conquer Egypt. The Athenians were defeated, but this defeat had dire consequences for all involved—in 357 B.C., Phillip II of Macedon conquered the city of Amphipolis, despite Athenian attempts to provide support for Amphipolis. By 354 B.C. Phillip also had taken Pydna and Methone, both of which were aligned with Athens and provided military bases for Athenian troops. The Social War so sapped Athens' economy that the Athenians were no longer capable of opposing Macedonian expansionism until 352 B.C.⁵⁴ Thus, Phillip of Macedon began to swallow up a world worn down by warfare between the Greeks, Egyptians, and Persians. However, Phillip's conquests may never even have begun had the Athenians not been locked in a standoff with the Persians as they contributed forces to Egypt and the Anatolian satrapies. Likewise, the Persians themselves probably would not have allowed Phillip to become so strong were they not struggling against Egypt and their Athenian and Spartan supported rebellious satrapies. The previously mentioned examples of Persian support for Sparta during the Peloponnesian War and Corinthian War, Thebes beginning in 366 B.C., and Athens' enemies in the Social War certainly indicate that the Persians were mindful of Greek affairs. However, Athens was no longer strong enough to oppose Macedon and the Persian Empire was too focused on retaking Egypt to counter Macedonian expansionism. The Carians, while strong enough to support Athens' enemies in the Social War, did not possess the military resources to fill in the power vacuum created by their Persian overlords. Only the Great King of Persia could take control of the situation in Greece, but he was ensuring that Egypt was removed as a major power in the Near East. Persian attempts to create a power vacuum in Greece that would prevent Spartan or Athenian aid to Egypt or the Anatolian satrapies paved the way for the rise of Macedonia.

After 360 B.C., the Persians established the already noted tighter control of Phoenicia during the reign of Artaxerxes III. This was part of their strategy to finally bring the war with

⁵³ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 154-158. Also Josiah Ober, "Disorder and Growth, 403-340 BCE" in *The Rise and Fall of Classical Greece* (Princeton, New Jersey: Princeton University Press, 2015), Kindle Edition. Also Ian Worthington, "Greece and Macedonia" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition.

⁵⁴ Ian Worthington, "Philip II and the Rise of Macedonia" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition.

Egypt to an end. However, Persia's invasion in 351 B.C. was repelled by Pharaoh Nectanebo II, who received aid from Spartan and Athenian commanders.⁵⁵

Persia's Own Creation: the Rise of the Macedonian Empire and the Persian Conquest of Egypt

Everything changed by 346 B.C. By this date, Phillip II of Macedon had ascended over Greece as the most powerful warlord in the land. Bold new military innovations and rich stores of precious metals and lumber powered the Macedonian state on a scale not seen before in the Greek world. Athens feared the growing power of the Macedonians, even more than she had feared the Persians in the past. But previous wars had left Athens weakened. This resulted in a drastic change of foreign policy. The Athenians would no longer provide military aid—covert or otherwise—to Egypt. By 341 B.C. the famous Athenian orator Demosthenes had gone as far as to advocate allying with the Persians against the Macedonians.⁵⁶ When Phillip began expanding into Thrace and towards the Hellespont around 342 B.C.,⁵⁷ Athenian fears grew all the more. The Hellespont, the trade route that was so vital for Athenian grain importation, was again threatened, but not by Persia. However, an alliance with Persia presented a problem; the Persians would never commit large-scale forces to checking Macedonian power unless their war with Egypt had been brought to a successful conclusion. The Persians later would learn that their focus on Egypt and negligence towards Macedonia was a mistake.

Knowing that the Persians would not seriously commit to checking Macedonian expansionism while Egypt remained independent, an Athenian commander named Phocion acted in Persian service during the campaigns of Artaxerxes III Ochus to put down another round of Egyptian supported revolts in Cyprus and Phoenicia in 345-344 B.C.⁵⁸ These revolts came to an end with the extremely violent sacking of Sidon in 344 B.C.⁵⁹ Harsh Persian treatment of many Phoenicians would have long lasting repercussions when most of the Phoenician cities willingly deserted Darius III of Persia after the Battle of Issus in 333 B.C. It is remarkable how this city, once prized by the Persians, was a frequent ally of Greek supported Egypt and an enemy of the Persian Empire by the middle of the fourth century B.C.

Diodorus Siculus indicates that the Persians finally retook Egypt by about 343 B.C., but Egyptian records and Manetho indicate that the reigning pharaoh of Egypt at this time,

⁵⁵ Diod. XVI, 48 in *The Complete Works of Diodorus Siculus* (Hastings, East Sussex: Delphi Publishing, 2014), Kindle Edition. Diophantus of Athens and Lamius of Sparta are said to have been in Egypt serving under Nectanebo II.

⁵⁶ Dem. Dem. 10 in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington (Cambridge, Massachusetts: Harvard University Press, 1962), 287-293.

⁵⁷ Ian Worthington, "The Gathering War Clouds" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition.

⁵⁸ Diod. XVI, 42 in *The Complete Works of Diodorus Siculus* (Hastings, East Sussex: Delphi Publishing, 2014), Kindle Edition.

⁵⁹ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 164-176.

Nectanebo II, continued to hold at least Upper (southern) Egypt into 341 B.C.⁶⁰ Interestingly, Diodorus does not say that Artaxerxes III campaigned into southern Egypt.⁶¹ Instead, as the Athenians probably had hoped, Persian forces were fighting in Thrace and around the Hellespont in 340 B.C. against Phillip II of Macedon.⁶² Persia's military priority had changed. Egypt was cut off from the coast and from contact with the Greeks. The confinement of any hostile Egyptian forces to Upper Egypt theoretically prevented the Egyptians from instigating any more revolts within the Persian Empire. Persian control of Lower (northern) Egypt appears to have been satisfactory for Artaxerxes III when other major threats were looming to the north. If Macedonia became the dominant power in Greece, then it could present a serious threat to Persia. Phillip might use the united force of the Greek states under his command to threaten Persian power in Anatolia. However, Persian opposition to Macedonia had come too late. By 340 B.C. the Kingdom of Macedonia was a powerful force in Greece.

The Breaking Point: How Persian Preoccupation with Egypt Aided in Alexander the Great's Victory over the Persian Empire

The year 338 B.C. saw a series of simultaneous problems that combined to cause paralysis and chaos in the Persian Empire that paved the way for renewed problems with Egypt and even the final victory of Phillip II over the Greeks. The chief official of Artaxerxes III, Bagoas, assassinated the Great King of Persia and placed Arses on the Persian throne as Artaxerxes IV.⁶³ Artaxerxes IV was merely a puppet of Bagoas and the Persian Empire once again plummeted into chaos. Meanwhile, despite Demosthenes of Athens' rhetoric and the vigorous efforts of the Persians and Athenians to curb Macedonian power in 340 B.C., Phillip II won an earth shattering victory against his Greek opponents at Chaeronea. After Chaeronea, Greece fell under the shadow of the Kingdom of Macedon and Phillip II now held power over all of the Greek city-states with the exception of Sparta. Under his leadership, the Greeks could be led in a united front against the Persians, something Phillip proclaimed clear intentions of doing in 337 B.C.

Another threat to the Persian Empire was emerging in 338 B.C. Artaxerxes III's apparent decision to leave Upper Egypt in Egyptian hands now proved to be a mistake. A new pharaoh had risen to power there named Chababash in 338 B.C.⁶⁴ Sometime before 336 B.C. he had driven the Persians completely out of Egypt once again. Egypt was once again independent. The court corruption created by Bagoas prevented the Persians from acting against either of these threats initially, so both the Macedonian king and the Egyptian king were allowed to consolidate their positions.

⁶⁰ David Warburton, Rolf Krauss, and Erik Hornung, *Ancient Egyptian Chronology* (Leiden: Brill, 2006), 268-271. Also Manetho in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington (Cambridge, Massachusetts: Harvard University Press, 1962), 182-185.

⁶¹ *Ibid.*, 177-199.

⁶² Ian Worthington, "The Gathering War Clouds" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition.

⁶³ Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York: Cambridge University Press, 2014), 197-199.

⁶⁴ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 199-210.

The Persian Empire was once again functional upon the accession of Darius III and the assassination of Bagoas in 336 B.C. Darius III was about to make a decision that would influence the course of history. The Persian Empire was seriously threatened from the northwest and southwest. If Chababash was allowed to retain control of all of Egypt, then the long history of warfare between the Persians and the Egyptians might be repeated. Endless Persian onslaughts against Egypt over the past several decades indicate that the Persian Empire's primary foreign focus was Egypt in the fourth century B.C. The Macedonian domination of Greece was a grave threat as well and Phillip II compounded Persia's Macedonian concern by dispatching a large army across the Hellespont into Anatolia with the intentions of establishing a bridgehead into the Persian Empire and pulling the Ionian Greek cities into his fold.⁶⁵ Darius knew of the Macedonian king's ferocity and he began assembling an army to confront Phillip's advance force.⁶⁶ But Phillip himself was assassinated in 336 B.C. and his son and successor, Alexander III, struggled to deal with multiple serious revolts in the Macedonian Empire throughout 336-335 B.C.⁶⁷

Darius noticed the young age of Alexander and the series of revolts that threatened the Macedonian Empire and he apparently dismissed Alexander as a minor threat. Seemingly, Greece had returned to its pre-Macedonian state: a power vacuum that the Persians would not fill because they were devoted to conquering Egypt. Darius III redeployed the army that had been preparing to confront Macedonia to Egypt to launch yet another Persian invasion of that country.⁶⁸ Persian desires to suppress Egypt once again prevailed. Darius III's decision to redirect Persian military assets away from Anatolia to Egypt during the beginning of Alexander's reign would decisively alter the course of human history. These same forces would no longer be available to drive back the Macedonian forces that Phillip II had dispatched into Anatolia just before his assassination.

Egypt was defeated, though it is uncertain how far reaching Darius III's control of Egypt was considering that Artaxerxes III had stopped short of subjugating the entire country a mere seven years ago. On the other hand, Persia's once again altered focus—from Egypt to Macedonia in about 340 B.C. and now back from Macedonia to Egypt in 336 B.C.—had dire consequences. The Persians failed to contribute substantial forces to dislodging Macedonian holdings along the Anatolian coast taken by Phillip II in 336 B.C.⁶⁹

Persian forces certainly could have returned from Egypt by 334 B.C. Despite the neutralization of Egypt, when Alexander III "the Great" began his invasion into the Persian Empire, the Persians had not been able to dislodge Macedonian holdings along Anatolia's northern coast, which Macedonia had taken while the main Persian army was fighting against

⁶⁵ Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York: Cambridge University Press, 2014), 199-201.

⁶⁶ Diod. XVII, 7.1-3, 8-10 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 428.

⁶⁷ Ian Worthington, "Alexander's Early Kingship-and Persia" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition.

⁶⁸ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 204.

⁶⁹ Edward M. Anson, "The Persian Fleet in 334" in *Classical Philology* Vol. 84, No. 1 (January 1989): <http://www.jstor.org/stable/270044>

Chababash in Egypt. The Persians been attempting to push the Macedonians back from Anatolia with a small force at least partially under the command of Memnon of Rhodes, but their Anatolian forces were insufficient for gaining back the northern coasts. This is interesting since Diodorus indicates that Darius III had assembled a massive force to deal with Macedonia.⁷⁰ After Phillip's death in 336 B.C., this force seems to have been redirected to Egypt, leaving only a minimal Persian military presence in Anatolia at least until 334 B.C. In that same year, with the beachhead in northern Anatolia still held by the Macedonian vanguard, Alexander crossed into Anatolia. That same year he defeated the armies of the local satraps at the Granicus River. The Persian navy would have had time to return from the Egyptian campaign of 336 B.C. in order to prevent Alexander's crossing, and if it had done this, Alexander almost certainly would have been driven back because he did not possess a fleet that could even hope to challenge the Persian navy in a direct battle. However, Alexander was familiar with the coastal territory that Phillip had conquered, which the Persians had failed to retake due to their commitment to suppressing Egypt. Alexander's strategy for conquering Anatolia due to his knowledge of Philip's conquests and Persia's failure to counter Phillip's successes proved fruitful for countering the Persian navy. Phillip's conquests extended too far inland in order for the Persian navy to safely moor anywhere close enough to prevent Alexander's crossing.⁷¹ Alexander knew that Persia's navy could not oppose him because he controlled the coast, so he relied on his land armies knowing that when he reached Anatolia the superior Macedonian army could carry the campaign.

Not that this strategy was without problems. Despite Alexander's successful crossing into Anatolia, the Persian navy launched several attacks on the Macedonian rear and attempted to cut Alexander's supply lines with a degree of success from 334-332 B.C. Alexander also faced fierce resistance from pro-Persian and Persian forces at Miletus and Halicarnassus that were aided by Persian naval contingents.⁷² The Persians even used their navy to encourage the Spartan king Agis III (338-331 B.C.) to begin a new war against Macedonia in Greece. Persia's strategy of playing Greek states off of each other had worked in the past. From the Persians' perspective, history provided no reason to doubt whether or not funding Spartan war efforts would hinder Macedonian expansionism. However, they failed to realize how powerful Macedonia had become while they were preoccupied with conquering Egypt.

They soon understood, but first they learned once again why autonomous satraps could be harmful to the empire as a whole. Royal dynasties often are prone to dynastic warfare. One such dynastic civil war had broken out in Caria, a semi-independent satrapy that had been a major source of Persian military power in Anatolia since roughly 392 B.C. Alexander took advantage of this conflict. He backed one of the rival claimants to the throne, a woman who was named Ada who wielded a great deal of influence over the common Carian people and established her as a puppet ruler in Caria.⁷³ The dynastic quarreling in Caria may have

⁷⁰ Diod. XVII, 7.1-3, 8-10 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 428.

⁷¹ Edward M. Anson, "The Persian Fleet in 334" in *Classical Philology* Vol. 84, No. 1 (January 1989): <http://www.jstor.org/stable/270044>

⁷² Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York: Cambridge University Press, 2014), 204-209.

⁷³ Diod. XVI, 24 in *The Complete Works of Diodorus Siculus* (Hastings, East Sussex: Delphi Publishing, 2014), Kindle Edition.

handicapped the current Carian dynast's ability to properly conduct war against Persia's enemies as previous Carian rulers/satrap had done in 390 and 345 B.C. against Cyprus and against Athens in 357-355 B.C. Not that the Carians were likely to present a serious threat to Alexander on their own, but they were capable of greatly slowing his consolidation of Anatolia, as the siege of Halicarnassus indicates. Furthermore, the Carian navy could have been used to counter Alexander's crossing into Asia. But Caria's navy did not hinder Alexander. Additionally, if Alexander had been forced to contest the whole of Caria, it may have taken him substantially longer to push eastwards. It is significant that this powerful Persian satrapy was also where Halicarnassus, which has already been noted as a site that fiercely resisted Alexander, was located. Alexander's installation of Ada as the ruler of the Carians not only ended a dynastic feud, but it also firmly placed Caria under the control of a pro-Macedonian ruler who was popular with the people and could keep the Carians in subjection to Alexander (with the exception of Halicarnassus, where pro-Persian forces were holding out). The dynastic feuding of the Carian royal house greatly aided Alexander's ability to conquer and pacify this once heavily militarized semi-independent Persian satrapy. It should be remembered that the semi-independent status of Caria resulted from the Persians' strategy of allowing more autonomy to the western Anatolian satraps while the Great King himself focused on Egypt. However, dynastic warfare in Caria and court intrigue in the palace of the Persian Great King may have utterly limited the Persian Empire's ability to respond to Phillip II as he consolidated his position in Greece. Likewise, Chababash's success in Egypt in the 330's B.C. may be due in part to the same internal struggles of the Persian Empire.

Despite continued resistance at Halicarnassus, Alexander charged his subordinates with conquering the city while he himself continued campaigning eastwards. Alexander defeated the army of Great King Darius III at Issus in 333 B.C. Now the consequences of Egyptian and Greek destabilizing efforts had one final part to play. It has been noted that the Persians had lost support in many coastal territories, Phoenicia in particular. This was partially due to increased contacts between these regions and an Athenian and Spartan backed Egypt. After Issus, Darius III dispatched a Macedonian mercenary commander named Amyntas to secure Persian holdings from Phoenicia down to troublesome Egypt.⁷⁴ He was accompanied by 4000 troops according to Diodorus Siculus—the same number of troops that Arrian says escaped from Issus with Darius III.⁷⁵ This fact suggests that Darius III gave Amyntas the remainder of the army and dispatched him to the south. The defeat at Issus would surely encourage new revolts. In addition to this tension, the Persian satrap of Egypt, Sabaces, had been killed at Issus. Sabaces' death likely caused confusion for Persian authorities in Egypt. The Persian king knew that maintaining the empire's coastal territories from Phoenicia to Egypt was vital because the Persian navy was one of the only great advantages Darius III still had over Alexander. This is why it was vital for Amyntas to maintain Persian power in these regions while the Great King himself retreated eastwards into the Persian heartland to assemble a new land army. Unfortunately for Darius III the coastal regions that he needed to maintain control of also were the same regions affected by warfare between the Persian Empire and Egypt. The backlash towards Persian policies was especially strong in Phoenicia and Egypt, both of which had suffered Persian attacks roughly

⁷⁴ Diod. XVII, 48. 2-4 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 436.

⁷⁵ Arr. *Anab.* II, 13.1 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 436.

within the previous decade. Amyntas abandoned any hopes of holding Phoenicia as Alexander was already fast approaching and most of the Phoenicians had no love for the Persians at this point. Diodorus Siculus says that the rebellious city of Sidon actually invited Alexander into Phoenicia.⁷⁶ Therefore, it may be safe to assume that Sidon had proclaimed all ties with Persia to be broken. The Persians had clearly lost much support here. The sack of Sidon in 344 B.C. following an Egyptian supported uprising in Phoenicia certainly was still fresh on many Sidonians' minds. Amyntas departed from Phoenicia with as many ships and troops as he could muster in 333-332 B.C., just before Alexander arrived in Phoenicia. He then moved on to Cyprus. He mustered more Persian forces there and then moved on to Egypt. Amyntas appears to have given up any hope of holding Palestine against Alexander's advance, but he does seem to have hoped to fortify Egypt with as many Persian troops as he could gather and use the country's natural defenses to block Macedonian expansion into Egypt. If Egypt remained in Persian hands, Alexander might not be able to move east against Darius III. Egypt could function as a naval base, in the eyes of the Persians, and from there Amyntas could raid Macedonian supply lines. Egypt could not be left as a threat to Alexander from behind while he marched on the Persian heartland. After all, Sparta was still at war with Macedonia. Alexander's communications with the home front could not be severed. Amyntas could hold Egypt until the Great King returned with a new army to confront Alexander again. The stout resistance of Tyre and Gaza to Alexander's advance indicates that Darius had not simply abandoned his western possessions.

But Darius' strategy had an unforeseen obstacle. The Egyptians had indeed revolted after Issus. The Persian official who had taken charge in the wake of Sabaces' demise, Mazaces, was unable to suppress the situation as Sabaces had departed from Egypt for Issus in 333 B.C. with the majority of the Persian forces who had been garrisoning the country after Persia retook it in 336 B.C. Amyntas attempted to take the city of Memphis, the heart of Egypt, which controlled all movement up and down the Nile River between Lower and Upper Egypt. But the Egyptians resisted Amyntas and killed him and the Persian relief force that he had brought to hold down the country for Darius III. Thus, Egypt was in a state of serious revolt against Darius III when Alexander arrived there later in 332 B.C.

Diodorus Siculus recounts the Amyntas episode. He indicates that Amyntas' actions were completely on behalf of Darius III and not those of a rogue commander seeking personal gain after the Persian defeat at Issus. However, he gives very vague details concerning who it was that opposed him when he tried to take control of Memphis. The Roman historian known as Quintus Curtius Rufus provides clear indications that the Egyptians themselves had revolted after Darius III's defeat at Issus.⁷⁷ Egypt was for all practical purposes out of Persian hands before Alexander even arrived. Now Alexander delivered the final blow to the Persians. While Tyre resisted staunchly, the rest of Phoenicia and its naval contingents at home or abroad harassing Macedonian supply lines near Anatolia, willingly defected from the Persian navy. Arrian gives a

⁷⁶ Arr. *Anab.* II, 13.7-8; 15.6 4 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 439.

⁷⁷ Quintus Curtius Rufus, IV, 7. 2-5 in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington (Cambridge, Massachusetts: Harvard University Press, 1971), 225-227. It should be noted that Rufus believed Amyntas to be a rogue commander seeking his own gain. This contradicts Diodorus' claim that he acted in Persian service, but I have chosen to believe Diodorus based upon what I have noticed concerning Amyntas' actions as he travelled to Egypt as described in this essay.

clear account of this series of events.⁷⁸ The heavy-handed policies of Artaxerxes II and Artaxerxes III Ochus came back to haunt Darius III. Cyprus, another troublesome possession as well as a source of naval strength for the Persians, also willingly defected to Alexander. The Macedonian king entered Egypt, the most troublesome Persian possession of all, in 332 B.C. to the shouts and applause of the Egyptians according to ancient historians. Alexander's efforts to fit the role of an Egyptian pharaoh, such as burying an Apis bull, may have stemmed from his need to appease a tense political situation. The large garrison that remained in Egypt after Alexander departed from Egypt in 331 B.C. may also have been intended to reestablish public order in the name of the new Macedonian pharaoh.⁷⁹ All of the Persian Empire's coastal possessions were now lost. Persian naval supremacy and the possibility of forcing Alexander to fall back by cutting off his communications with Greece and Macedonia were now things of the past. Alexander the Great continued eastwards in order to finish his war with Persia. The culminating encounter was at Gaugemela in late 331 B.C. where Alexander shattered Persian military strength once and for all. The great Macedonian conqueror of the known world had achieved the impossible. The military and economic system created by Phillip II and capitalized on by Alexander the Great was a phenomenal success.⁸⁰ Phillip's careful coordination of cavalry and infantry, the reforming of the phalanx by introducing the *sarrissa* to the infantry (among other infantry reforms), the formation of a highly trained professional standing army, and the establishment of a sound economy had been the most obvious and perhaps important factors in the rise of Macedonia. Under Alexander's unequalled generalship, Macedonia's powerful economy and military proved decisive in Persia's final defeat.

Summary: How Athens, Sparta, and Egypt Mutually Created Circumstances Favorable to the Fall of the Persian Empire

The instability caused in the Persian Empire by Egyptian and Greek actions preceded and contributed to Alexander's final victory. With the support of Athens and Sparta, Egypt became a threat that persuaded the Persians to focus their military efforts on Egypt throughout the fourth century B.C. Persian efforts to counter Greek incursions into Anatolia by allowing the satraps a degree of autonomy made the satrapal revolts of the 360's B.C. possible. Even the loyal satraps only had enough strength to check any Greek state that became strong. They did not possess the strength to fill the power vacuum in Greece that the kings of Persia created. The lack of a dominant power in Greece aided Macedonia in its rise to power.

Heavy-handed policies towards the Phoenicians and perhaps also the Cypriots in order to curb growing Egyptian influence there only agitated the situation further. The culminating sack of Sidon in 344 B.C. so destroyed Persian support in Phoenicia that most of the Phoenicians were loathe to support Darius III in 333-332 B.C. (though certainly Tyre's devotion to Darius III is notable). Finally, the simultaneous threats of Phillip II of Macedon, his son Alexander, and Chababash' rise to power in Egypt could not be met by the Persians at the same time. The Persian Empire's inability to act properly against either one in 338 B.C. due to court intrigue

⁷⁸ Arr. *Anab.* II, 20.1-3 in *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, ed. Amelie Kuhrt (New York, New York: Routledge, 2010), 442-443.

⁷⁹ Stephen Ruzicka, *Trouble in the West: Egypt and the Persian Empire* (New York, New York: Oxford University Press, 2012), 199-209.

⁸⁰ Ian Worthington, "Philip and the Rise of Macedonia" in *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire* (Oxford, New York: Oxford University Press, 2014), Kindle Edition

permitted both to consolidate their power. When the Persians did react against Chababash in 336 B.C., it was at the expense of their territory on the northern coast of Anatolia. Darius III practiced the typical anti-Egyptian focus of Persia's kings during the fourth century B.C.

The Persians developed this focus due to the power of a Greek-supported Egypt in the fourth century B.C., but their undying commitment to this focus in the 330's B.C. became their undoing. By allowing the Macedonians to maintain their control of part of Anatolia, the Persians permitted the Macedonians to maintain their bridgehead when Alexander the Great crossed into Asia in 334 B.C. From there, he campaigned against an empire that consisted of subjects worn down by constant Persian war demands and reprisals for anti-Persian revolts that had been connected to the Persian Empire's attempts to contain incursions and growing influence from Athens, Sparta, Egypt, or rebellious satraps or cities that were also colluding with Persia's Egyptian and Greek enemies. Most of these same subjects were not willing to fight against Alexander for their Persian oppressor. Darius III found enemies within some regions of his empire as easily as he did within the Macedonian army.

The point of this paper is not to point out the faults of the Persian Empire. The Persians, in reality, had established a sophisticated system of governance that from an organizational or administrative perspective was probably far superior to anything else in the ancient world besides the system utilized by the Romans later on—hence the immense size and longevity of the Persian Empire.⁸¹ The Persians initially exercised a policy of carefully respecting and integrating themselves into the cultures of the peoples subjected to them.⁸² The Persians also created a remarkable empire on a scale that no previous power even could have imagined. The power of the Great King to project his might across most of the known world at the time was unsurpassed by anything in the ancient world besides the Roman Empire. However, in the fourth century B.C., the Persians faced the unprecedented challenge of having to contest their power not only with the Greek city-states, but also with an Egyptian kingdom that quickly reestablished itself as a force to be reckoned with in the ancient Near East. Egypt's vast natural resources coupled with well-timed revolts in the Persian Empire and wars against the Greeks allowed Egypt to become a major political player during this time in history. Spartan and Athenian support furthermore helped Egypt insure its position as an independent power in the fourth century B.C.

Historically, the power of an Egyptian pharaoh was manifested in his building record and Egypt's fourth century B.C. kings, above all Pharaoh Nectanebo II, constructed monumental and temple buildings on a scale that was witnessed in all of Egypt's major cities. His constructions even surpassed many of Egypt's imperial rulers from its New Kingdom days in the late second millennium B.C. Nectanebo II's construction work, which epitomized the architectural programs of Egypt's kings from this period, is attested at countless sites throughout Egypt:⁸³ the capital of Egypt during his reign of Sebennyus, Behbeit el-Hagar where a massive temple to Isis was built, Busiris, Athribis, Tanis, Pharaithos, Bubastis, Saft el-Henna, Tell el-Maskhuta,, Heliopolis, Memphis, Abusir el-Meleq, Herakleopolis Magna, Nebwy-Bilifia, Hermopolis Magna, Abydos, Koptos, Thebes, Hermonthis, el-Kab, Edfu, Elephantine-Aswan, Hibis at the Kharga Oasis, and

⁸¹ Matt Waters, *A Concise History of the Achaemenid Empire, 550-330 BCE* (New York: Cambridge University Press, 2014), 92-113.

⁸² *Ibid.*, 35-49, 142-145.

⁸³ Olivier Perdu, "Saïtes and Persians" in *A Companion to Ancient Egypt* ed. Alan B. Lloyd (Malden, Massachusetts: Blackwell Publishing, 2010), 153-157.

the Siwa Oasis where the oracle to Zeus Ammon was located. Construction at these sites bear witness to a massive building program not seen in Egypt for many centuries. Nectanebo II's building program is important because it shows that Egypt had once again emerged as a major power in the Near East even if such a phenomenon was partially by default due to events happening elsewhere in the world. Egypt's reawakening and its connections with Greek states were major obstacles for Persian imperialism in the fourth century B.C.

Egypt's commercial influence can be seen in the discoveries at the recently excavated Egyptian coastal city of Thonis-Heracleion.⁸⁴ This newly discovered city also attests to the intense contact that occurred between the Egyptian and Greek worlds at this time in history.

It was through alliances with Athens and Sparta that Egypt maintained its independence by means of unparalleled Athenian and Spartan military expertise and often outright military support. However, when Athens began to fear Macedonia more than it had feared the Persians and Sparta fell into a state of decline following military setbacks against Thebes, Egypt lost its valuable Greek allies. The result was another Persian conquest of Egypt between 343-341 B.C., but the threat of Macedonia and Egyptian hatred of the Persians meant that Egypt could never again be stabilized as a Persian satrapy. Persian forces remained poised to counter any new disturbances in Egypt, but Bagoas' treachery in 338 B.C. disrupted any initial attempts to contain the power of Chababash. The result was another round of warfare between the Persians and Egyptians in 336 B.C. that drew Darius III's attention away from the increasingly unstable situation on the empire's northwestern borders.

Finally, after seeing the success of Alexander the Great at Issus, the Egyptians again revolted, contributing to the collapse of Persia's navy, the only military force that brought meaningful success for Darius III. The Persian Empire that had operated off of so efficient a system in the past was forced to respond to new threats it had not known previously and it was frequently fighting on two fronts. New threats persuaded the Persians to adopt new defensive policies. However, whether it provided more or less autonomy to its subjects, both policies failed as the Greeks and Egyptians exploited either resentful and oppressed Persian subjects or opportunistic Anatolian satraps. The Persians became far more militaristic and they zeroed in on Egypt, the primary source for all anti-Persian activity. Meanwhile, the Persians focused on isolating Athens and Sparta from Egypt or the western Anatolian satrapies by creating a power vacuum in Greece. The results were fatal. Egypt was the primary concern no matter what else happened in the world and Alexander the Great took advantage of this in 334 B.C. when he began his invasion of the Persian Empire at the head of an army that quite simply outclassed any other military force in existence at that time. However, this military force marched over the remains of a Persian Empire that was worn down by conflict with Athens, Sparta, and Egypt. Not necessarily every Persian satrapy was opposed to Persian rule, but Cypriot cities, most of Phoenicia, and Egypt were more than ready to side with any enemy of Persia. Alexander the Great found many coastal subjects of Persia eager for freedom and vengeance against Persia.

⁸⁴ *Thonis-Heracleion in Context*, ed(s). Damian Robinson and Franck Goddio (Oxford, England: Oxford Book Projects, 2015). This book discusses some of the most recently published findings from Thonis-Heracleion.

Bibliography

- Anson, Edward M. "The Persian Fleet in 334" in *Classical Philology* Vol. 84, No. 1 January 1989: <http://www.jstor.org/stable/270044>.
- Boedeker, Deborah. "Athenian Religion in the Age of Pericles," in *The Cambridge Companion to the Age of Pericles* ed. Loren J. Samons II. Cambridge, New York: Cambridge University Press, 2007.
- Bury, J. B. and Picard, A. W. *The Ancient World: Egypt, Greece, and Persia in the 4th Century BC* Cambridge: Ozymandias Press, 2016, Kindle Edition.
- Dem. Dem. 10 in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington. Cambridge, Massachusetts: Harvard University Press, 1962.
- Kuhr, Amelie, ed. *The Persian Empire: A Corpus of Sources from the Achaemenid Period*, New York, New York: Routledge, 2010.
- Lloyd, Alan B. "The Late Period" in *The Oxford History of Ancient Egypt*, ed. Ian Shaw New York: Oxford University Press, 2002.
- Manetho in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington. Cambridge, Massachusetts: Harvard University Press, 1962.
- Mysliwiec, Karol. *First Millennium B.C.E.: The Twilight of Ancient Egypt*. Translated by David Lorton. Ithaca, New York: Cornell University Press, 2000.
- Ober, Josiah. *The Rise and Fall of Classical Greece*. Princeton, New Jersey: Princeton University Press, 2015, Kindle Edition.
- Perdu, Olivier. "Saïtes and Persians" in *A Companion to Ancient Egypt* ed. Alan B. Lloyd Malden, Massachusetts: Blackwell Publishing, 2010. *Plut. Ages*. in *Plutarch: Lives of the Noble Grecians and Romans*, Kindle Edition. *Poly. Strat.* III, 11 (1973): <http://www.attalus.org/translate/polyaenus3.html>
- Quintus Curtius Rufus, IV, 7. 2-5 in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington. Cambridge, Massachusetts: Harvard University Press, 1971. Rhodes, P. J. and Osborne, Robin, ed(s). *Greek Historical Inscriptions*. Oxford, New York: Oxford University Press, 2003.
- Ruzicka, Stephen. *Trouble in the West: Egypt and the Persian Empire*. New York, New York: Oxford University Press, 2012.
- Strauss, Strauss. *The Battle of Salamis: the Naval Encounter that Saved Greece-and Western Civilization*. New York, New York: Simon and Schuster, 2004.
- The Complete Works of Diodorus Siculus*. Hastings, East Sussex: Delphi Publishing, 2014, Book XIV, 92. 1-2, Kindle Edition.
- Thonis-Heracleion in Context*, ed(s). Damian Robinson and Franck Goddio. Oxford, England: Oxford Book projects, 2015.
- Trigger, B. J., Kemp, B. J., O'Conner, D. and Lloyd, Alan B. "The Late Period, 664-323 BC" in *Ancient Egypt: A Social History*, 2nd ed. Cambridge, New York: Cambridge University Press, 2012.
- Warburton, David, Krauss, Rolf, and Hornung, Erik. *Ancient Egyptian Chronology*. Leiden: Brill, 2006.
- Waters, Matt. *A Concise History of the Achaemenid Empire, 550-330 BCE*. New York, New York: Cambridge University Press, 2014.
- Woodward, A. M. "Athens and the Oracle of Ammon," *The Annual of the British School at Athens*, Vol. 57 (1962): <http://www.jstor.org/stable/30104495>

Worthington, Ian. *By the Spear: Philip II, Alexander the Great, and the Rise and Fall of the Macedonian Empire*. Oxford, New York: Oxford University Press, 2014, Kindle Edition. Xen. *Ages*. in *The Loeb Classical Library* ed(s). T. E. Page, E. Capps, W. H. D. Rouse, L. A. Post, E. H. Warmington. Cambridge, Massachusetts: Harvard University Press, 1962. Xen. *Hell*. VII, 1.33-40 in *The Landmark Xenophon's Hellenica* ed. Robert B. Strassler New York: Pantheon Books, 2009.

Predicting Conflict in Sino-British Relations in the Time of the Opium Wars

Lillie Mermoud

Abstract

This research examines the relationship between Britain and China in the early 19th century prior to the First Opium War in 1839, as observed from the art produced by both nations. While it is common to look at this historical event from a political or social perspective, art history has seldom been used to study the Opium Wars, a turning point in modern Chinese history. This research offers an overview of the political atmosphere in China prior to the Opium Wars, as well as necessary knowledge on Sino-British relations leading up to 1839. The focus of this paper rests on analysis of Chinese porcelain ware made for export, as well as British and Chinese caricature drawings. The art discussed throughout the paper illustrates Britain's souring opinions of the Chinese people, while revealing Britain's growing oppression of traditional Chinese culture. This analysis reveals that the art produced by both nations not only illustrates worsening diplomatic relations, but can be used to predict the military conflict known as the Opium Wars.

In the early 19th century, prior to the Opium Wars, Western intrusion in Chinese affairs was on the rise. What political structure that had survived amid the crumbling Qing dynasty was mainly preoccupied with foreign diplomatic relations and was unable to protect the Chinese people against European interference. This paper focuses mostly on Sino-British relations during this time period. Even though many Western powers scrambled and fought for ownership of Chinese land and demanded the right to trade with China – most notably, France, Germany, Russia, and eventually the United States – the British encroached upon China the most. This interference caused the Opium Wars, a forceful attempt to legalize the sale of opium in China, while Britain also negotiated for the status of “most favored nation” in Chinese trade policies.¹

The British public supported this push into Chinese affairs with an increasing disdain for the Chinese people, and British art reflects these views. British commissions for Chinese art shifted during the 19th century from the desire for idealistic scenes of the Yellow Mountains to portraits of stereotyped, seemingly backwards and uneducated Chinese characters. In turn, traditional Chinese art, normally prized among the Chinese as a noble cultural heritage, shifted drastically from its traditional techniques to more Western styles as it was made to be exported to Europe. Chinese artists and artisans visibly adopted Western artistic techniques in order to make their products more marketable, forgoing traditional techniques and subject matter. The Chinese also produced art designed largely for propaganda against British forces, portraying the British as fearsome monsters. This paper explores the possibility of predicting the occurrence of a conflict such as the “unequal treaties,” otherwise known as the Opium Wars, by examining these changes in both Chinese and British art from the late 18th century leading up to the 1830s.

During the years leading up to the Opium Wars, an important source of tension between China and the West was the forceful introduction of opium into China by the British. Britain imported opium from its Indian colonies to the Chinese black market, where opium was illegal due to the risks that it presented to addicts as well as the Chinese economy. It was also considered a vile practice closely associated with the British “barbarians.” However, it quickly became popular at all levels of Chinese society and became a source of income for the British East India Company. Wanting to preserve their profits in the trade, the British forced the sale of opium in China by corrupting Chinese officials who worked near Hong Kong and Macau. These officials would offer their protection to the British opium dealers, as well as allow them entry for a certain cut of the opium. In essence, the British were corrupting the inner workings of the Qing government.²

The British were also exerting pressure on China to participate in more trade with the West than China had previously allowed. The British pressed the Qing government to allow

¹ John Delury and Orville Schell, *Wealth and Power, China's Long March to the Twenty-First Century* (New York:

Random House Trade Publications, 2013): 12-36. The Qing dynasty was China's last dynasty, preceding Sun Yat-Sen's rule, and eventually Mao Zedong's. Its fall in 1911 was due in part to the Qing government's incapacity to rule but also to Western intrusion.

Throughout this paper, “Sino” refers to China. The word is used in academic circles to denote the study of Chinese affairs, or simply something that comes from or relates to China.

² John King Fairbank, *Trade and Diplomacy on the China Coast: The Opening of the Treaty Ports, 1842-1854*, (Stanford: Stanford University Press, 1964): 68-69.

Immanuel C. Y. Hsü, *The Rise of Modern China*, 6th ed. (New York: Oxford University Press, 2000): 169-173.

foreigners to live in cities other than Macau, as well as open more trade ports. This would not only have allowed the British more income from Chinese trade, but it would also have allowed the British to have a stronger military presence within China to protect British commercial interests. Furthermore, the decay of the Qing dynasty initiated the deterioration of overall traditional Chinese culture, making China more vulnerable than it otherwise would have been to foreign assault. All of this culminated in the First Opium War, which lasted from March 1839 to August 1842, and ended in the Treaty of Nanjing. The Chinese consider this the beginning of a century of “unequal treaties,” a just name considering the treatment China received from the British after its defeat. The art produced around this time period from both the Chinese and the British reflects these escalating tensions.³

China and European powers such as France, the Netherlands, and England were engaged in trade for centuries since the 2nd century BCE via the Silk Road. The European market had always been captivated by Chinese tea, silks, textiles, and visual art. In particular, the iconic blue and white qing hua porcelain technique, like the set in [Fig. 1] was in high demand. China did not share the West’s fervor for trade, as it regarded virtually every culture other than Chinese culture as barbaric and culturally inferior. Western technology, literature, and philosophy were viewed as undesirable. Up until the Treaty of Nanjing in 1842, trade was exclusively limited to Hong Kong; foreigners were not allowed into the interior of China. They were only allowed to live within the Portuguese-claimed city of Macau, as the Chinese greatly disdained any foreign presence. China’s trade policies reflected this outlook: China gladly parted with whichever goods the Europeans desired, in return not for foreign products and good relations, but simply for gold. The only object the Chinese imperial court wanted from Europe was its money, viewed as an inferior culture’s monetary recognition of the Middle Kingdom’s superiority.⁴

³ Fairbank, 59. Delury and Schell, 11-14.

⁴ Katrina Hill, “Collecting on Campaign, British Soldiers in China during the Opium Wars,” *Journal of the History of Collections*, (2012): 1-29.

Delury and Schell, *Wealth and Power*, 12-36. Delury and Schell describe Chinese leaders prior to the Opium Wars as referring to foreign powers “pejoratively... as “barbarians,” as well as culturally inferior to China. Many Chinese texts from this time also refer to Europeans specifically as barbarians.

“Early American Trade with China,” Atlas Teaching Resources for the University of Illinois at Urbana Champaign, last modified 2006, accessed March 29, 2016,

<http://teachingresources.atlas.illinois.edu/chinatrade/introduction04.html>. Fairbank, 8-11.

Edward H. Schafer, *Great Ages of Man, A History of the World’s Cultures: Ancient China* (New York: Time Inc., 1967): 171.

Hsü, 92-97.



Fig. 1: Chinese, Garniture, Kangxi period c. 1700. Blue underglaze on porcelain, height of vases: 15 ½”, height of beakers: 13 ½”. Private collection, Spain.

Traditional Chinese export art consisted largely of porcelain serving sets, decorative vases and urns. The most desirable porcelain that Westerners sought out was the qing hua technique, [Fig. 1], also known as blue-and-white porcelain. Vases such as those in [Fig. 1] were widely popular in Dutch, French, Portuguese and English households due to their serene and exotic appearance. The visions of China’s enticing mountains, rivers, and various plants and animals made traditional Chinese art highly marketable. Millions of similar items were exported every year and were used by Europeans as a way of showcasing a family’s wealth. These particular vases show traditional Chinese subject matter on porcelain ware. Each piece in the set is either decorated with a large flowering peony or a Chinese riverscape with crested mountains in the background. Such vases were often placed on elaborately designed mantelpieces that were built to showcase fine Chinese porcelain.⁵

The image of China as a peaceful and mystical land was one that Europeans, in particular the British, favored from the 15th through the 18th centuries. Not only is this seen in the British penchant for purchasing traditional art, it is also seen in their tendency to order Chinese art – mostly porcelain – that reinforced their existing view of China as an exotic, idyllic, simple, yet mysterious land. The British market sought idyllic scenes of pagodas and content mandarin scholars, exotic Chinese landscapes, and blissfully ignorant Chinese peasants [Fig. 2]. This enabled the art to be more readily sold in British markets. It must be noted that this particular type of Chinese art that the British sought was already in existence, and the British simply perceived it as the epitome of Chinese culture. However, in the 19th century, the Qing government was weakened vis-à-vis the ever-growing British military forces and the Chinese

⁵ Margaret Kaelin, *Important Chinese Export Art* (New York: The Chinese Porcelain Company, 1998): 2-3.

Carol Stepanchuk, “Ming Blue-and-White Porcelain,” *Calliope* 5, no. 5 (May/June 1995): 32. The qing hua pattern

became so popular in the Netherlands that they incorporated that style of porcelain into their own art, a style that is still seen in Dutch porcelain today.

Michael Sullivan, *The Arts of China* (Berkeley: University of California Press, 2009): 278.

were not in a position to decline the British demand for Chinese goods. In order to ship larger quantities of marketable goods more rapidly, the Chinese saw themselves producing art that supported the British vision of China, as dictated by Britain. In other words, the Chinese were stereotyping their own culture so that their culture would sell.⁶



Fig. 2: Chinese, Platter, Qianlong period c. 1765. Porcelain, 14¾” length. Private collection, South America.

[Fig. 2] exemplifies the type of art that was commissioned by the English for export from China. Not only does the platter boast the coat of arms of the English Boynton family quartering that of the Topham family, it also depicts the British vision of a Chinese court scene. The woman in the center blushes coyly behind her fan as she greets her lover in a typical Chinese garden – to British eyes, at least – while a female attendant kneels before a man presenting a painted fan. This shows the British understanding of the Qing court, as well as their own fantasies of Chinese courtly interactions. In a sense, this type of export art could be considered the Chinese export ware version of Rococo *fête galante*: an idealistic depiction of aristocratic flirtation amid an exotic and heavenly landscape.⁷

As captivating as this romantic scene might be, Chinese courtly interactions forbade a noble woman such as the one in the center of [Fig. 2] to receive a man alone. In fact, it was considered indecent for a woman to act too enthusiastic around a man she fancied, even if he was her husband, as she would appear too eager for attention and pleasure. Eagerness made a woman appear immodest, which was considered a gross social indiscretion. Had this platter been created for a Chinese audience, the woman would have been painted as almost blushing with shyness around the presence of a man. The couple would also have been chaperoned by an elderly woman or a palace eunuch, depending on the social standing of the two parties involved. Such

⁶ Catherine Pagani, “Chinese Material Culture and the British Perceptions of China in the Mid-Nineteenth Century,” *Colonialism and the Object: Empire, Material Culture, and the Museum* (New York: Routledge, 1998): 28-36.

Stepanchuk, 32. Fairbank, 8-11. Hill, 2.

Clare Le Corbeiller, *China Trade Porcelain: Patterns of Exchange* (New York: The Metropolitan Museum of Art, 1974): 3-4.

⁷ Kaelin, 69. Hill, 2.

scenes as this, however, sparked the imagination and curiosity of British viewers, and households greatly sought to put their coat of arms on such treasured pieces.⁸

Another example of the Westernization of Chinese art is the artistic techniques that the artist used in the production of this platter. Hatching is used as a way of defining the lines of the tree to the left as well as those of the rocks in the lower right corner of the platter. One must note that hatching is not an artistic technique that is original to Asian art as a whole, though here it is used as a way of creating shadows under the figures in the garden. Furthermore, Asian art, and especially Chinese art, traditionally chose not to represent shadows. Hatching and shadows are both aspects of Western art that the artist has adopted in order to make the platter more marketable to a European audience.⁹

The Chinese adoption of Western techniques and subject matter in order to make art more valuable to British buyers shows the position of weakness that China was placed in when facing the English. Chinese artists altered their centuries-old artistic traditions in order to avoid military conflict with Britain – which China was sure to lose – because the British demanded the production of the type of art that would sell at home in order to reap the most profits. The British push for art that misrepresented Chinese culture further increased tensions between the two nations in the years leading up to the First Opium War.¹⁰



Fig. 3: Chinese, *Plate*, Qianlong period c. 1739-43. Hard-paste porcelain, 9” diameter. The Metropolitan Museum of Art, New York.

⁸ Kaelin, 69.

Michael Harbsmeier, “Le Monde Renversé, Quelques Expériences Chinoises de la Modernité Européenne au XIXème siècle,” *Revue Germanique Internationale* 21, (2004): 163-179.

⁹ Kaelin, 69. Harbsmeier, 163-164. Delury and Schell, 29.

David Hockney, “Painting and Photography,” lecture at Getty Perspectives, Los Angeles, California, September 2015. Accessed May 9, 2016.

¹⁰ Harbsmeier, 163-164.



Fig. 4: Chinese, *Plate*, English market c. 1805. Hard-paste porcelain, 8" diameter. The Metropolitan Museum of Art, New York.

In those last years prior to the war in 1839, China felt belittled as the British attitude visibly shifted against China. The British looked more and more disparagingly upon the people of the Middle Kingdom. Chinese art continued to Westernize so it could maximize its marketability overseas, while the English placed increasing pressure on China to maintain trade through growing British military presence. The Chinese had no choice but to continue to cater to Western demands for art. During this time period, Chinese porcelain made for export was Westernized to the extent that it kept no characteristics of traditional Chinese patterns. The porcelain wares were Chinese only in the respect that they were made in China with Chinese firing techniques. They often showed individual European families' coats of arms, decorated with French or English words and other European patterns, such as [Fig. 3] and [Fig. 4]. Upon first glance, it would be near impossible to tell that these two pieces were made by Chinese hands.¹¹

The plate in [Fig. 3], though made a little before the time of the Opium Wars, is a gorgeous example of the extreme westernization that is observed in later Chinese export art. Its ornate patterns were originally drawn by the English Okeover family who commissioned this set. The family sent the pattern for this design to China so that a commissioned artist could replicate it. British demand truly dictated the direction that Chinese art took. This piece would not have been used to eat on but rather as a decorative piece that showcased a family's wealth and artistic connoisseurship. A simpler plate bearing a coat of arms made closer to the time of the Opium Wars, [Fig. 4], shows just as much westernization as the aforementioned plate. The plate bears the coat of arms of the English East India Company. It is painted in a light mauve, and is geometrically pierced around the rim. The color and the piercing of the porcelain is thought to originate from a popular Wedgwood pattern in England during this time period, further exemplifying Western dictation of Chinese export art. These dishes paint the picture of a China that was forced to abide by Western trade demands in order to avoid military conflict, yet

¹¹ Chinese Porcelain Company, *Chinese Export Art from the Eckenhoff and other Collections* (New York: Chinese Porcelain Company, 1997): 50-51. Le Corbeiller, 3.

propaganda posters drawn during the Opium Wars show the extent to which these two countries disdained each other.¹²



Fig. 5: Unknown Artist, *A Chinese Cage*, Illustrated London News, 13 August 1842.

As tensions increased between the two nations, the British felt more and more culturally superior to the Chinese. The English market no longer commissioned porcelain depicting blissful peasants and content scholars. British art became increasingly demeaning to Chinese people, who were interpreted as lazy, cruel, and sinfully backwards in their ways. On August 13, 1842, The Illustrated London News published an image to accompany an article that described the “revolting cruelty” with which the Chinese treated the British, [Fig. 5]. The cage shown in this image, according to the article, was used to capture a British woman by the name of Mrs. Noble. The Chinese allegedly imprisoned Mrs. Noble in this tiny cage and paraded her around the village in order to humiliate her. The article stated that Mrs. Noble was held captive in the cage for six weeks. The story was never proven factually true, yet the journal’s readers took it as a story supporting their existing view of the Chinese.¹³



Fig. 6: Chinese, *The Foreign Devils*, c. 1839-1860. Engraving. The British Library, London.

¹² Chinese Porcelain Company, 50-51. Le Corbeiller, 3.

¹³ Pagani, 30. Hill, 2.

The Chinese artistic response looked similar to that of the British, though Chinese artists were forced to take a more defensive position than the British. The Chinese response came from a place of weakness and a desire for retaliation. Chinese prints depicted the British as pure barbarians, [Fig. 6], as they had always been viewed in traditional Chinese culture, instigating a sense of nationalism and xenophobia in Chinese citizens. A Chinese caricature of a British soldier during the time of the Opium Wars depicts a British man as a fearsome, hairy, fire-breathing monster. It was no doubt meant as propaganda to reinforce the Chinese revulsion for foreigners, but it also clearly reflects the light under which the Chinese viewed the British. The Chinese maintained the view of the British which they had held for centuries, yet the difference was that the British were no longer tolerated as the annoying intruders they once were. They were now viewed as the source of destruction of Chinese culture.¹⁴

After examining the progression of British commissions of Chinese art, as well as the Chinese artistic response, it is clear that the British progressively altered their view of China and its people. What began as an infatuation with Chinese art, which the British originally viewed as incomparable to Western art, quickly degenerated into a way to gain control of China by forcing it to alter its own art. China was already weakened by the decay of the Qing dynasty, as well as the continued British push for more trade and legalized opium, and adopted a position of submission to British demands, yet not without resentment. Under these circumstances, conflict was inevitable. The First Opium War broke out in 1839, ending with the Treaty of Nanjing in 1842 which appropriately became known by the Chinese as the first “unequal treaty.”

The treaty of Nanjing shows that Britain clearly had the upper hand in Sino-foreign relations all along: five treaty ports were officially opened to the West, Hong Kong was given to the British, England was given the status of “most favored nation,” and opium was legalized. The Chinese did not gain anything from the conflict. In the years after the Opium Wars, Chinese art soon fell out of vogue with Britain and the rest of Europe due to political tensions. Cast aside, Chinese art had to learn to move forward after being forced to change an artistic culture that had remained untouched for centuries prior to the arrival of the British.¹⁵

¹⁴ Hsü, *The Rise of Modern China*, 92-97.

It is interesting to observe that the British drawing uses shadows and hatching to create lines, yet the Chinese drawing features neither of these techniques. However, the platter mentioned earlier in [Fig. 2] does show shadows and hatching. Chinese art was westernized when it was made for export, but art made for solely for Chinese eyes abides by traditional techniques.

¹⁵ Delury and Schell, 29. Fairbank, 59. Harbsmeier, 165. Pagani, 30. Hsü, 189-190.

Bibliography

- Delury, John and Orville Schell. *Wealth and Power: China's Long March to the Twenty First Century*. New York: Random House Trade Publications, 2013.
- Fairbank, John King. *Trade and Diplomacy on the China Coast: The Opening of the Treaty Ports, 1842-1854*. Palo Alto: Stanford University Press, 1953.
- Harbsmeier, Michael. "Le Monde Renversé, Quelques Expériences Chinoises de la Modernité Européenne au XIXème Siècle." *Revue Germanique Internationale* 21 (2004): 163-79.
- Hill, Katrina. "Collecting on Campaign, British Soldiers in China during the Opium Wars." *Journal of the History of Collections* (2012): 1-29.
- Hockney, David. "Painting and Photography." Lecture at Getty Perspectives, Los Angeles, California, September 2015. Accessed May 9, 2016.
- Hsü, Immanuel C. Y. *The Rise of Modern China*, 6th ed. New York: Oxford University Press, 2000.
- Kaelin, Margaret. *Important Chinese Export Art*. New York, NY: The Chinese Porcelain Company, 1998.
- Le Corbeiller, Clare. *China Trade Porcelain: Patterns of Exchange*. New York: The Metropolitan Museum of Art, 1974.
- Pagani, Catherine. "Chinese Material Culture and British Perceptions of China in the Mid-Nineteenth Century." *Colonialism and the Object: Empire, Material Culture, and the Museum*. New York: Routledge, 1998.
- Schafer, Edward H. *Great Ages of Man, A History of the World's Cultures: Ancient China*. New York: Time Inc., 1967.
- Stepanchuk, Carol. "Ming Blue-and-White Porcelain." *Calliope* 5, no. 5 (May/June 1995): 32.
- Sullivan, Michael. *The Arts of China*. Berkeley: University of California Press, 2009.
- University of Illinois at Urbana Champaign. "Early American Trade with China." *Atlas Teaching Resources*. Last modified 2006. Accessed 30 March, 2016.
<http://teachingresources.atlas.illinois.edu/chinatrade/introduction04.html>.

A Comparison of Certain Beliefs and Attitudes of Men Who Pay for Sex Vs. Men Who Do Not

Meghan Shuffett

Abstract

As researchers, policymakers hoping to stop prostitution, and feminist anti-prostitution activists began studying prostitutes, they started to notice some alarming trends in the lives of prostitutes. Childhood abuse was common for many of the women and many times this violence continued once they grew up and began prostituting. This discovery of violence in the lives of prostitutes led to an increase in the research on men who pay for sex, or “johns.” Many of these studies looked for characteristics or beliefs that these men have that set them apart from the rest of society. This study uses social learning theory to look at views of women, number of sexual partners, and views on pornography of men who purchase sex, to see if they differ from men who do not. The results show that johns have a higher average of sexual partners, but do not have more conservative views on women or less conservative views on pornography. Other than having more partners, they do not differ from men who do not purchase sex, prompting even more questions on how to differentiate johns from men who do not pay for sex, on how to protect prostitutes from violent men, and on how to make and enforce policies to stop men from paying for sex since prostitution is illegal.

Prostitution, or the buying or selling of sexual acts, is illegal in most of the United States, but it is still present. Sometimes considered the oldest profession, it can be found almost anywhere in the world. Most of the focus on stopping prostitution in America has been directed towards the prostitutes themselves, arresting the women instead of the men soliciting the women – often known as johns (Preston & Brown-Hart, 2009). However, the rise of feminism and the spreading of awareness of issues such as sex trafficking have produced a rise in the focus on johns. The goal of this study is to examine some of the views, experiences, and qualities of these clients and, through social learning theory, to discover if there is a relationship between them and understand why these men choose to purchase sex. Having a better understanding of these johns could play a big role in helping end prostitution and sex trafficking.

Literature Review

Violence in the Lives of Prostitutes

As researcher's interest in prostitution developed, they followed the lead of law enforcement and mostly focused on the prostitutes rather than the johns. One of the reoccurring themes in this research was violence in the lives of the prostitutes. It is clear from past research that childhood abuse was prevalent in the lives of many of the women who engage in prostitution (Farley & Barkan, 1998; Nixon, Tutty, Downe, Gorkoff, & Ursel, 2002). In the 1998 study conducted by Farley and Barkan, 57% of the 130 prostitutes interviewed had a history of childhood sexual abuse, with an average of three abusers per person. Also, 49% of the women reported childhood physical abuse. Another study was done by Nixon et al (2002) who interviewed 47 women that had started prostituting before the age of 18. Of these women, over two thirds of them started when they were 15 years old or younger. Of the women who mentioned childhood abuse in their interviews, 90.9% of them reported being abused. Most of them, 80.3%, were sexually abused, two were physically abused, and one was seriously neglected. Sixteen of them were abused by more than one person.

Results from studies such as these has caused a rise in the interest of prostitution, especially by feminist anti-prostitution activists (Monto & McRee, 2005). These feminists take the stand that, based off results like these, prostitution can only be seen as a form of male domination, and should be viewed as sexual violence against women (Farley & Barkan, 1998; Joseph & Black, 2012). This is supported by research that shows that violence in these women's lives did not stop after they grew up. It continued, becoming especially prevalent through these women's roles as prostitutes (Barnard, 1993; Farley & Barkan, 1998; Nixon, Tutty, Downe, Gorkoff, & Ursel, 2002).

As mentioned earlier, Nixon et al (2002) interviewed 47 women. Not only did they look at childhood abuse, but they also examined the presence of violence from pimps – men who control prostitutes who work for them – and intimate partners, customers, police and other professionals, members of the public, and by other prostitutes in the lives of these women. Half of these women were abused or threatened with abuse by their pimps while in prostitution, and many were abused by their intimate partners at home and were intimidated, forced, or coerced into prostitution by them. Once they became involved, many of the women would be beaten if they did not work. Once these women had started working, the chance of having a “bad date” was very high with over half the women reporting violence from customers. “Such assaults included being stabbed or cut, raped, gang-raped, raped at gunpoint, forced to engage in degrading sexual acts, choked/strangled, beaten, kidnapped, stalked, held with a gun to head, tied up, tortured, beaten with objects (e.g., baseball bat, crowbar), and run over” (pgs. 1027-1028). Only three of the women who were interviewed claimed to never have had a bad date –

out of forty-seven. When these women were abused, many refused to turn to the police or professionals for help for fear of being victimized by them as well. Nine of the women had been assaulted, sexually assaulted, or propositioned by police officers who were supposed to be helping them and several had been propositioned or assaulted by professionals in institutions or group homes.

In Barnard's (1993) study she looked at the issue of violence in street-working prostitutes in Glasgow, Scotland. Prostitutes in this research reported being abducted, punched or raped. During this study, one of the women was actually murdered. There are two main kinds of attacks that happen to these women: rape and robbery. Neither are very difficult to do with the women being so vulnerable in that they are alone with these men, mostly in deserted places. Practically all of the women in this study had faced violence by a client on at least one occasion, many more than once.

Farley and Barkan (1998) researched one hundred and thirty people working as prostitutes in San Francisco, with 13% being men and 12% transgendered. Eighty-two percent of these interviewees reported physical assault since becoming a prostitute, with 55% of it coming from customers. Eighty-three percent had been physically threatened with a weapon and 68% reported rape since entering prostitution. Of those raped, 46% had been raped more than five times. At the end of the interviews, 68% of the respondents met criteria to be diagnosed with post-traumatic stress disorder, and 76% met criteria for partial PTSD. The interviewees involved in prostitution ended up having a higher PCL score (index of PTSD severity) than Vietnam War veterans. Eighty-eight percent of the respondents expressed the desire to leave prostitution.

Research on Johns

As many different researchers continued to show similar results in their data, a rise in the interest of the men soliciting the prostitutes rose. It was obvious there was violence in the lives of these women, many of it coming from the johns who solicit them. This began to raise questions about these men. Is it innately violent men that purchase sex? Does prostitution draw out those men who have violent or exploitative tendencies? Is it all or most men who prey on these women? Is it a small number of violent men that commit a large amount of attacks? Research on johns has begun to rise; however, there is still not much, and it is far less than the research done on prostitutes themselves. Again, some feminists claim this is a double standard where women are blamed above men (Monto & Hotaling, 2001).

Much of the research that has been done on johns is focused on what attributes and feelings these men have that set them apart from those who do not solicit prostitutes. Many researchers hope to begin to get answers for questions like the ones mentioned above to better understand these men who purchase sex. Research done by Blevins and Holt (2009), Bucher, Manasse, and Milton (2015), Monto and Mcree (2005), and Preston and Brown-Hart (2005) studied these men in the hopes of uncovering their qualities and beliefs. Blevins and Holt (2009) found three subcultural norms that shaped the attitudes of johns toward sex work: experience, commodification, and sexuality. Experience, in this context, refers to sex work being an interest or pastime they enjoyed doing. Commodification is viewing sex and the workers themselves as a commodity. Sexuality was how important the sexual acts were to the men. Bucher, Manasse, and Milton (2015) looked at why men solicit prostitutes through Agnew's General Strain Theory and discovered that the main factors in relationship with strain and prostitution for johns were: stress in relation to work, money, and relationships and wanting more and/or less-committal sex to feel more like a man. Engaging in sex was a coping mechanism for the men. Monto and Mcree (2005) found that clients are less likely to be married, and if they are they are less happy in their

marriage, that they have a higher number of sexual partners in the past year, that they are more likely to use pornography, and that they are more sexually liberal than most men.

Preston and Brown-Hart (2005) took a slightly different approach in that they separated clients of prostitutes by race and ethnicities when looking at their attitudes toward women, sex, and prostitution. They found two motives for paying for sex: instrumental eroticism and symbolic eroticism. Instrumental eroticism is where the client gets a direct benefit through a physical or emotional release. Symbolic eroticism uses a broader framework to give justification for soliciting a prostitute. These justifications include the men's views on sexual acts and prostitutes, things from their past such as childhood victimization, and whether or not they believe the prostitutes enjoy their work. Each of the races/ethnicities fell into one, if not both of these categories.

Not only was research being done to discover the overall attributes and feelings of these men who purchase sex, but research on specific harmful qualities and beliefs began to be studied. Specifically, feelings on support for rape myths and sexual assault. In studies done by Monto and Hotaling (2001) and Joseph and Black (2012), it was found that most men who purchase sex have low support for rape myths and sexual assault. However, both found a subcategory in the population of johns that had qualities correlated with acceptance of rape myths and sexual assault. In Monto and Hotaling's (2001) study, the five variables that were correlated with rape myth acceptance were: participating in sexual violence, using pornography, finding violent sexuality appealing, thinking about sex less often, and being sexually conservative. In Joseph and Black's (2012) study the correlation was with a fragile model of masculinity – meaning a man who is insecure around women and is usually rejected by women in everyday life – and was not found in the consumer model of masculinity – a man who sees prostitution as an exchange of goods, likes having a variety of partners, does not want a traditional relationship, and gets excited by soliciting a prostitute. These studies seem to point to violence in the lives of prostitutes coming from a specific subcategory of johns who have certain qualities and beliefs that make them dangerous to these women, rather than these women being in danger from all johns.

These results are consistent with an article by Weitzer (2005) who claims that research has been focused on only one type of prostitution – female streetwalkers – and that this is the minority of prostitutes. He claims that “workers differ in their risk of victimization: Assault, robbery, and rape are occupational hazards for streetwalkers and for those coercively trafficked into prostitution, but are relatively uncommon among off-street workers who have not been recruited by force or fraud” (pg. 216). He states that it is not all johns who are predatory and violent, but that this specific type of prostitution attracts these men. According to him, that is why much of the research on prostitutes has stories and results of extreme violence against these women.

While Weitzer (2005) argues that the focus on prostitution should shift to focus on those types of prostitution that are less studied (call girls, massage parlor workers, brothel workers, escorts, male prostitutes, and transgender prostitutes), the fact is that violence is still happening in the lives of some of these women. Thus, clients should still be studied in hopes of finding why streetwalking prostitution may attract violent men and what qualities and beliefs these men have that may invoke them to commit violent acts. This study will look at some qualities and beliefs of johns and examine them through social learning theory in the hopes of contributing more research on men who solicit prostitutes.

Theory/Hypothesis

In 1985, Akers looked at deviant behavior through the lens of social learning theory. He stipulated that deviant behaviors such as addiction, alcoholism, suicide, violence, sexual deviance, etc. are learned behaviors. Akers posited that these deviant behaviors are not only learned, but continued through observation, reinforcement, and imitation. These behaviors are learned from the social groups that individuals interact and have relationships with, and that social group teaches them what is acceptable and what is not. As the individuals observe these behaviors from those with whom they identify, it is reinforced through rewards and lack of punishment or consequence. Behavior that is not accepted is met with a penalty, consequence, or lack of reward. Then, the individuals imitate these behaviors and accept them as they become part of the group from which they learned the behaviors.

Boeringer, Shehan, and Akers (1991) used this social learning theory of deviant behavior to study sexual coercion and aggression in fraternities. The main concepts from this theory that they used were “differential association (interaction with primary groups), definitions (attitudes toward sexual aggression), imitation, and differential reinforcement (balance of actual or anticipated rewarding and punishing consequences)” (pg 59). These authors claim that these four variables account for significant amounts of deviant behavior among adolescents when looking at their behavior through social learning theory. At the end of the study it was found that when these four variables were controlled for, being a member of a fraternity had no significant impact on sexual coercion and aggression.

This study will use Akers (1985) social learning theory of deviant behavior to study attitudes and qualities of men who purchase sex and to assess what, if anything, separates them from ordinary men. I will observe these men’s attitudes toward gender roles, the number of partners they have had in the past five years, and their views on pornography to see what, if any, affect it has on them paying for sex. These views on women and pornography are often socially taught behaviors that are learned from parents and peers through “differential association” while growing up that influence these men’s “definitions.” Then, once these men have learned these “definitions,” they may “imitate” behaviors that could stem from unequal views of women and less conservative views on pornography, such as having a lot of partners and ultimately perhaps paying for sex. As this last behavior, the one I am studying, is illegal, the “differential reinforcement” aspect of it means they would believe this behavior of paying for sex would outweigh the consequences of being caught and any stigma that may come from hiring a prostitute (Boeringer, Shehan, & Akers, 1991).

Based on social learning theory of deviant behavior, and the four variables Boeringer, Shehan, and Akers (1991) used, I hypothesize that:

1. If a man has a higher number of sexual partners, then he will be more likely to pay for sex.
2. If a man has more conservative views on gender roles and women's abilities, then he will be more likely to pay for sex.
3. If a man has less conservative views on pornography, then he will be more likely to pay for sex.

These hypotheses will be used to discover qualities and behaviors of men who purchase sex to see if there are any distinguishably different characteristics between them and those who do not pay for sex.

Methods

Data

Data for this study was obtained from the 1994 General Social Survey (GSS). Since 1972, the GSS surveys attitudes, beliefs, behaviors, etc. of people aged 18 and older living in the

United States. It is run by the National Opinion Research Center (NORC) at the University of Chicago. The geographic areas of respondents are mixed, and respondents are randomly selected from households. If the respondents agree to participate, they are interviewed in person by the NORC. U.S. Census data is used to guarantee that correct demographics are represented in the sample. The year 1994 was chosen for this study because of the need to have overlapping years for variables, as the GSS does not ask the same questions every time. The number of respondents is 717.

Variables

The dependent variable in this study was *evpaidr*, which is a recoded version of *evpaysx*. *Evpaysx* is the variable name for “Ever have sex paid for or being paid since 18.” The respondents answered either “Yes, No, Don't Know, or Not Applicable” to this question. If they answered yes they were coded as a 1, and no's were coded as a 0. Since the question specifies both paying for and being paid for sex, males was controlled for as studies show that men are overwhelmingly the ones who pay for sex (Blevins & Holt, 2009; Bucher, Manasse, & Milton, 2015; Joseph & Black, 2012; Monto & Hotaling, 2001; Monto & Mcree, 2005; Preston & Brown-Hart, 2005). The independent variables can be separated into three groups. The first is views on women, which has three variables. The first is *fepresd*, which is the recoded variable for “Vote for woman president” with the answers being either “Yes, No, Would Not Vote, Don't Know, and Not Applicable.” No's were coded as 1, and yes's were coded as 0. The second is *fepold*, which is the recoded variable for “Women not suited for politics” with the answers being “Agree, Disagree, Not Sure, No Answer, and Not Applicable.” Agree was coded as 1, and disagree was coded as 0. The third variable is *fehlpd*, which is the recoded variable for “Wife should help husbands career first” with the answers being “Strongly Agree, Agree, Disagree, Strongly Disagree, Don't Know, No Answer, and Not Applicable.” Strongly agree and agree were coded as 1, while disagree and strongly disagree were coded as 0. The second group is number of sexual partners. The variable is *partners2*, which is the recoded variable for “How many sex partners respondents had in last 5 years” with the responses being “0, 1, 2, 3, 4, 5-10, 11-20, 21-100, 100+, 1 or More/Number Unknown, Not Applicable, and Don't Know.” Anyone who reported having 2 or more sexual partners in the last 5 years was coded as a 1, and anyone who reported having 1 or no sexual partners in the past 5 years was coded as a 0. The third group is views on pornography. The variable is *pornlawd2*, which is the recoded variable for “Feelings about pornography laws” with the answers being separated into either pornography should be “Illegal to all, Illegal under 18, Legal, Don't Know, No Answer, and Not Applicable.” Anyone who answered that porn should be legal or illegal for those under 18 (which would be legal for those over 18) was coded as a 1, and anyone who answered that porn should be illegal was coded a 0.

The control variables were *males*, *whites*, *educ* and *income*. *Males* is the recoded variable sex, with the answer categories being “Male or Female.” Males were coded as 1 and Females as 0. *Whites* is the recoded variable race where responses answered “White, Black, or Other.” Those who responded White were coded a 1 and Blacks and Others were coded a 0. *Educ* measured respondents “Highest year of school completed,” with answers going from “0 to 20, Don't Know, and No Answer.” Finally, *income* measures respondent's “Total family income,” with answers being “Less than \$1000, \$1000 to \$2999, \$3000 to \$3999, \$4000 to \$4999, \$5000 to \$5999, \$6000 to \$6999, \$7000 to \$7999, \$8000 to \$9999, \$10000 to \$14999, \$15000 to \$19999, \$20000 to \$24999, \$25000 or more, Refused, Don't Know, No Answer, and Not Applicable.”

Analysis

Descriptive statistical analyses were run on all of the variables to find the mean or modal percentage, and the standard deviation if applicable. Then, a bivariate correlation was run on all the variables to examine the correlations between each of the variables with one another to see any significance between them. After that, a logistic regression was run on all of the variables to discover the significance, if any, between the independent/control variables and the dependent variable. This analysis showed the significance level of the model, the Nagelkerke R Square, the overall percentage of respondents covered by the independent/control variables, and the odds ratios, significance level, and standard error of each one of the independent/control variables.

Results

Table 1 shows the descriptive statistics of the variables. If they were nominal or ordinal the percentage of the responses coded 1 were presented and no standard deviation was reported. If the variables were interval or ratio the mean and standard deviation were reported. As shown on the table, only 9% of respondents answered that they had “Ever had sex paid for or being paid since age 18,” 8% said they would not “Vote for woman president,” 21% said “Women not suited for politics,” and 22% agreed that a “Wife should help husband's career first.” For the question “How many sex partners respondent's had in last 5 years,” 32% answered 2 or more. For “Feelings about pornography laws,” 63% said porn should be legal or legal for those over the age of 18. Forty-three percent of the respondents were males, 83% were white, the average number of years of education was 13.16 years, and the average income was 10.57, which averages about \$15000-\$25000. The sample size ended up being a total of 717 men.

Table 1: Descriptive Statistics

| Variable | Mean/Modal Category | Std. Deviation |
|---|---------------------|----------------|
| Dependent Variable | | |
| Ever paid for prostitute | .09 | |
| Independent Variables | | |
| Would not vote for female president | .08 | |
| Women not suited for politics | .21 | |
| Wife should help husband's career first | .22 | |
| 2 or more partners in the past 5 years | .32 | |
| Porn should be legal/legal over 18 | .63 | |
| Control Variables | | |
| Males | .43 | |
| Whites | .83 | |
| Education | 13.16 | 2.97 |
| Income | 10.57 | 2.43 |

Table 2 shows the bivariate correlations. All of the correlations were weak; however, there were quite a few with significance. It shows that conservative views on women in politics and conservative views on women as president were significantly correlated. The same is true for conservative views agreeing that a wife should help husband's career before her own, believing women should not be in politics, and not voting for a women as president. Also, having two or more partners was significantly correlated with paying for/been paid for sex. Also, having two or more partners in the past five years has a negative, but significant correlation with believing a woman should not be in politics, should not be president, and should help her husband's career before her own. Belief that pornography should be legal or legal for those over age eighteen was positively significantly correlated with having two or more sexual partners in the past five years and with ever paying for/ been paid for sex. It also has a negative significant correlation between all three of the views on women variables. Being male is significantly correlated to paying for/been paid for sex and having two or more partners in the past five years. It also has a negative significant correlation with views on pornography. Being white is negatively correlated with two or more partners and is positively correlated with conservative views on pornography. Education is negatively, significantly correlated with all three views on women variables and with views on pornography. It is positively, significantly correlated with two or more partners and being white. Finally, income has a negative and significant correlation with conservative views of women in politics, two or more partners in the last five years, and less conservative views on pornography. It has a positive and significant correlation with believing wife should help husband's career first, being male, being white, and having a high level of education. While there were several significant correlations here, once a logistic regression was run, many of them lost their significance.

| | Ever paid for prostitute | Would not vote for woman president | Women not suited for politics | Wife should help husband's career first | 2 or more partners in past 5 years | Porn should be legal/legal over 18 | Males | Whites | Education | Income |
|---|--------------------------|------------------------------------|-------------------------------|---|------------------------------------|------------------------------------|-----------------|-----------------|-----------------|--------|
| Ever paid for prostitute | 1 | | | | | | | | | |
| Would not vote for woman president | .038 .115 | 1 | | | | | | | | |
| Women not suited for politics | .013 .606 | .318*** .000 | 1 | | | | | | | |
| Wife should help husband's career first | .015 .527 | .315*** .000 | .345*** .000 | 1 | | | | | | |
| 2 or more partners in the past 5 years | .139*** .000 | -.071** .003 | -.073** .003 | -.177*** .000 | 1 | | | | | |
| Porn should be legal/legal over 18 | .080** .001 | -.160*** .000 | -.118*** .000 | -.210*** .000 | .248*** .000 | 1 | | | | |
| Males | .278*** .000 | .015 .499 | .016 .479 | .013 .562 | .129*** .000 | .199*** .000 | 1 | | | |
| Whites | -.037 .055 | .011 .619 | -.041 .079 | -.022 .333 | -.119*** .000 | -.053* .018 | .033 .070 | 1 | | |
| Education | .017 .381 | -.132*** .000 | -.143*** .000 | -.255*** .000 | .047* .015 | .140*** .000 | .032 .084 | .095*** .000 | 1 | |
| Income | -.011 .594 | -.044 .072 | -.097*** .000 | -.142*** .000 | -.136*** .000 | .068** .004 | .111*** .000 | .134*** .000 | .330*** .000 | 1 |

***: Correlation is significant at the 0.001 level (2-tailed).
 **: Correlation is significant at the 0.01 level (2-tailed).
 *: Correlation is significant at the 0.05 level (2-tailed).

Table 2: Bivariate Correlations

Table 3 shows data from the logistic regression. Once this analysis was completed, only being male and having two or more partners in the past five years was significantly correlated with paying for/being paid for sex. This supports the statement that this variable is measuring men who have paid for sex, especially as men are twelve times more likely than women to answer yes to this question. Those who have had two or more partners in the past five years were almost three times as likely as those who have only had one or no partners in the past five years to pay for sex. This number of partners may not seem meaningful; however, 68% of the sample have only had one or no partners in the past five years. The Nagelkerke R Square was .252. Overall, the independent and control variables accounted for 90.8% of those that responded yes to the dependent variable “Ever have sex paid for or being paid since 18.” This means that only 8.2% is unaccounted for.

Table 3: Logistic Regression

| Variables | Odds Ratios (Standard Error) |
|---|---------------------------------|
| Theoretical Variables | |
| Would not vote for woman president | .907 (.625) |
| Women not suited for politics | .617 (.423) |
| Wife should help husband’s career first | 1.825 (.378) |
| 2 or more partners in the past 5 years | 2.822** (.307) |
| Porn should be legal/legal over 18 | .827 (.336) |
| Control Variables | |
| Males | 12.294*** (.402) |
| Whites | .697 .386 |
| Education | 1.051 .052 |
| Income | .912 .058 |
| Constant | .023 (.910) |
| Nagelkerke R square | .252 |
| Overall Percentage | 90.8 |

Discussion/Conclusion

My first hypothesis, stating that if a man has a higher number of sexual partners, then he will be more likely to pay for sex, was supported. As just mentioned, those who had two or more partners in the past five years were about three times more likely to pay for sex than those with one or no partners. My second hypothesis, stating that if a man has more conservative views on gender roles and women's abilities, then he will be more likely to pay for sex, was not supported. There was no significant correlation between any of the views on women and ever paying for sex. My third hypothesis, that if a man has less conservative views on pornography, then he will be more likely to pay for sex, was also not supported. While it was significant during the bivariate correlation, once a logistic regression was run, it was found to not be significant anymore.

Overall, the only significant characteristics of johns who pay for sex is that they are overwhelmingly males and have a higher than average number of sexual partners. In regards, to social learning theory, it cannot be used to adequately describe men who pay for sex when looking at views on women and views on pornography. According to this study, these johns are just like other men – except for having more sexual partners. It does not appear that certain types of behavior or feelings on women or pornography is taught to them by groups with which they interact. The only belief that could possibly be taught is that to them sex is possibly less committal than for the average man. This makes sense, as paying for sex has a complete lack of commitment other than committing to paying. This less committal view of sex could be absorbed through social learning as men interact with those that teach them this behavior through differential association, then form their definitions believing that sex need not be committal, then imitate this belief by having many partners and paying for sex.

As this study only has one attitude or characteristic that sets johns apart from other men, it is not consistent with those that state johns have deviant personalities that draw them to prostitution and many times cause them to victimize the women. Studies that show prostitutes are victimized continuously throughout their lives are many times led to believe this is the case. (Barnard, 1993; Farley & Barkan, 1998; Nixon, Tutty, Downe, Gorkoff, & Ursel, 2002). This study agrees more with the research that shows prostitution is used by johns as a enjoyable past time (Blevins & Holt, 2009) or releases some stress or has an added benefit (Bucher, Manasse, & Milton, 2015; Preston and Brown-Hart, 2005). Based off the findings of this study, it is more likely that Monto and Hotaling (2001), Joseph and Black (2012), and Weitzer (2005) are correct in that it is not all, or even most of the johns, who victimize these women, but rather a subgroup that are drawn to this lifestyle.

Some drawbacks to this study are that the number who answered yes to “Ever had sex paid for or being paid since age 18,” was relatively small – only 9% of the sample. This could be from lack of men who pay for sex or from men not answering truthfully on the survey. This left only 717 men in the sample. Also, the variable used to measure paying for sex was not separated from being paid for sex. The overwhelming number of men who answered versus women allowed me to use it as a variable just looking at men paying for sex; however, this possibly skewed the data some. Also, this data is older and may not be as accurate as it was seventeen years ago. A final drawback is that I was not able to look at all the characteristics I would have liked to as they were not asked or the sample size was too small.

Future research should continue to study the men who pay for sex to determine if there are certain characteristics or attitudes that either separate them from ordinary men or cause some of them to victimize these women. Regardless of the results of studies like mine, women who

participate in prostitution are still being victimized. Therefore, working to find the types of men who engage in this violence is very important. Also, knowing certain characteristics and attitudes of men who pay for sex could help policymakers as they attempt to put an end to prostitution, since prostitution is illegal in the United States of America.

Bibliography

- Akers, Ronald L. 1985. *Deviant Behavior: A Social Learning Approach* (3rd ed). Belmont, CA. Wadsworth.
- Barnard, Marina A. 1993. "Violence and Vulnerability: Conditions of Work for Streetworking Prostitutes." *Sociology of Health & Illness*. 15(5):683-705.
- Blevins, Kristie R., and Thomas J. Holt. 2009. "Examining the Virtual Subculture of Johns." *Journal of Contemporary Ethnography*. 38(5):619-648.
- Boeringer, Scot B., Constance L Shehan, and Ronald L. Akers. 1991. "Social Contexts and Social Learning in Sexual Coercion and Aggression: Assessing the Contribution of Fraternity Membership." *Family Relations*. 40(1):58-64.
- Bucher, John, Michelle Manasse, and Jeffrey Milton. 2014. "Soliciting Strain: Examining Both Sides of Street Prostitution Through General Strain Theory." *Journal of Crime and Justice*. 38(4):435-453.
- Farley, Melissa, and Howard Barkan. 1998. "Prostitution, Violence, and Posttraumatic Stress Disorder." *Women & Health*. 27(3):37-49.
- Joseph, Lauren J., and Pamela Black. 2012. "Who's the Man? Fragile Masculinities, Consumer Masculinities, and the Profiles of Sex Work Clients." *Men and Masculinities*. 15(5):486-506.
- Monto, Martin, and Norma Hotaling. 2001. "Predictors of Rape Myth Acceptance Among Male Clients of Female Street Prostitutes." *Violence Against Women*. 7(3):275-293.
- Monto, Martin A., and Nick McRee. 2005. "A Comparison of the Male Customers of Female Street Prostitutes With National Samples of Men." *International Journal of Offender Therapy and Comparative Criminology*. 49(5):505-529.
- Nixon, Kendra, Leslie Tutty, Pamela Downe, Kelly Gorkoff, and Jane Ursel. 2002. "The Everyday Occurrence: Violence in the Lives of Girls Exploited Through Prostitution." *Violence Against Women*. 8(9):1016-1043.
- Preston, Pamela, and Alan D. Brown-Hart. 2005. "John Court: Comparison of Characteristics, Sexual Behavior and Sexual Attitudes of Clients of Prostitutes." *Journal of Ethnicity in Criminal Justice*. 3(4):49-65.
- Weitzer, Ronald. 2005. "New Directions in Research on Prostitution." *Crime, Law & Social Change*. 43(4-5):211-235.

Depression and Anxiety: The Influence of Stress, Coping, and General Health

Emma Paine, Judy Latta, Kristin Fowler, Elizabeth Batson, Sandra Willis Ph.D.

Abstract

This paper explores how coping strategies, perceived stress, and perceived general health affect the rates of anxiety and depression among college students. The hypotheses include 1) that anxious and depressed students will engage in more dysfunctional coping styles and fewer adaptive coping styles resulting in increased reports of anxiety and depression, 2) that more anxious and depressed students will have a higher perceived level of stress, 3) if students report higher levels of anxiety and depression, then they are likely to rate their overall health as lower than a person not experiencing anxiety and depression.

The study recruited mostly undergraduate students who attended universities in central Alabama. They were given the Brief Cope, Perceived Stress Scale, and Positive and Negative Affect Scales along with a symptom checklist. This study concluded that the sophomores had the highest reported feelings of anxiety and depression and dysfunctional coping styles. Sophomores were also found to have the lowest scores in adaptive coping use. This supports the well-known phenomenon called the sophomore slump. The research also suggests that as age increases in college, the use of adaptive coping methods also increases which could mean a gradual improvement and recovery from the “sophomore slump.”

Introduction

The physical and psychological implications of stress have a multitude of effects on a person's overall well-being. Anxiety and depression are two common indicators, as they frequently accompany high levels of stress (Seiffge-Krenke, Weidemann, Fentner, Aegheister, & Peoblau, 2001). Stress is often a cycle which becomes increasingly detrimental since higher levels of stress are related to decreased quality of sleep. Less sleep can be the first step in a causal sequence of sickness, anxiety, and depression (Sadeh, Keinan, & Daon, 2004). Finding alternative relationships connecting stress to other symptoms could be a key step in finding innovative ways to manage the stress that plagues many college students. Understanding the relationship between anxiety, depression, and the ways people cope could lead to better treatments for individuals dealing with these issues. Changing coping methods could be an alternative treatment for many dealing with the negative side effects of stress.

College can be an extremely stressful time for many people. Therefore, many attempt to manage this stress by using various coping methods. The study, "Effects of a Brief Mindfulness Meditation Intervention on Student Stress and Heart Rate Variability," created by researchers, Shearer, Hunt, Chowdhury, and Nicol (2016), aims to test the effectiveness of four sessions of mindfulness meditation intervention per week compared to that of an active control condition and a no treatment control. The active control condition was interacting with a therapy dog during a group study break. The main hypothesis for this study is that the mindfulness training and active control treatments would both have a significant effect on the anxiety of participants when compared with the control group who received no treatment at all. (Shearer, Hunt, Chowdhury, & Nicol, 2015)

The 74 participants for this study came from several Psychology courses at the University of Pennsylvania. All participants were randomly assigned to either the mindfulness meditation training group, the active control group, or the no-treatment control group. Those that were assigned to a group receiving treatment participated in four weekly sessions of treatment, each lasting an hour long. Prior to the first treatment session, participants were given a battery of measures assessing variables such as mood, mindfulness personality trait, and general demographics. Within the mood inventory, there was a 20 item subscale of the Spielberger State-Trait Anxiety Inventory (STAI-S; Spielberger, Gorsuch, & Lushene, 1970) and six dysphoric affect items from the expanded form of the Positive and Negative Affect Schedule (PANAS-X; Watson & Clark, 1999). These six items chosen from the PANAS included guilty, sad, distressed, ashamed, lonely, and angry. Symptoms of depression were measured by using the 21 item, forced choice questionnaire, Beck Depression Inventory (BDI-II; Beck, Steer, & Brown, 1996). Mindfulness was measured by the 39 item Five Facet Mindfulness Questionnaire (FFMQ; Baer, Smith, Hopkins, & Toney, 2006). The five measured facets of mindfulness included observing, describing, acting with awareness, non-judging of inner experience, and non-reactivity to inner experience. The researchers also collected data through the use of an electrocardiogram during a cognitive challenge that was included in the post-treatment assessment to determine the participant's heart rate variability (HRV), which measures the body's ability to handle physical stress. (Chowdhury et. al, 2015)

The active control group treatment was brought into a room with a circle of chairs, a table with healthy snacks, and therapy dogs to interact with. They were allowed to interact with both the dogs and their fellow participants in the group. The facilitator of this group would organize group games involving playing with the therapy dogs and feeding them bits of popcorn. Cell phone use was not allowed during the hour of treatment. The mindfulness treatment included

breathing exercises, basic yoga, light stretching, balancing exercises, meditation sessions, and education about the stress response of the body. The participants in the no treatment control group received the mood inventory survey at the same time as the other group meetings and received a challenging, possibly even stressful cognitive test as a post-assessment (Chowdhury et.al, 2015).

This study found a significantly lower level of anxiety in the mindfulness group than in the other groups after the four weeks of treatment. One surprising finding of this study is that mindfulness training did not show a positive connection with self-reported mindfulness. The researchers attributed this discrepancy to the inconsistencies in measuring levels of mindfulness in various tests. Moreover, those in the mindfulness group were found to have a significantly higher HRV in the cognitive challenge than the other groups. This shows that the participants who received mindfulness training intervention had a more adaptive response to stress and therefore, it can be inferred that mindfulness training could help students in college with managing the academic and cognitive stress that they experience in a more positive and beneficial manner (Chowdhury et. al, 2015).

Chronic stress can cause many unwanted physiological and psychological symptoms such as headaches, low energy levels, and insomnia. The study, “Effect of Perceived Stress on Cytokine Production in Healthy College Students” by Vorachai Sribanditmongkol, Jeremy L. Neal, Thelma E. Patrick, Laura A. Szalacha, and Donna O. McCarthy (2014) investigates the relationship between chronic psychological stress and cytokine production and glucocorticoid sensitivity (GCS) of leukocytes after a treatment of an influenza vaccine from healthy college students. Researchers hypothesized that psychological stress and cytokine production in participants would have a significant direct relationship, meaning that as psychological stress increases cytokine production would increase as well. It was also hypothesized that those with greater levels of stress would have lower GCS in a laboratory model of the influenza vaccine challenge (Sribanditmongkol, Neal, Patrick, Szalacha, & McCarthy, 2015).

This study used 61 full-time, military college students with an average age of 22 from the ROTC program at a large, Midwestern university. The subjects were required to be in good health with no pre-existing conditions, acute illnesses, or current use of any serious medications. Participants were asked to complete a demographics form, several questionnaires, and a blood sample. The researchers chose to measure the student’s psychological stress by using the 10-item Perceived Stress Scale (PSS) questionnaire which utilizes a Likert-type scale. A blood sample was taken and sent to labs to be analyzed within two hours after the blood was drawn. The researchers used Pearson correlations to estimate the relationship between perceived stress (PSS) and cytokine production. (McCarthy et. al, 2015)

This study found that among the ROTC students who chose to participate, PSS will reduce GCS and, while not listed as statistically significant, increase production of cytokines in response to the treatment of an influenza vaccine as the researchers hypothesized. Similar to the U.S. norms, the average score found on the 10-item PSS was 12.82. However, this average was lower than the mean for young adults in the general population (aged 18-25). There was no correlation between the PSS and demographic variables found. Therefore, according to this specific study, demographic variables did not play a role in influencing the participant’s levels of perceived stress. In conclusion, this study found that the psychological stress experienced by 61 ROTC university students does influence cytokine production and GCS level, however, the researchers mentioned that for some points of interest, a larger, more diverse sample of

participants could be beneficial in determining true statistical significance. (McCarthy et. al, 2015)

Avi Sadeh, Giora Keinan, and Keren Daon (2004) studied how coping styles of college students could relate to stress levels and sleep behaviors. The aim of the study was to determine how different coping strategies would moderate the levels of stress and sleep in college students applying to an exclusive graduate program. The first of two hypotheses was that people with high emotion-focused coping would have more disrupted sleep in the high-stress time of the experiment than in the low-stress time. The second hypothesis stated that people with high disengagement would have more successful sleep during the high-stress period. There was no specific hypothesis about problem-focused coping, but it was included in the measure and in the results to see what contribution it would make to the study (Sadeh, Keinan, & Daon, 2014).

This specific quasi-experiment consisted of thirty-six undergraduate students applying to the same graduate program. The students reported data on both a low-stress normal week and a high-stress week when they were interviewing for the graduate school. During the course of the experiment, the participants completed the COPE inventory, wore an ActiGraph, and recorded stress and sleep. The COPE inventory was given during the low-stress week of the study, and the fifty-two items were used to assess the student's predominant coping strategies. The three global strategies that were considered for use in the study were emotion-focused coping, problem-focused coping, and disengagement. The students were also instructed to wear an ActiGraph while they were sleeping in both the low and high-stress weeks. The ActiGraphs recorded sleep onset time, sleep period, true sleep time, sleep percentage, awakening, and end of the sleep period. This device allowed for an objective measure of time and quality of sleep. The subjective measures of sleep and stress were in the form of the daily logs that were completed twice daily during the testing weeks. These scales included perceived morning stress, perceived evening stress, sleep quality, and difficulty falling asleep. With both objective and subjective measures to compare to the established coping strategies, several strong conclusions can be drawn from the quasi-experiment (Sadeh, et. al, 2014).

The first comparison the researchers did was to compare the stress levels of the two weeks to be certain that the students were, in fact, experiencing increased stress levels in response to the graduate school application process. They found significant levels of increased perception of stress during the week of the graduate school interviews. This is the foundational comparison that all the other findings can be related to. According to the findings, the people who exhibited high emotion-focused coping got less sleep and reported poorer sleep in the high-stress week than those with lower emotion-focused coping. The relationship between disengagement and sleep was not significant and therefore could not support their second hypothesis. The researchers also found that high problem-focused coping was related to more time sleeping no matter the stress level and a better reported quality of sleep in high stress (Sadeh, et. al, 2014).

Dealing with stress is something that college students are all too familiar with. The preliminary signs of the harmful effects of chronic stress are seen in even considerably healthy individuals. In the study by Carol Dolan, Andrew Sherwood, and Kathleen Light (1992), the links between coping strategies and the physical signs of stress on the body of college students. The hypothesis is that self-focused coping and other coping strategies that involve suppression and ineffective coping may be related to blood pressure elevations. The concentration on self-focused coping includes the tendencies to self-blame and continually put oneself in stressful

situations, and these particular traits have been linked to hypertension in previous studies (Dolan, Sherwood & Light, 1992).

The study included 30 college males who filled out a questionnaire and had their blood pressure measured on both a high stress and low-stress day. As part of the survey, the participants filled out a Ways of Coping Questionnaire which was a 66 item survey from which they determined whether the students were inclined to the problem-focused coping, wishful thinking, detachment, seeking social support, focus on positive, self-blame, tension reduction, or keep to self. From the Ways of Coping Questionnaire, the researchers derived the Self-Focused Coping Scale by combining the self-blame and keep to self-subscales. Within this newly created subscale, researchers focused on the most extreme cases for the comparison to attempt to magnify the differences in the groups of self-focused coping and not self-focused coping individuals. Participants also completed a family health history questionnaire and the Spielberger Anger Expression Scale, which were used to support findings from the coping scale and blood pressure correlation. In addition to the coping questionnaire, the participants had their blood pressure monitored for two days during the study. One day was a low stress, normal class day, and the other was a high-stress exam day. They compared daily activity for each of the testing days to ensure that any difference was in fact due to stress instead of another factor. They also split the groups so some of the coping styles were on their high-stress day on the first day of testing to remove the stress of having their blood pressure taken as a factor in the results (Dolan, et. al, 1992).

On the low-stress days, both groups experienced increases in blood pressure during class and study time then went back to a decreased level after the mild stressors were accomplished for the day. On the high-stress exam day, the high self-focused coping group had higher blood pressure during the study time and the exam and experienced a much smaller drop in blood pressure after the stressors were over for the day in comparison to the high self-focus coping group (Dolan, et. al, 1992).

A study performed in 2001 at the Department of Psychology at the University of Mainz in Germany by Inge Seiffge-Krenke, Sandra Weidemann, Sonja Fentner, Nicole Aegenheister, and Mirian Poebrau (2001) and published in the article “Coping with School-Related Stress and Family Stress in Healthy and Clinically Referred Adolescents” investigated the perceived stress and coping style of subjects in response to two normative stressors, school-related stress and family stress, in patients with varying health status, age, and gender. The two health statuses compared were clinically and non-clinically referred subjects. The age groups investigated were early and late adolescents. The experimenters hypothesized that clinically referred adolescents perceive higher stress levels, especially when exposed to parental conflict, that clinically referred adolescents employ a more dysfunctional coping style, like withdrawal or avoidant coping, in response to stress, that late adolescents perceive more school-related stress than early adolescents, and that females perceive higher stress than males.

The sample consisted of 77 adolescents, with 35 of the participants being early adolescents (ages 12-15) and 42 being late adolescents (ages 16-17). Roughly half of the participants were inpatient at a psychiatric unit for adolescents and children. Participants in the clinically referred group had been diagnosed with a phobia, depression, antisocial behavior, drug abuse, anorexia or bulimia. Anticipatory coping and coping after a stressful event were analyzed both using the Coping Process Interview and traditional surveys. The Coping Process Interview was administered shortly after exposure to a stress and included 7 open-ended questions: definition of the situation, the context of the situation, subjective interpretation of the causes,

appraisal of the event as threatening, challenging or loss, the coping process, evaluation of intended and achieved effects, and reappraisal. In this study, the subjects were asked to answer these questions concerning the most recent stressful event at school and at home. Stress was then measured using the Problem Questionnaire. The Problem Questionnaire consists of 64 items covering seven domains, though this study focused on two domains: problems with school, which included eight items, and problems at home, which included ten items. The participants ranked each item's perceived stressfulness on a scale of 1, "not stressful at all," to 5, "highly stressful." Lastly, anticipatory coping was analyzed using the CASQ, which measures twenty coping strategies over eight domains, though this study only used the domains of school and parents.

The results of the Coping Process Interview varied both by the health status of the group and the stressor type. The clinically referred group reported higher chronic stress and being more affected by the event than the control group in response to school-related stressors. In response to parental conflict, the clinically referred group reported more negative emotions and perceived the conflict as more stressful and unpleasant than the control group. Healthy adolescents reported more cognitive coping responses to school-related stressors while clinically referred adolescents reported more cognitive coping styles in response to family conflicts and had a more negative perception of the problem. While both groups named a similar amount of active coping responses, the clinically referred group's actions were more ambivalent, characterized by an equal amount of seeking support and passivity. Notably, both health status groups reported that they still think about conflicts with parents more than school-related stressors.

The results of the standardized questionnaires suggest both that late adolescents perceived school-related stressors as more stressful than early adolescents and that clinical participants perceived both stressors as more stressful than the control group. Furthermore, the results of the CASQ showed that the clinical group engaged in more dysfunctional coping styles, like withdrawal, when exposed to both stressors, though clinical participants were also less likely to engage in active coping styles when exposed to school-related stressors.

These results suggest both that clinically referred adolescents engage in more dysfunctional coping styles and that late adolescents experience higher stress levels in response to school-related stressors. These results are clinically important, especially in stress management for clinically referred adolescents. Treatment must focus on changing their coping style from withdrawal and avoidance to more active styles (Seiffge-Krenke, Weidemann, Fentner, Aegheister, & Peoblau, 2001).

In 2015, Asim Çivitci, of the Department of Counseling and Guidance at Pamukkale University, investigated the relationship between stress, affect, and social support of college students in the article "The Moderating Role of Positive and Negative Affect on the Relationship between Perceived Social Support and Stress in College Students." Previous research has shown that positive affect has a positive relation with social support and negative relation with stress while negative affect has a negative relation with social support and positive relation with stress (Green et al., 2012; Hamama, Ronen, Shachar, and Rosenbaum, 2013; Al Nima, Rosenberd, Archer and Garcia, 2013; Brannan, Boswas-Diener, Mohr, Mortazavi and Stein, 2013; Wright et al., 2014; Zhou, Zhu, Zhang, and Cai, 2013; Jou and Fukada, 2002). This previous research spurred Çivitci to investigate a possible moderator effect of positive and negative affect between social support and stress. Accordingly, he hypothesized that as social support increases, stress decreases, though positive affect will result in a higher decrease in stress while negative affect will result in a lower decrease in stress.

To test these hypotheses, Çivitci randomly surveyed 479 undergraduate students from Pamukkale University in the School of Education. 329 participants were female and 150 were male, with their ages ranging from 17 to 33 years old. Of the participants, 111 were freshmen, 93 were sophomores, 166 were juniors, and 109 were seniors. The survey included the Multidimensional Scale of Perceived Social Support (MPSS), the Perceived Stress Scale, and the Positive and Negative Affect Scale (PANAS). MPSS, developed by Zimet et al. in 1988, measures social support using three sub-scales: family support, friend support, and significant other support, with each subscale having four items for a total of twelve items. The subject is scored both on the subscale and total scale level, using a 7-point Likert scale ranging from “very strongly disagree” to “very strongly agree” with higher scores relating to higher perceived social support.

The Perceived Stress Scale, developed by Cohen, Kamarck and Mermelstein in 1983, is used to determine how uncontrollable, unpredictable, and burdening one perceives his or her life to be. These factors are determined by measuring how often, on a 5-point Likert scale ranging from “never” to “very often” a stressful situation has occurred in the last month with a higher score indicating higher perceived stress. Traditionally, the Perceived Stress Scale includes 14 items, though Çivitci uses the shortened version with only 10 items.

Lastly, the Positive and Negative Affect Scale, developed by Watson, Clark, and Tellegen in 1988, is a 10 item scale that measures positive and negative affect using responses to how the subject felt in the last two weeks on a 5-point Likert Scale ranging from “very slightly” to “very much.” A high score in the positive or negative affect dimensions indicates high positive or negative affect, respectively.

The data was analyzed using hierarchical multiple regression analysis based on Baron and Kenny’s moderating model. The results indicate that positive affect is negatively correlated with perceived stress and positively correlated with social support, with the opposite also being true for negative affect. There was also an inverse relationship found between perceived stress and perceived social support. When analyzing the moderating effects of positive and negative affect, however, positive affect had no moderating effect between social support and perceived stress while negative affect did. From these results, Çivitci concluded that as one’s negative affect increases, social support’s positive effect on perceived stress decreases. As one’s positive affect increases, however, there is no difference in social support’s positive effect on perceived stress. These findings are significant for clinical application, especially in patients with high negative affect. Stress management may be more successful if treatment focuses on decreasing negative affect in especially high negative affect patients (Çivitci, 2015).

In Roddenberry and Rank’s study (2010) they stated that most research examines the relationship between stress and illness and assess the biological phenomena that mediates the relationship. Their study tested the mediating effects of locus of control and self-efficacy in the relationship between real world and academic stress, illness, and the utilization of health services.

The researchers’ hypotheses were that 1) Stress would be related significantly and positively to illness, 2) That external locus of control and negative self-efficacy will be related positively to increased stress, illness, and the utilization of health services, and 3) That the locus of control and self-efficacy will mediate the relationship between stress and illness when stress levels are high.

For this study, students were recruited through in-class announcements in upper-level psychology classes. It was hoped that by recruiting from advanced undergraduate psychology

courses that it would provide a sample that would be experiencing stress mainly due to the administration of final exams (versus new college students who were experiencing more stress about transitioning). In total, there were 159 participants, with an average age of 24.8 years. It was predominately female (77%). The majority of the sample said they were juniors, Caucasian, and full-time students, taking an average of 12 credits.

Participants completed the Perceived Stress Scale (PSS) to assess general stress. The scale consisted of 14 items that are scored on a four-point Likert scale. Participants also completed the Academic Stress Scale (AS) to measure academic stress by assessing the degree to which academic events are deemed stressful. To measure locus of control, participants completed the Multidimensional Health Locus of Control Scale (MHLC). This measures 18 items on a 6-point Likert scale. To measure self-efficacy, participants completed the General Self-Efficacy (GSE) subscale from the Self-Efficacy Scale. This subscale had 17 items which were scored on a 14-point Likert scale. To measure illness, participants took the Brief Symptom Inventory (BSI) to measure psychological illness and the Pennebaker Inventory of Limbid Languidness (PILL) to measure symptoms of physical illnesses. Utilization of health services was examined by assessing participants' answers to questions regarding their utilization of the services.

The findings of this study support the relationship between an increased level of general stress and increased reports of illness with both psychological and physical symptoms. One unique finding was that the students' experiencing more academic stress had more psychological, but not physical, symptoms. Consistent with the hypotheses proposed, participants' stress levels are related to both their locus of control and self-efficacy. In particular, higher levels of general and academic stress are related to higher levels of external locus of control and lower levels of self-efficacy. Also, higher levels of external locus of control were related to higher levels of illness and higher levels of self-efficacy are related to lower levels of illness. Higher levels of academic self-efficacy are related to lower levels of illness. Overall, the variables examined in this study are not significantly related to participants' utilization of health services. Failure to find significant results may have been because of the limited use of health services by participants in this sample. With regard to the hypothesis proposed for the structural equation model, the path coefficients suggest that there are direct relationships between stress and illness and as well as between illness and the utilization of health services. The hypothesis that stress and illness are correlated was not completely supported by the data. Limitations of this study include the age and gender of the sample, the use of self-report data, and the participants limited used of health services (Roddenberry & Rank, 2010).

Ling, He, Wei, Cen, Zhou, and Zhong (2016) focused on the moderating roles that intrinsic and extrinsic goals can have on hassles and depressive symptom development among undergraduate students using a five-wave longitudinal design with new assessments every three months. The researchers hypothesized that lower levels of intrinsic goals and higher levels of extrinsic goals would be associated with greater within-subject increases in depressive symptoms following within-subject increases in stress levels.

Participants were drawn from a five-wave longitudinal study of late-adolescent students from a university in Guangzhou, Guangdong, mainland China. All participants were living on their university's campus. The participants must have participated in at least two of the five measurement waves. The ethnic distribution was consistent with that of China. Personal consent was required before participants were allowed to be a part of the study. 533 students were offered a chance to participate in the study, but only 502 students chose to participate, and only 427 students had useful data. Students were given 1) The Center for Epidemiologic Studies

Depression Scale (CES-D 2) The revised Aspiration Index (AI-R 3) The General, Academic, and Social Hassles Scale for Students (GASHSS) questionnaires during free time after school. Follow-up assessments were performed every three months for twelve months. The average time to take the first assessment was 20 minutes, and for the following assessments, the average time to finish was 10 minutes. The students had just begun their first year at the time of the first assessment and had just begun their second year at the time of the last assessment (assuming they followed a normal pace).

The CES-D is a 20-item self-report instrument that was used to assess students' depressive symptoms. Students were asked to rate the occurrence of each depressive symptom on a scale ranging from 1 (rarely) to 4 (most of the time). Total scores ranged from 20 to 80. Higher scores indicated higher levels of depressive symptoms. The AI-R measured intrinsic and extrinsic goals using 35 items. The intrinsic goal score was the sum of scores for self-acceptance, affiliation, community feeling, and physical fitness subscales. The extrinsic goal score was the sum of scores for the subscales for financial success, social recognition, and appealing appearance subscales. Stress was measured on the GASHSS. The 30-item test had students rate the frequency and duration of general, academic, and social hassles. Hassle scores could range from 0 – 180, with higher scores reflecting a greater frequency of hassles.

When analyses were carried out, the dependent variable was participants' fluctuations in depressive symptoms during the follow-up period, as predicted by goal orientation style and fluctuations in hassles. Stress was found to be highest at the initial assessment. Consistent with their hypotheses, lower levels of intrinsic goals were associated with greater increases in depressive symptoms following a time with many hassles— meaning that having an intrinsic locus of control might help protect students from depressive symptoms. Intrinsic students interpret events negative events as a possibility for growth. Contrary to their hypothesis, undergraduate students with higher levels of extrinsic goals did not report greater increases in depressive symptoms following the occurrence of hassles relative to those with low extrinsic goals. Limitations of this study include the use of self-report, the homogeneity of the sample of undergraduate students, other types of stress are likely to cause depressive symptoms that are not accounted for, and the fact that no specific negative life events were identified.

The first hypothesis is that more anxious and depressed students will engage in more dysfunctional coping styles and fewer adaptive coping styles. Students with more maladaptive coping styles should, in turn, be coping with stress less effectively, resulting in increased reports of anxiety and depression.

The second hypothesis to be tested states that more anxious and depressed students will have a higher perceived level of stress. Since dysfunctional coping styles are maladaptive, students will perceive a higher level of stress.

The last hypothesis is that if students report higher levels of anxiety and depression, then they are likely to rate their overall health as lower than a person not experiencing anxiety and depression. Generalized anxiety and depression should be indicators of how a person is coping and dealing with stress and, therefore, should be indicative of how effectively a person is adapting to life's stressors. If this general life adaptation also correlates with general health ratings then this would be a strong indicator of how stress, coping, and the psychological indicators of anxiety and depression dictate overall health (Ling, He, Wei, Cen, Zhou, & Zhong, 2016).

Method

Participants

We recruited mostly undergraduate students ($N = 44$) who attend universities in central Alabama. Students' ages ranged from 18 - 60 years old ($M = 21.4$, $SD = 6.07$). 42% of respondents were female. With regard to academic level, 14% ($N = 6$) of the respondents were freshmen, 14% ($N = 6$) were sophomores, 34% ($N = 15$) were juniors, 34% ($N = 15$) were seniors, 2% ($N = 1$) was a 5th year senior, and 2% ($N = 1$) was a 6th year senior (see Figure 1). See Table 2 for sample means and standard deviations of dysfunctional coping styles, adaptive coping styles, feelings of depression in the last three months, grade point average, perceived stress, and anxiety.

Predictor Measures

Brief Cope (Carver, 1997; as adapted from Carver, Scheier, & Weintraub, 1989). This scale is a shortened version of the COPE, which was developed to assess participants' various coping strategies. Participants responded to each of the items with a rating from 1, *I usually don't do this at all*, to 4, *I usually do this a lot*. Adaptive coping was an aggregate of scores for active coping, planning, and positive reframing. Dysfunctional coping was an aggregate of scores for substance use, behavioral disengagement, and denial.

Perceived Stress Scale (Cohen, S; Kamarck T; Mermelstein R, 1983). The PSS was created to measure the number of situations in one's life that are viewed as stressful. In this case, stress is defined as the extent to which a person believes the demands exceed their resources and coping skills. The test consists of 13 situations and asks the participant to respond by indicating the frequency, on a scale from 0 = *Never* to 4 = *Very Often*, that the participant has felt this way in the last three months. The scores are then calculated by reversing the four positively stated items and then summing all of the items on the scale.

Positive and Negative Affect Scale (PANAS; Watson, Clark, & Tellegen, 1988). On this scale, participants used a 5-point rating scale (1 = *very slightly or not at all* to 5 = *extremely*) to indicate how frequently the participant has felt these feelings or emotions in the last year. This scale consists of two subscales: one with nine items designed to measure positive affect and the other nine items designed to measure negative affect. The items in each subscale are summed and averaged.

Symptom Checklist Participants were asked to mark any conditions they have ever experience or are currently experiencing.

Dependent Measures The dependent measures in this study were the students' anxiety, depression, coping styles, perceived stress, and general health which were all assessed from measures on the Health Belief and Behavior Studies (HBBS) battery.

Procedure

Participants were asked to respond to the HBBS battery (Willis, 2016), which is composed of scales designed to test the strength of faith (SCSORF; Plante & Boccacini, 1997), optimism, positive and negative affect, coping strategies, and multiple dimensions of health locus, including the God locus of Health Control scale (Wallston & Willis, 1999). After participants' responses were collected, the data was entered and analyzed in the Statistical Package for the Social Sciences (SPSS). We tested for Pearson correlation between anxious feelings in the last year, depressive feelings in the last year, general health rating, and perceived stress score. We also tested for group differences of gender, year in college, and anxiety and depressive feelings in the last year, general health rating, grade point average, perceived stress, dysfunctional coping style, and adaptive coping style.

Results

Characteristics of the Sample

In our sample, students had an average perceived stress score of 1.95 (SD=.671) out of a possible 4. They also reported an average amount of times they felt anxious in the last year as 2.80 (SD=1.50) and depressed in the last year as 1.90 (SD=1.63) out of 5, or “extremely often.” The students reported an average score of dysfunctional coping strategies as 1.84 (SD=.628) and adaptive coping strategies of 2.72 (SD=.424) out of 4. Females had slightly higher average dysfunctional coping score (female- \bar{x} =1.94, SD= .749; male- \bar{x} =1.69, SD=0.329), higher perceived stress score (female- \bar{x} =2.11, SD= 0.585; male- \bar{x} =1.69, SD=.734), higher anxiety score (female- \bar{x} =3.33, SD=1.18; male- \bar{x} =1.94, SD=1.60), and depression score (female- \bar{x} = 2.44, SD= 1.74; male- \bar{x} =1.06, SD=.966). Male and female students were equally likely to use adaptive coping styles (female- \bar{x} =2.71, SD=.427; male- \bar{x} =2.75, SD=.430) and similar in general health ratings (female- \bar{x} =7.62, SD=1.88; male- \bar{x} =7.76, SD=1.86). Although not significant, females had higher perceived stress, $F(1,42)=3.655$, $p=.063$ $\eta_p^2=.086$. There was no significant difference between genders on dysfunctional coping, adaptive coping, general health, or anxiety.

Test of Hypotheses

We found that anxious feelings in the last year and depressive feelings in the last year were significantly positively correlated with each other ($r(44)=.696$, $p<.001$). In contrary with the first hypothesis, anxiety was significantly correlated with dysfunctional ($r(44)=.297$, $p=.050$) and adaptive coping styles ($r(44)=-.164$, $p=.288$), however, in accordance with the first hypothesis, depression was significantly positively correlated with dysfunctional coping styles ($r(44)=.337$, $p=.025$) but not with adaptive coping styles ($r(44)=-.290$, $p=.057$). Adaptive and dysfunctional coping style were also significantly negatively correlated, ($r(44)=-.362$, $p=.016$) (see Table 1)

In support of the second hypothesis, anxiety and depression were both significantly positively correlated with perceived stress ($r(44)=.634$, $p=.000$; $r(44)=.477$, $p=.001$). Dysfunctional coping styles were also positively correlated with perceived stress ($r(44)=.455$, $p<.05$), though adaptive cope was not significant ($r(44)=-.183$, $p=.234$).

No significant difference was found between people who reported symptoms of depression and anxiety with those who did not when considering dysfunctional coping or adaptive coping. However, there is an interaction effect of depression and anxiety on general health, $F(1,39)=31.28$, $p<.001$ $\eta_p^2=.445$, which supports hypothesis three (see Table 2).

Interesting differences were found between years in college for depression, anxiety, coping style, perceived stress and grade point average (see Table 3). These differences were most notable for the sophomores. Sophomores reported the highest use of dysfunctional coping styles, feelings of depression in the last three months. Sophomores also had the lowest use of adaptive coping styles.

Discussion

One of the most interesting findings in our research was the relationship between reported year in college and variables such as feelings of anxiety and depression in the last three months, and the use of dysfunctional coping styles. We found that the sophomore class had the highest reported use of each of these factors. Sophomores were also found to have the lowest scores in adaptive coping use, perceived stress, and the second lowest anxiety score. This research would support the well-known phenomenon on university campuses called the “sophomore slump”, referring to the difficulty that sophomores experience as they continue to adjust to college life. However, our research also suggests that as age increases in college the use of adaptive coping methods increases while rates of anxiety and depression decrease, which

could mean a gradual improvement and recovery from the “sophomore slump” (see Figures 2-5).

As expected, we found a significant positive correlation between reported depression and the use of dysfunctional coping methods. However, no correlation was found between reported feelings of depression in the last three months and the use of adaptive coping methods. This suggests that students reporting depression in the last three months are more likely to use dysfunctional methods of coping. Therefore, a treatment for depression that aims to change dysfunctional coping methods could decrease levels of perceived stress and positively impact an individual experiencing depression. Our findings did not support a similar relationship between anxiety and coping style. We found that reported anxiety in the last three months was not significantly correlated with either dysfunctional or adaptive methods of coping with stress. These findings were surprising in that they imply that anxiety is not linked to the selective use of either adaptive or dysfunctional methods of stress. As expected, females reported having increased dysfunctional coping styles, perceived stress, and anxiety or depressive feelings in the last three months more than males. Males reported a greater use of adaptive coping styles and reported higher general health ratings.

The significant positive correlation between perceived stress, anxiety, and depressive feelings in the last three months implies a link between reported high levels of stress and high ratings of anxiety and depression. One alternate explanation for the correlation is that people who are more aware of their stress levels may also exhibit greater awareness of their levels of anxiety and depression. This awareness of their physiological and psychological states could be a personality trait of mindfulness rather than a possible causal relationship between stress and anxiety or depression. As depression and dysfunctional coping are also related, it could be argued that changing one’s coping styles could influence levels of depression, possibly also reducing perceived stress. The connection between altering coping styles and its effects on stress levels still needs to be explored further, though the data implies a possible interdependence between the two variables.

The general health ratings reported by participants were found, as expected, to have inverse relationships with both anxiety and depression. This connection may imply that some college students would define general health as including anxiety levels and depression along with physical well-being. Stress, however, is not included in this idea due to a lack of significant findings, positive or negative, between perceived stress and the participant’s general health ratings. Consequently, this absence of a correlation could indicate that some college students view stress as a temporary state that should not be included as a factor of general health.

Despite the many significant correlations found in our study, there are many substantial limitations as well. For instance, our sample was collected using a self-report method which could lead to some bias due to social desirability. This bias could present as a fear of judgment from the researcher or a fear of false anonymity. Moreover, all collected data was inserted into the SPSS program by hand, which means that human error could be a factor. Furthermore, our variables such as anxiety levels, depression, and perceived stress could have not been clearly defined for the participants which could lead to incorrect data. In addition, our study consisted only of a limited number of college-age students at Samford University and various neighboring colleges and therefore cannot be generalized to the general population.

Bibliography

- Al Nima, A., Rosenberg, P., Archer, T., & Garcia, D. (2013). Anxiety, affect, self-esteem and stress: Mediation and moderation effects on depression. *Plos One*, *8*(9), e73265. Retrieved from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0073265>
- AT Beck, RA Steer, GK Brown - San Antonio, TX: Psychological Corporation, 1996
- Baer, R. A., Smith, G. T., Hopkins, J., Krietemeyer, J., & Toney, L. (2006). Using self-report assessment methods to explore facets of mindfulness. *Assessment*, *13*(1), 27-45.
- Brannan, D., Biswas-Diener, R., Mohr, C. D., Mortazavi, S., & Stein, N. (2013). Friends and family: A cross-cultural investigation of social support and subjective well-being among college students. *e Journal of Positive Psychology*, *8*(1), 65-75.
- Carver, C. S., Scheier, M. F., & Weintraub, J. K. (1989). Assessing coping strategies: A theoretically based approach. *Journal of Personality and Social Psychology*, *56*, 267-283.
- Çivitci, A. (2015). The Moderating Role of Positive and Negative Affect between Perceived Social Support and Stress in College Students. *Educational Sciences: Theory and Practice*, *15*(3). doi:10.12738/estp.2015.3.2553
- Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of health and social behavior*, 385-396.
- Dolan, C. A., Sherwood, A., & Light, K. C. (1992). Cognitive coping strategies and blood pressure responses to real-life stress in healthy young men. *Health Psychology*, *11*(4), 233-240. doi:10.1037/0278-6133.11.4.233
- Green, M., DeCourville, N., & Sadava, S. (2012). Positive affect, negative affect, stress, and social support as mediators of the forgiveness-health relationship. *e Journal of Social Psychology*, *152*(3), 288-307.
- Hamama, L., Ronen, T., Shachar, K., & Rosenbaum, M. (2013). Links between stress, positive and negative affect, and life satisfaction among teachers in special education schools. *Journal of Happiness Studies*, *14*(3), 731-751.
- Jou, Y. H., & Fukada, H. (2002). Stress, health, and reciprocity and sufficiency of social support: a case of university students in Japan. *e Journal of Social Psychology*, *142*(3), 353-370.
- Ling, Y., He, Y., Wei, Y., Cen, W., Zhou, Q., & Zhong, M. (2016). Intrinsic and extrinsic goals as moderators of stress and depressive symptoms in Chinese undergraduate students: A multi-wave longitudinal study. *BMC Psychiatry*, *16*(1). doi:10.1186/s12888-016-0842-5
- Roddenberry, A., & Renk, K. (2010). Locus of Control and Self-Efficacy: Potential Mediators of Stress, Illness, and Utilization of Health Services in College Students. *Child Psychiatry & Human Development*, *41*(4), 353-370. doi:10.1007/s10578-010-0173-6
- Sadeh, A., Keinan, G., & Daon, K. (2004). Effects of Stress on Sleep: The Moderating Role of Coping Style. *Health Psychology*, *23*(5), 542-545. doi:10.1037/0278-6133.23.5.542
- Seiffge-Krenke, I., Weidemann, S., Fentner, S., Aegenheister, N., & Poebblau, M. (2001). Coping with school-related stress and family stress in healthy and clinically referred adolescents. *European Psychologist*, *6*(2), 123-132. doi:10.1027//1016-9040.6.2.123.
- Shearer, A., Hunt, M., Chowdhury, M., & Nicol, L. (2016). Effects of a brief mindfulness meditation intervention on student stress and heart rate variability. *International Journal of Stress Management*, *23*(2), 232-254. doi:10.1037/a0039814
- Spielberger, C. D., Gorsuch, R. L., & Lushene, R. E. (1970). Manual for the state-trait inventory. *Consulting Psychologists, Palo Alto, California*.

- Sribanditmongkol, V., Neal, J. L., Patrick, T. E., Szalacha, L. A., & Mccarthy, D. O. (2014). Effect of Perceived Stress on Cytokine Production in Healthy College Students. *Western Journal of Nursing Research*, 37(4), 481-493. doi:10.1177/0193945914545658
- Wallston & Willis, S. (1999). God locus of Health Control scale: the MHLC Scale.
- Watson, D., & Clark, L. A. (1999). The PANAS-X: Manual for the positive and negative affect schedule-expanded form.
- Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of personality and social psychology*, 54(6), 1063.
- Willis, S. (2016). Health Beliefs and Behaviors Studies: the HBBS Scales.
- Wright, K. B., King, S., & Rosenberg, J. (2014). Functions of social support and self-verification in association with loneliness, depression, and stress. *Journal of Health Communication*, 19(1), 82-99.
- Zhou, X., Zhu, H., Zhang, B., & Cai, T. (2013). Perceived social support as moderator of perfectionism, depression, and anxiety in college students. *Social Behavior & Personality: An International Journal*, 41(7), 1141-1152.

Table 1

Correlations Between General Health Ratings, Anxiety, Depression, Adaptive Coping, and Dysfunctional Coping

| | | Correlations | | | | |
|-----------------------|---------------------|-----------------------|---------|------------|--------------|-------------------|
| | | General Health Rating | Anxiety | Depression | AdaptiveCOPE | DysfunctionalCOPE |
| General Health Rating | Pearson Correlation | 1 | -.184 | -.304* | .015 | .081 |
| | Sig. (2-tailed) | | .237 | .047 | .923 | .608 |
| | N | 43 | 43 | 43 | 43 | 43 |
| Anxiety | Pearson Correlation | -.184 | 1 | .587** | .012 | .073 |
| | Sig. (2-tailed) | .237 | | .000 | .937 | .637 |
| | N | 43 | 44 | 44 | 44 | 44 |
| Depression | Pearson Correlation | -.304* | .587** | 1 | .154 | .009 |
| | Sig. (2-tailed) | .047 | .000 | | .320 | .955 |
| | N | 43 | 44 | 44 | 44 | 44 |
| AdaptiveCOPE | Pearson Correlation | .015 | .012 | .154 | 1 | -.362* |
| | Sig. (2-tailed) | .923 | .937 | .320 | | .016 |
| | N | 43 | 44 | 44 | 44 | 44 |
| DysfunctionalCOPE | Pearson Correlation | .081 | .073 | .009 | -.362* | 1 |
| | Sig. (2-tailed) | .608 | .637 | .955 | .016 | |
| | N | 43 | 44 | 44 | 44 | 44 |

*. Correlation is significant at the 0.05 level (2-tailed).

** . Correlation is significant at the 0.01 level (2-tailed).

Table 2

General health rating means and standard deviations as related to depression and anxiety

| | Depression | Anxiety | Mean | Std. Deviation | N |
|-----------------------|------------|---------|--------|----------------|----|
| General Health Rating | no | no | 8.2917 | 1.12208 | 24 |
| | | yes | 7.1250 | 2.16712 | 8 |
| | | Total | 8.0000 | 1.50269 | 32 |
| | yes | no | .0000 | . | 1 |
| | | yes | 7.4000 | 1.07497 | 10 |
| | | Total | 6.7273 | 2.45320 | 11 |
| | Total | no | 7.9600 | 1.98914 | 25 |
| | | yes | 7.2778 | 1.60167 | 18 |
| | | Total | 7.6744 | 1.84805 | 43 |

Table 3

Means and standard deviations of dysfunctional and adaptive coping, GPA, perceived stress, and anxiety

| Measures | Freshmen | | Sophomore | | Junior | | Senior | |
|-----------------------------|----------|------|-----------|------|--------|------|--------|------|
| | M | SD | M | SD | M | SD | M | SD |
| Dysfunctional Coping Styles | 1.89 | .300 | 2.07 | 1.21 | 1.85 | .687 | 1.77 | .357 |
| Adaptive Coping Style | 2.81 | .340 | 2.56 | .558 | 2.67 | .415 | 2.75 | .426 |
| Feelings of Depression | 1.83 | 1.84 | 2.17 | 1.84 | 2.13 | 1.46 | 1.87 | 1.73 |
| GPA | N/A | N/A | 3.56 | .305 | 3.41 | .416 | 3.31 | .50 |
| Perceived Stress Score | 1.93 | .524 | 1.73 | .398 | 2.04 | .684 | 2.07 | .727 |
| Anxiety | 3.33 | 1.51 | 2.67 | 1.63 | 3.20 | 1.08 | 2.53 | 1.69 |

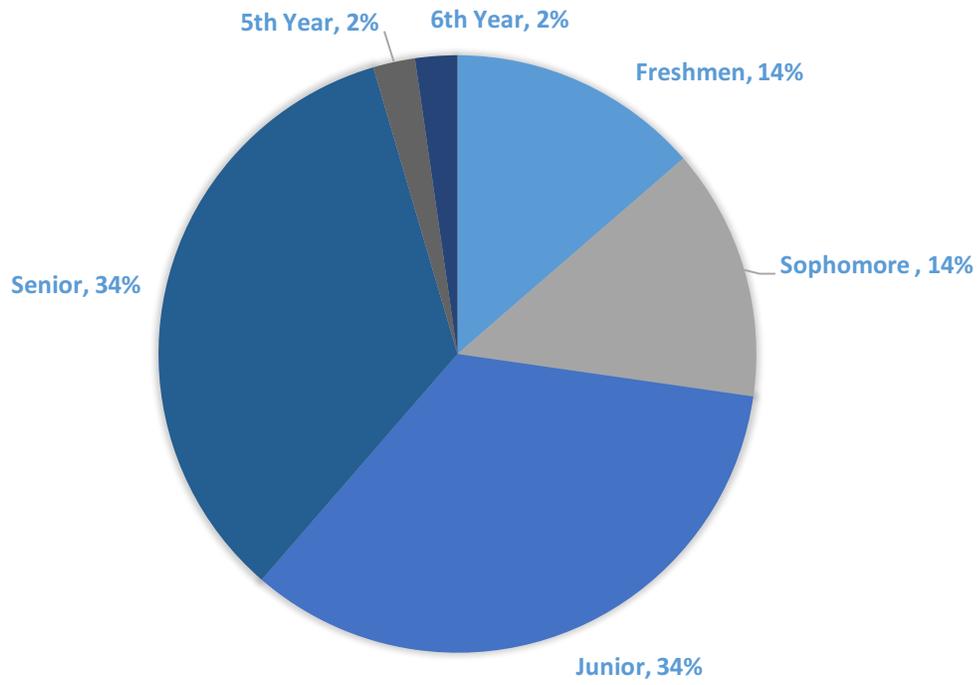


Figure 1. Percentage of participants by year in college

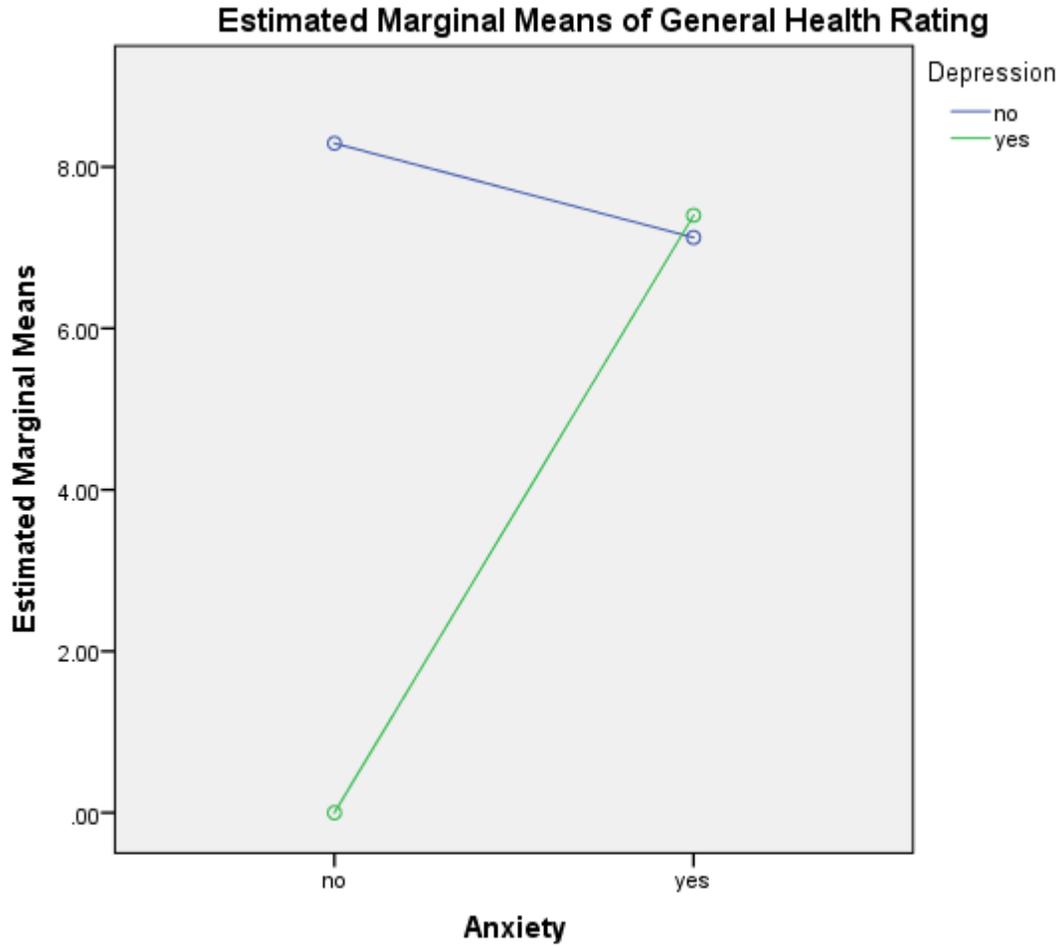


Figure 2. Interaction of depression and anxiety on general health

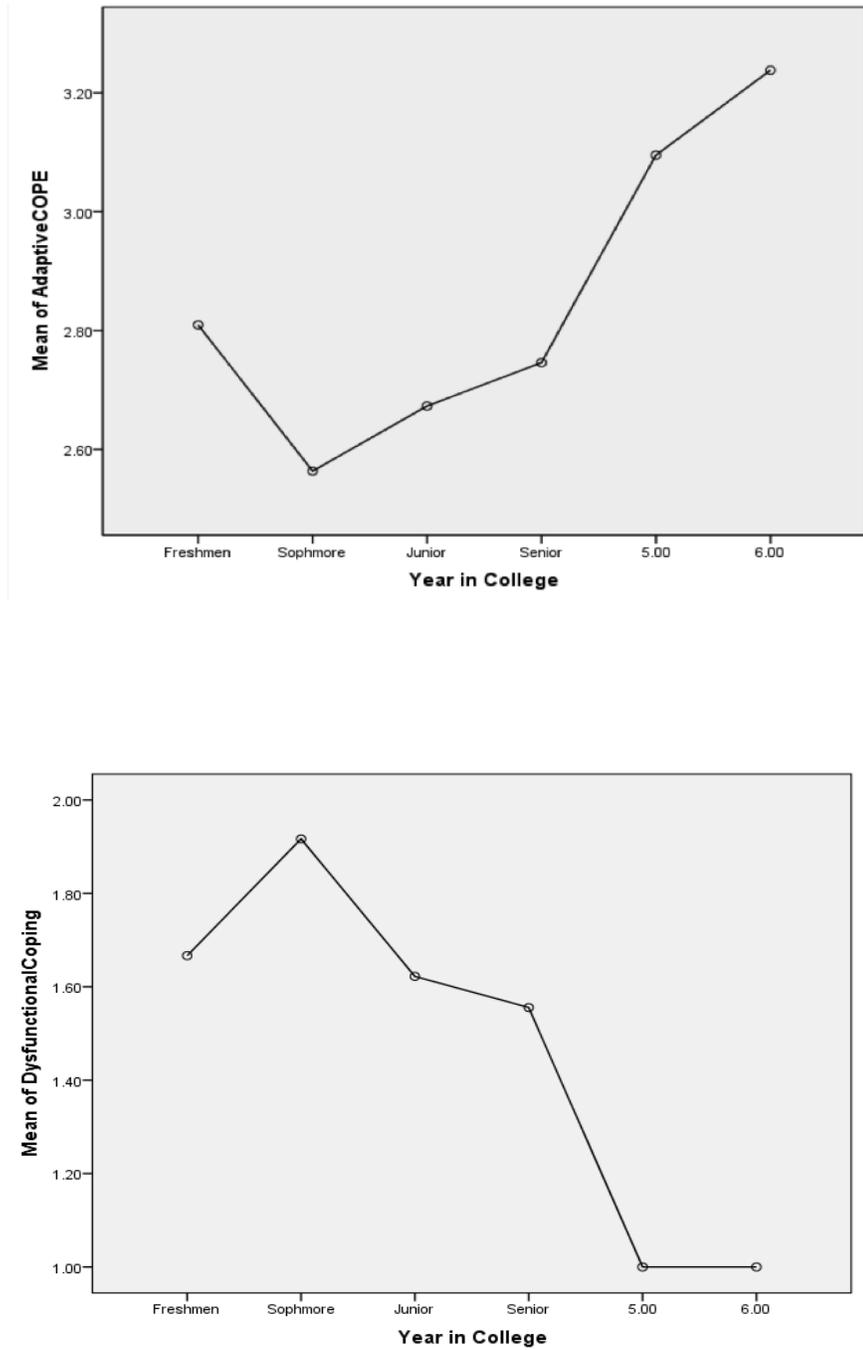


Figure 3. Means of adaptive and dysfunction across years in college

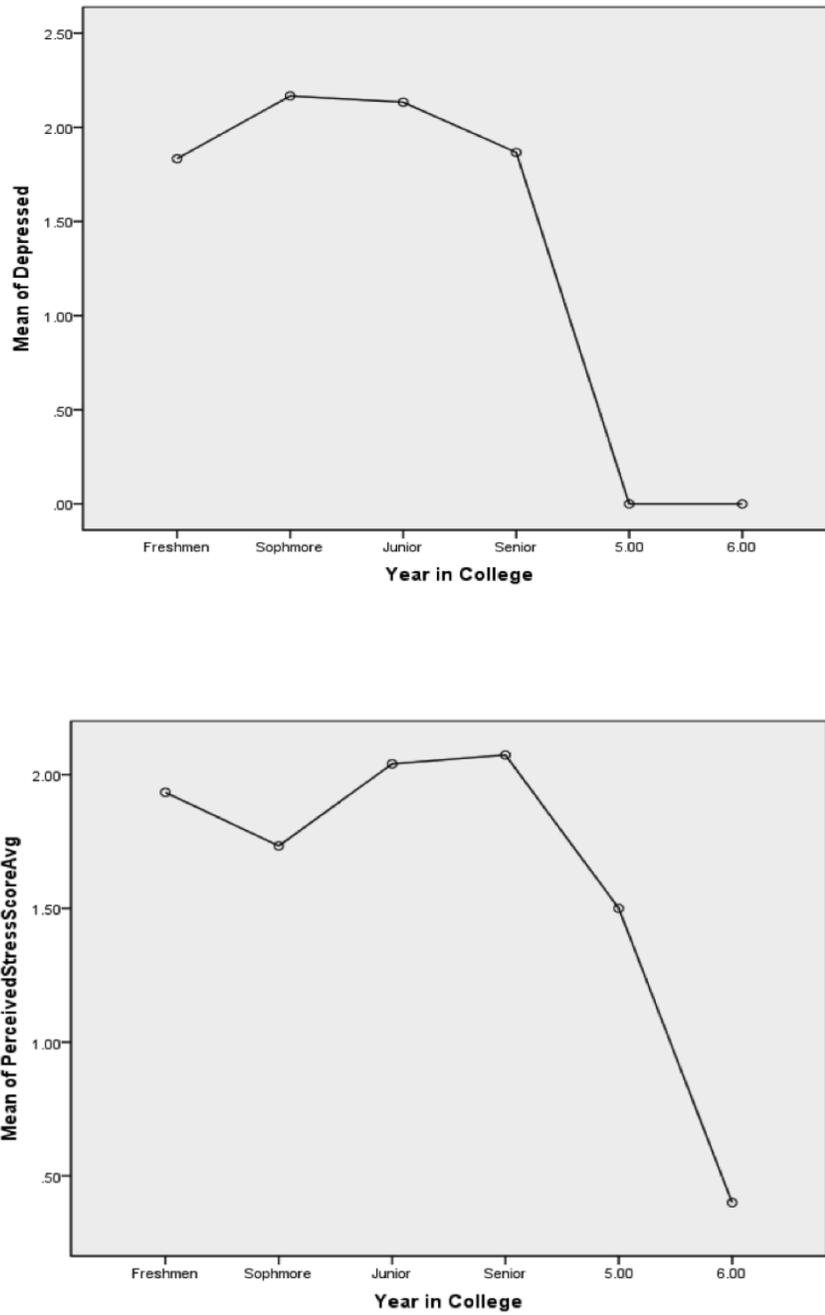


Figure 4. Means of depression and perceived stress across years in college

NASCAR Fans' Recognition of Driver Sponsorship

Olivia Williams

Abstract

The National Association of Stock Car Auto Racing (NASCAR) is an industry that has evolved tremendously over the years in terms of sponsorship. The purpose of this paper was to figure out how well fans recognize the drivers' sponsors, and to determine whether or not they purchase products and/or services based on the sponsorships of these drivers.

The researcher surveyed 88 male and female NASCAR fans, ages 19 and older. A 20-question survey was conducted with questions related to drivers and their respective sponsors. The survey included specific questions designed to examine fans' recognition of sponsorship based on a list of possible correct and incorrect choices.

Results showed that older fans were less likely than younger fans to recall sponsorships. Also, fans of the Camping World Truck Series (NASCAR'S "beginner" series) had much higher knowledge of XFINITY sponsorships (NASCAR'S "intermediate series) than those who were not Camping World Truck fans. Thus, as drivers moved through the levels of NASCAR, fans from the start carried knowledge of sponsorship into the next level. Also, results showed that fans tended to buy more small, cheap products associated with their favorite drivers' sponsors (for example, M&Ms candies) rather than more expensive products.

Each week NASCAR drivers have multiple primary and associate sponsors, based on the target markets and other demographics. Companies will choose to sponsor drivers at specific races because they want to gain market recognition, or perhaps they have a pre-existing market presence that they want to reinforce. Sponsors fuel the funding for drivers to race their cars every week, and without them it's safe to say there would be no NASCAR.

In 2013, Adweek estimated that primary sponsorships can range anywhere from \$5 million to \$35 million ("Explaining the difference in sponsors," 2015). Primary sponsors usually have five placement areas on the car: the hood, the rear-quarter panel, the deck lid, the TV panel, and the roof panel ("Explaining the difference in sponsors," 2015). The lower rear-quarter panel is designated for a team that has rotating primary sponsors ("Explaining the difference in sponsors," 2015). An example can be seen in the graphic on the right (Klara, 2013).

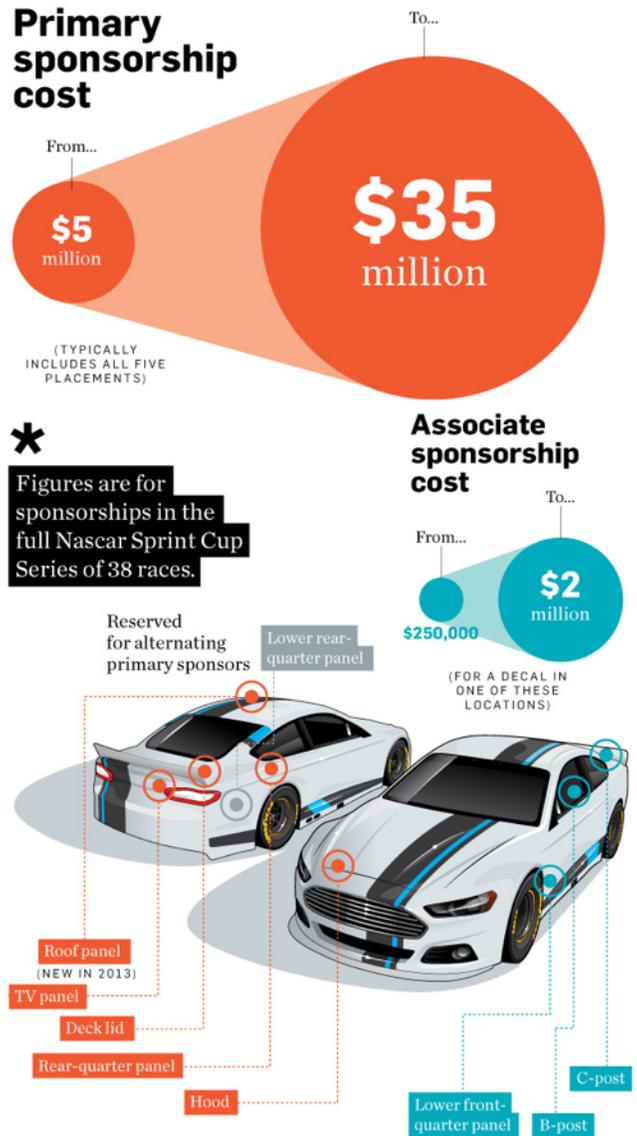
Associate sponsorship can be recognized by a decal in one of three areas: the lower front-quarter panel, the B-post, which is the area next to the driver's shoulder, or the C-post, which is located next to the rear window on both sides of the car ("Explaining the difference in sponsors," 2015). The quarter panel is usually the most expensive location for an associate sponsor, followed by the C-post and B-post ("Explaining the difference in sponsors," 2015).

The level of commitment from NASCAR fans is highly unrivaled. Whether they're watching races at home or travelling weekend after weekend to follow their favorite drivers, NASCAR fans remain some of the most loyal people in the world of sports. All the while, sponsors work tirelessly to vie for their attention.

But how well do NASCAR fans know the drivers' sponsors? Do they purchase products and/or services based on the sponsorships of these drivers? This paper attempts to answer these questions.

Literature Review

Many scholars have studied NASCAR'S history and NASCAR as it relates to marketing. Menzer (2001) studied the history of NASCAR, everything from how the sport began when moonshine runners were trying to outrun IRS agents in the 1930s and how NASCAR president Bill France started out with \$75 and a dream, all the way up to the beloved driver Dale Earnhardt Sr.'s final breath. Menzer concluded that while the unique sport of NASCAR started with redneck roots and initial profits of \$220, it has evolved into a corporate giant that has stayed true to its humble country beginnings.



Another scholar, Varvus (2007), examined the influence of the phrase “NASCAR dad,” which was especially prevalent in the 2004 presidential campaign season. Varvus’ study examined TV news stories that used the term “NASCAR dad” between January 1, 2002, and April 1, 2005. After ruling out stories with less than 100 words and repeated stories, she discovered that there were 76 transcripts of news segments using the popular phrase during that time. Interestingly, 50 of the 76 segments aired were during the 2004 presidential campaign season. She concluded that the phrase was a form of emotional branding used to distinguish white, blue-collared Southern men between the ages of 18-49 who tended to support the Republican party.

In another study Rotthoff, Depken, and Groothuis (2014) examined what factors influence broadcasters to mention NASCAR sponsors. The study identified a list of drivers from 2000-2007 with a season-long VTOC, or video time on camera, valued at \$38.6 million on the average. The scholars found that many factors contributed to a broadcaster’s decision to mention NASCAR sponsorships, including laps led, post wins, past experience, finishing races, and even being the son of a former driver.

Meanwhile, Dodds and DeGaris’ study (2011) examined “mobile marketing and its relationship to fan avidity, event enjoyment and purchase intent for sponsors’ products” (p. 64). A national telephone survey of 1,000 NASCAR fans was conducted. The survey revealed a strong correlation between a positive fan experience and sponsorship; for instance, Dodds and DeGaris (2011) reported that 73.7% of race attendees said that sponsors made attending the race more enjoyable. Another 79.9% of fans surveyed in Dodds and DeGaris’ study reported that when attending NASCAR races, they visit sponsors’ exhibits pre-race.

In another article, Cadwallader, Boyd, and Thomas (2012) pointed out “fan equity is driven by relationship equity created through relationship-building activities generated by the NASCAR governing body, sponsors, tracks, teams, drivers, the media, and fans themselves” (p. 193). The scholars said that it is very important for the different entities in this industry to work together to provide an invaluable and memorable experience for the fans. The scholars concluded by reiterating the importance of NASCAR teams, drivers, and sponsors’ collaborative efforts and stated that fans are ultimately “a product to be purchased” (p. 206).

These sources provide valuable insight to the history of NASCAR over the years and the importance of the sponsor-fan relationship. However, none of these sources has looked into how strongly fans recognize driver sponsorship. This paper will examine that topic.

Methodology

The researcher studied NASCAR fans’ recognition of driver sponsorship by using survey research. A 20-question survey was distributed through Samford University’s Qualtrix platform. Only those ages 19 and older were invited to participate. Participation was voluntary, and all answers were anonymous.

The survey was distributed to NASCAR fans via two popular online NASCAR fan websites: NASCAR Nation and Racing Forums, available at <http://nascarnation.us/> and <http://racing-forums.com>. It was also posted to the “Nascar Fans” page on Facebook. Participants were asked to forward the survey to other interested NASCAR fans.

Participants were asked various questions to determine their level of loyalty to the sport and how well they know the sponsors. There were basic questions that asked how many races fans had attended in the last 12 months and how many races they’d watched on television in the last 12 months. Then the questions examined how much fans knew about sponsorship – for example, one question asked which company is the primary sponsor for Kyle Busch in the

NASCAR Sprint Cup Series. For questions such as that one, participants had to identify the appropriate sponsors from a list provided. The survey concluded with questions about which NASCAR-sponsored services and products the fans purchased and how much money they spent per year on those services and products.

The survey was ultimately designed to determine how well fans knew NASCAR drivers' sponsors and whether they bought products based on sponsorship of these drivers.

Copies of the recruitment script and survey are attached.

Results

The survey reached a total of 88 participants. Ages varied across the board. There were 11 participants, or 12.5%, in the 19-24 category; 22 participants, or 25% in the 25-34 category; 14 participants, or 15.9%, in the 35-44 category; 12 participants, or 13.6%, in the 45-54 category; 16 participants, or 18.2% in the 55-64 category; and 13 participants, or 14.8%, that were ages 65+. The mean age of participants was 43.

The ages were somewhat surprising. The researcher expected that the majority of participants would fall into the 55 and up categories since that seems to match the age of the crowd at various NASCAR races.

The largest category was in fact 25-34 years old, which surely is good news for NASCAR.

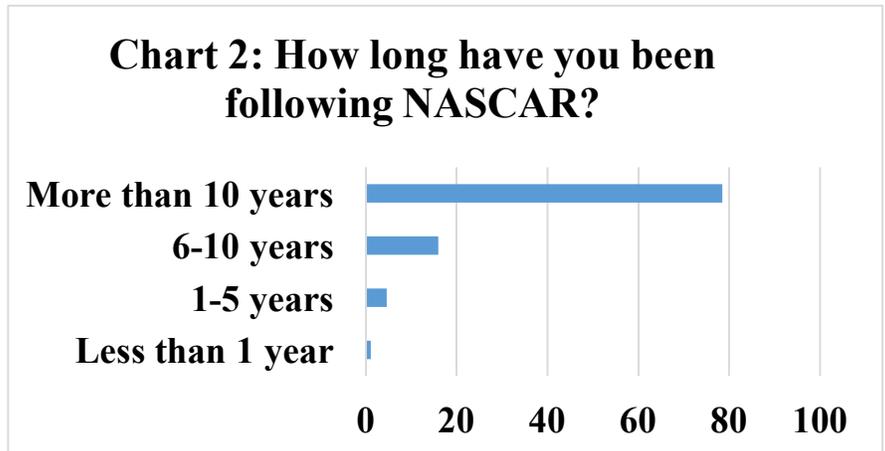
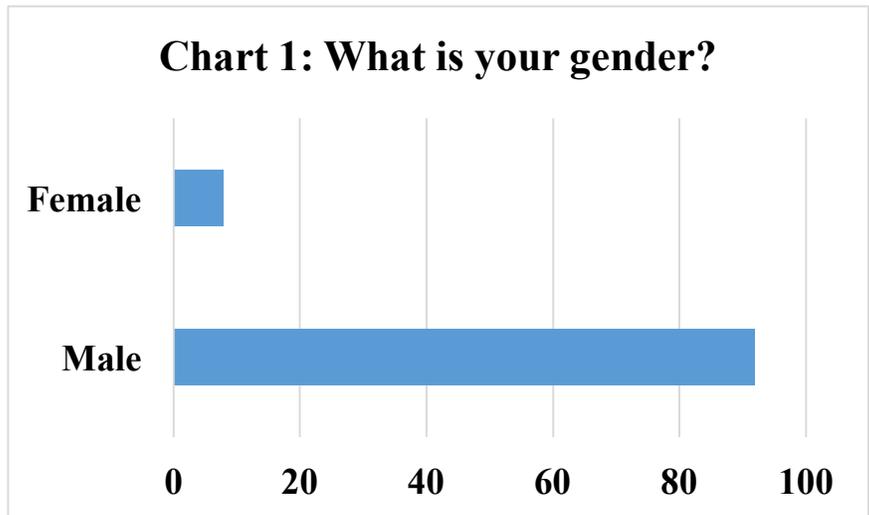
Of the 88 participants, the gender breakdown was 81 males, or 92.05%, and 7 females, or 7.95%, as seen in Chart 1.

These results confirmed what is already known about the sport of NASCAR in general—it's male dominated. Both the Camping World Truck and XFINITY Series are solely made up of male drivers. Danica Patrick is the only female driver in the Sprint Cup Series

(Weaver, 2016). Thus, the male-dominated statistics like these aren't unusual.

The majority of the crowd surveyed has been following NASCAR for a long time, as conveyed by Chart 2. Sixty-nine participants, or 78.41%, have been following NASCAR for more than 10 years. 14 participants, or 15.91%, have been following NASCAR for 6-10 years. Only four participants, or 4.55%, have been following the sport for 1-5 years, and 1, or 1.14% of participants, has been following the sport for less than a year.

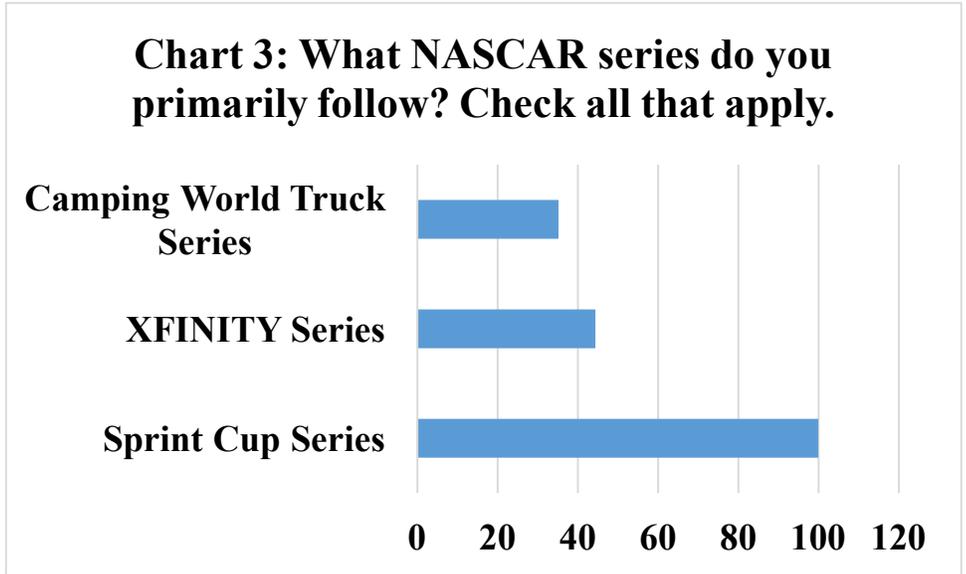
These numbers show that the participants have a long-



standing history of following the sport. They could also be a fairly strong indicator of why many participants chose the correct sponsors for drivers in the questions that appeared later in this survey.

The participants were then asked what NASCAR series they primarily followed. The three choices were the Sprint Cup Series, XFINITY Series, and Camping World Truck Series. Participants were told to check all that apply.

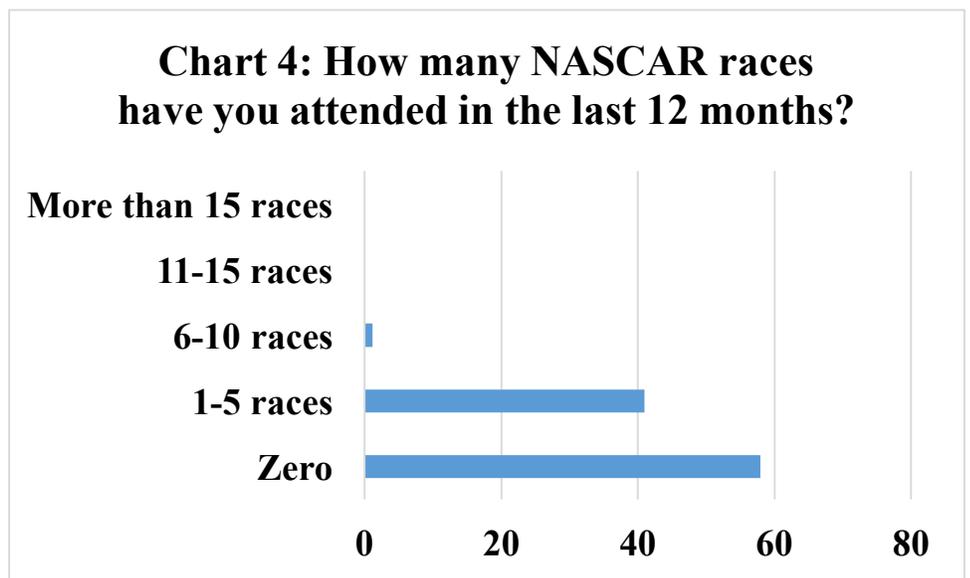
Not surprisingly, 100% of the participants chose Sprint Cup; 39 participants, or 44.32%, chose XFINITY, and 31 participants, or 35.23%, chose Camping World Truck. NASCAR series follow a certain ranking order; most drivers start out on the truck level, unless they are exceptionally talented, and work their way to XFINITY and then Sprint Cup (“NASCAR’s



different series,” 2015). It makes sense that the Sprint Cup Series, which boasts the best of the best, would be the most popular. The results are shown in Chart 3.

The next set of questions that participants were asked covered race attendance and television viewership. Fans were asked how many NASCAR races they have attended in the last 12 months. 51 participants, or 57.95%, marked zero, meaning they hadn’t been to a single race in the last year. 36 participants, or 40.91%, responded saying they’d been to 1-5 races in the last year. 1.14%, which was only 1 participant, responded that he’d been to between 6-10 races in the last year. Lastly, no one had been to 11-15 races or more than 15 races.

These results, shown in Chart 4, demonstrate that the majority of fans who responded to this survey do not attend races. Race attendance has been an issue in NASCAR for years, so these numbers aren’t all that surprising (Gluck, 2016). There are many factors that could explain this: cost of tickets, cost of travelling to the track, and the fact that it’s more convenient to watch races on television.

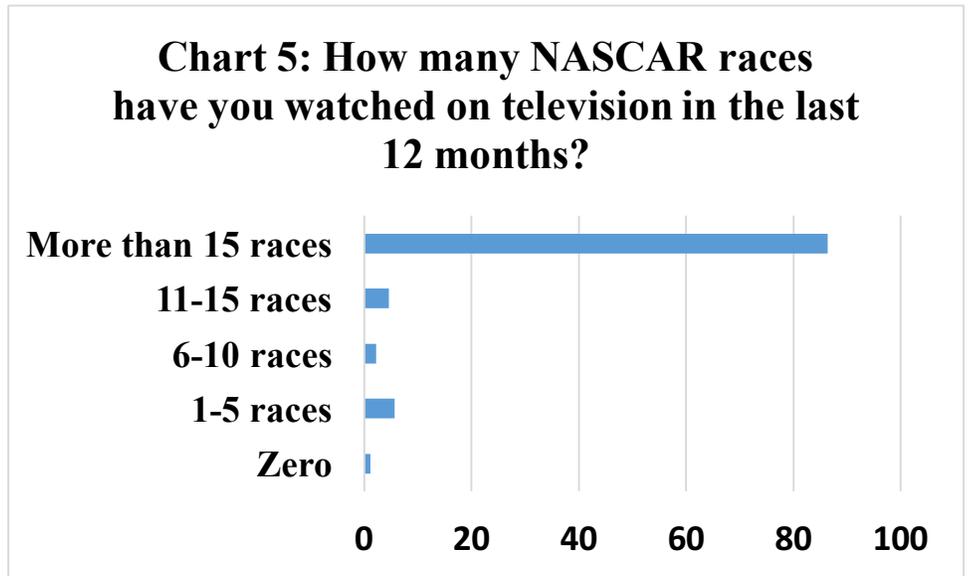


The next question

asked participants how many NASCAR races they'd watched on television in the last 12 months, and the answers were true to expectation. 86.36%, which translated to 76 of the 88 participants, had watched more than 15 races on television in the last 12 months. That number is pretty astonishing when you take into account that this year's Sprint Cup season had 39 races over a 10-month period ("NASCAR announces 2016 Sprint Cup Series schedule," 2015). Clearly a large number of NASCAR fans follow the racing season closely on TV.

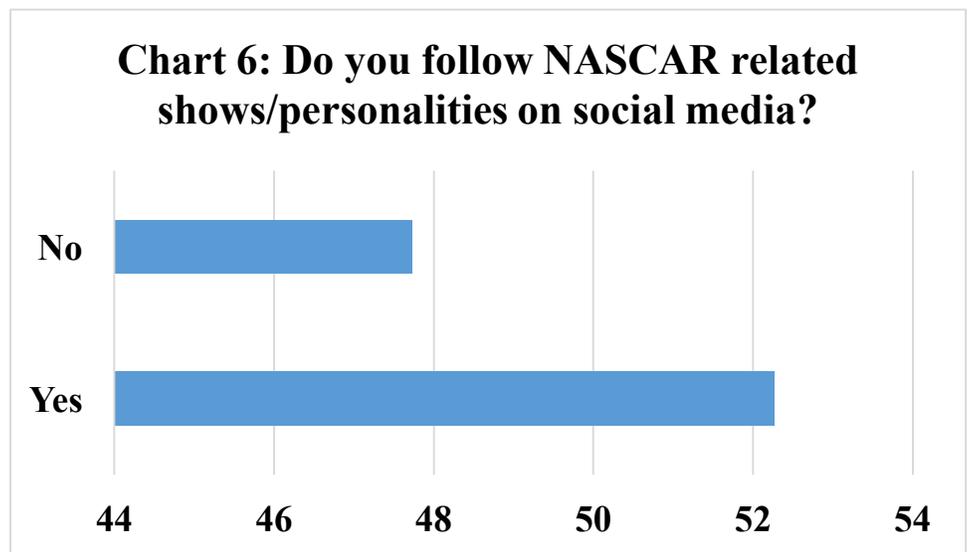
Four participants, or 4.55%, had watched 11-15 races in the last 12 months; 2 participants, or 2.27%, had watched 6-10 races in the last 12

months; 5 participants, or 5.68%, had watched 1-5 races in the last 12 months, and only 1.14% of participants (which translated to one participant), didn't watch any races in the last 12 months. Results are shown in Chart 5.



In an effort to see if social media had an influence on fans and knowledge of driver sponsorships, participants were asked if they follow NASCAR-related shows/personalities on social media. If the participants responded yes, they were asked to list their top 1, 2, or 3 favorite NASCAR-related social media. As seen in Chart 6, 46, or 52.27% of participants, responded yes and 42 participants, or 47.73%, responded no.

Of the 46 participants who responded yes, the sites that came up most frequently were Racing Forums, Jayski, Facebook, and Twitter. These responses demonstrated that the following drivers were most popular on social media: Jimmy Johnson, Brad Keselowski, Kevin Harvick, and Dale Earnhardt Jr.



The information from this question was then used to draw conclusions in the follow-up question, which asked participants to list their top 1, 2 or 3 favorite drivers from any series and

their sponsors in the space provided. Since there were too many drivers and sponsors named to list here, three of the most popular drivers will be examined: Jimmy Johnson, Brad Keselowski, and Kevin Harvick.

Of the 15 participants who listed Jimmie Johnson as their top first, second, or third favorite driver, all 15 of them correctly listed Lowe's as his sponsor. Of the 11 participants who listed Brad Keselowski as their top first, second, or third favorite driver, all 11 of them correctly listed Miller Lite as his sponsor. Finally, of the 15 participants who listed Kevin Harvick as their top first, second, or third favorite driver, all 15 correctly listed at least one or more of his primary sponsors (which include Jimmy John's, Busch Beer, Outback Steak House, and Budweiser).

It's important to note that all of the drivers listed above—Johnson, Keselowski, and Harvick—have multiple primary sponsors and multiple associate sponsors. For the purpose of analysis, however, participants weren't asked to list all of the primary or associate sponsors. The goal was to see if they could list any of them correctly, which all did very well. Clearly, fans of these popular drivers are taking in the sponsorship message.

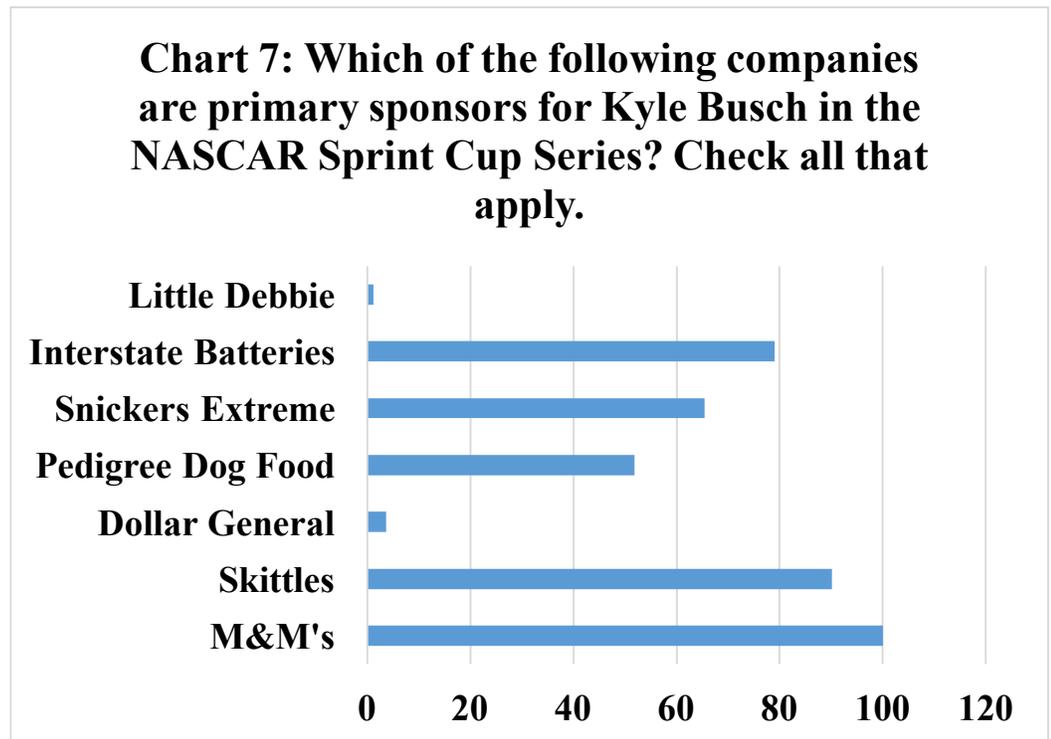
At this point in the survey there was a shift in the questions. Participants were asked to answer questions pertaining to driver sponsorship based on which NASCAR series they primarily follow, as listed in response to question 4. This means that participants who responded that they primarily follow the Sprint Cup Series would only be asked questions pertaining to driver sponsorship in that series. If participants responded that they primarily follow the XFINITY Series or Camping World Truck Series, they would only be asked questions about driver sponsorship in those two series. Participants were asked to answer these questions honestly based on their own knowledge, without consulting outside sources.

The first section of questions in this driver-sponsorship identification section related to the Sprint Cup Series. While 88 participants, or 100%, said they follow the Sprint Cup Series, only 81 of those participants answered the following questions.

The first question asked which companies on a given list are primary sponsors for Kyle Busch in the NASCAR Sprint Cup Series.

Participants were asked to check all that applied from a list of seven sponsors, where five of them were correct and two were incorrect. The correct answers were M&M's, Skittles, Pedigree Dog Food, Snickers Extreme, and Interstate Batteries. As seen in Chart 7, the majority of participants answered correctly.

All of the participants correctly chose M&M's; 73 participants, or 90.12%,



correctly chose Skittles; 64 participants, or 79.01%, correctly chose Interstate Batteries; 53 participants, or 65.43%, correctly chose Snickers Extreme, and 42 participants, or 51.85%, correctly chose Pedigree Dog Food. Only 3 participants, or 3.70%, chose Dollar General and 1 participant, or 1.23%, chose Little Debbie, both of which were incorrect.

Clearly, M&M's is a winner—everyone correctly identified M&M's. They are colorful, delicious, and have an attractive candy spokesman, all of which are attention getters. They're inexpensive, and likely every survey participant has bought and eaten them. Participants likely have had many opportunities to eat them—and associate them with Kyle Busch. Additionally, Busch's car has an eye-catching paint scheme that features the M&M's logo and that is sure to capture fans' attention.

Virtually every participant named Skittles, which has the advantage of looking a lot like M&M's, and of being an inexpensive buy. Repeated opportunities to purchase and eat the ubiquitous candy offer many chances to recall the NASCAR link.

The next most-often mentioned sponsor was Interstate Batteries, with about 80% recalling this sponsor. This product broke the pattern of the M&M's and Skittles candy; a car battery is not a frequent purchase, nor is it inexpensive. It's not very likely that everyone has bought one recently and could "connect" with it. Thus, this result indicates that the sponsorship has done its job; it has planted the name in viewers' minds. Next time they need a car battery, NASCAR fans might think of Interstate Batteries first.

Snicker's Extreme was next in line with sponsorship recognition. Like the other candies mentioned, Snicker's Extreme is cheap, tasty, and can be purchased where most food items are sold. A little more than half of the participants recognized it as a sponsor, so the company is doing a decent job connecting the product to NASCAR fans.

About half of the participants recognized Pedigree as a sponsor. This product is unique among Kyle Busch sponsors in that it's only applicable to dog owners, so non-dog owners might not recognize the tie to Kyle Busch. Another consideration is that there are many different types of dog foods on the market, and Pedigree is a name brand. People who have a dog or even multiple dogs might prefer a cheaper kind of dog food, such as Ol' Roy, since this is a frequent purchase. It could also be that nobody watching the race personally consumes Pedigree, so products that fans don't personally taste might not forge as strong a bond.

Only three participants chose Dollar General and one person chose Little Debbie, both of which are incorrect. Dollar General was a teaser, as it is the primary sponsor of NASCAR Sprint Cup Series driver Matt Kenseth. Little Debbie was also a trick answer; like M&M's and Snickers, it is an inexpensive, easily accessible snack food, but nevertheless has no sponsorship in any of the NASCAR series to date.

The next question asked which of the companies on a list are primary sponsors for Dale Earnhardt Jr. in the NASCAR Sprint Cup Series. Participants were asked to check all that applied from a list of seven sponsors, where four were correct and three were incorrect. The correct answers were Nationwide Insurance, Mountain Dew, TaxSlayer.com, and Axalta Coatings Systems. As seen in Chart 8, the majority of participants answered correctly.

77, or 95.06% of participants, correctly chose Nationwide Insurance; 73, or 90.12% of participants, correctly chose Mountain Dew; 47, or 58.02% of participants, correctly chose Axalta Coatings Systems, and 43, or 53.09% of participants, correctly chose TaxSlayer.com. Very few of the participants answered incorrectly; 9, or 11.11% of participants, chose Pepsi; only 1, or 1.23% of participants, chose State Farm, and 1, or 1.23% of participants, chose TurboTax, all of which were incorrect.

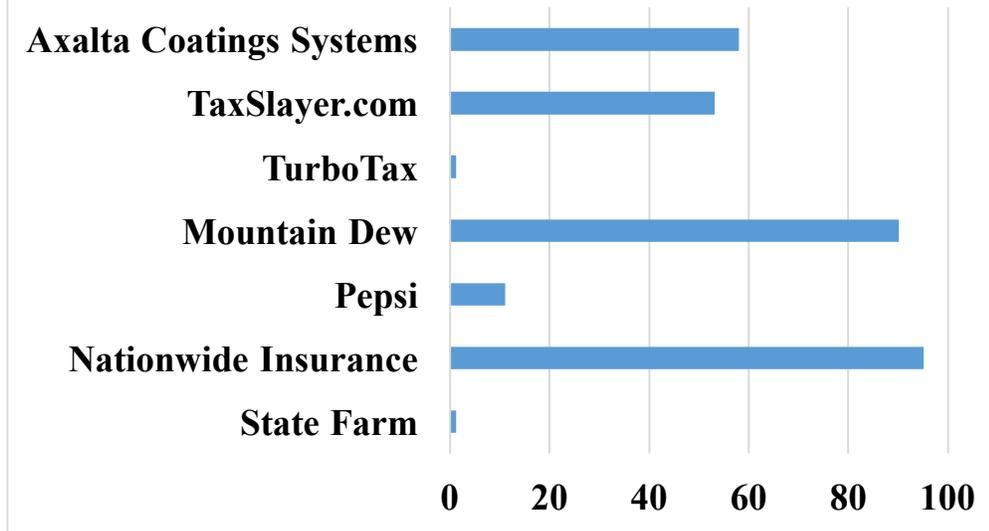
These answers, as with the previous answers, mixed big-ticket purchases with everyday products that might connect on a daily trip to the vending machine. The one everyday type of purchase, Mountain Dew, was correctly chosen about 90% of the time. But the other products in the list showed that the connection didn't come from buying an inexpensive product regularly, such as with Kyle Busch's sponsors, M&M's and Skittles, and with Dale Earnhardt Jr.'s sponsor, Mountain Dew.

Like Interstate Batteries for Kyle Busch, Nationwide Insurance is not an everyday, cheap purchase. However, like Interstate Batteries, it is an auto-related purchase, so the NASCAR link is a strong one, as race fans will be thinking about cars and matching the logo to the need for auto insurance.

The same is true for Axalta Coatings Systems, a company that specializes in coatings for light and commercial vehicles ("Customers," n.d.). This service is expensive, and it's not something you purchase very often. However, nearly 60% of fans associate that name with Dale Earnhardt Jr., so clearly the message is getting across. This could be because his name is so recognizable in the sport and easy to remember, so when fans think of Axalta Coatings Systems they associate it with him.

TaxSlayer.com had much less recognition than Nationwide Insurance and Mountain Dew. One possible influencer could be competitors, such as TaxAct, who is a partner for Danica Patrick in the NASCAR Sprint Cup Series. With other services out there, TaxSlayer.com might not be the most popular choice. However, a more likely possibility is the fact that this survey was distributed in the month of October, and people pay taxes in the spring. It's really not relevant this time of year. Additionally, it's a once-a-year purchase for something that makes a lot of people unhappy, which could also influence the lack of connection.

Chart 8: Which of the following companies are primary sponsors for Dale Earnhardt Jr. in the NASCAR Sprint Cup Series? Check all that apply.



Pepsi was included as a teaser choice since it falls under the PepsiCo company, which is also home to Mountain Dew (“Top global brands,” n.d.). There seems to be a pattern of soft drink companies sponsoring NASCAR drivers, most likely because of the factors mentioned earlier; accessibility, low price, taste, and everyday use. Pepsi might want to follow the lead of other competitors by sponsoring a driver in the near future. If 11.11% of the participants already thought Pepsi was sponsoring Dale Earnhardt Jr., sponsorship may not be a bad move.

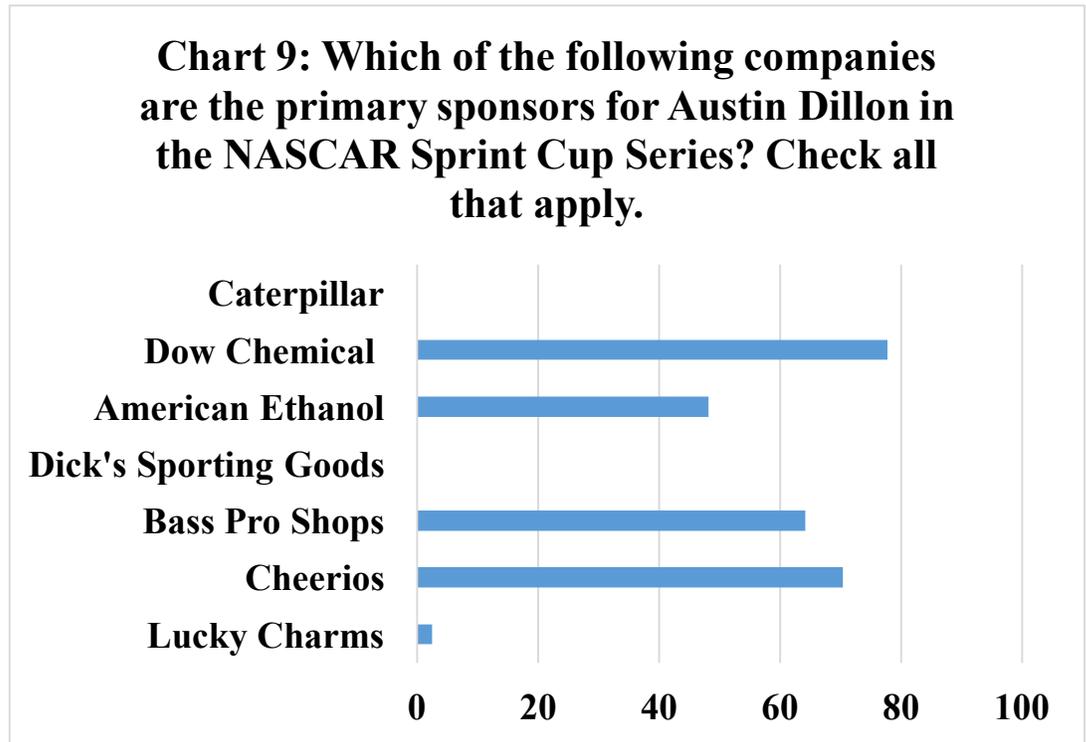
TurboTax and State Farm were also teaser choices that only one person each answered incorrectly. These tax and insurance companies were listed amongst the other competitors, such as Nationwide and TaxSlayer.com, in an effort to throw people off. The fact that only one person each chose these is a testament that sponsorship at large is very effective.

In the last NASCAR Sprint Cup Series question,

participants were asked which of the companies in a list were primary sponsors for driver Austin Dillon. Participants were asked to check all that applied from a list of seven sponsors, where four of them were correct and three were incorrect. The correct answers are Cheerios, Bass Pro Shops, American Ethanol, and Dow Chemical. As seen in Chart 9, the majority of participants answered correctly.

Sixty-three, or 77.78% of participants, correctly chose Dow Chemical; 57, or 70.37% of participants, correctly chose Cheerios; 52, or 64.20% of participants, correctly chose Bass Pro Shops and 39, or 48.15% of participants, correctly chose American Ethanol. Very few of the participants answered incorrectly; 2, or 2.47% of participants, incorrectly chose Lucky Charms and none of the participants chose Dick’s Sporting Goods and Caterpillar, which were also incorrect.

The largest number of correct answers was once again for the everyday type of product, Cheerios. Bass Pro Shops seemed to connect with the “NASCAR dad” type—the blue-collar worker who is likely to love hunting and fishing as much as auto racing. Dow Chemical is a huge leader in the automotive industry, and it seems clear here that most fans saw that connection. With a wide range of products, including adhesives, primers, cleaners, fluids, lubricants, and more, Dow is a name that many in the NASCAR community have come to know and trust (“Products,” Dow, 2016).



American Ethanol, a huge producer of ethanol fuel in NASCAR, didn't seem to fare as well when it came to sponsorship recognition. Of all the correct answers on the sponsorship recognition section, this was the only sponsor to receive below 50% recognition. This is rather interesting since "every lap of every NASCAR Camping World Truck Series, NASCAR XFINITY Series, and NASCAR Sprint Cup Series race is fueled by Sunoco Green E15—a fuel that includes 15% renewable American Ethanol" ("American Ethanol—Powering the NASCAR fan experience," 2016). It is possible that the average NASCAR fan is not familiar with the environmental benefits of ethanol and thus didn't distinguish American Ethanol as a sponsor of Austin Dillon. Another possible factor is that Austin Dillon is a younger driver in the sport and hasn't hit the level of success like Dale Earnhardt Jr., for example.

Caterpillar was thrown in as a trick answer choice because Caterpillar is the primary sponsor for Austin Dillon's Richard Childress Racing teammate Ryan Newman. Clearly the participants who answered this question are loyal to Austin Dillon, because no one chose that answer. The same goes for Dick's Sporting Goods; some of their products are similar to those available at Bass Pro Shops so that's why it was listed as a choice, but 100% of fans recognized that Dick's doesn't sponsor Austin Dillon. Lucky Charms was also thrown in as a trick answer choice, as it is a similar product to Cheerios. Again, loyal fans in general knew the truth about Cheerio's sponsorship.

Overall, the researcher was surprised at how strong the fan-product connection was. The fans often recognized the products. An overall breakdown of fan-sponsorship recognition is seen in Chart 10, starting with the highest percentage first.

Chart 10: Fan-Sponsorship Recognition Breakdown

| Sponsor | Recognition of Sponsor (%) | Category of Sponsor Recognition |
|----------------------|-----------------------------------|--|
| M&M's | 100% | Very High |
| Nationwide Insurance | 95.06% | Very High |
| Mountain Dew | 95.06% | Very High |
| Skittles | 90.12% | Very High |
| Interstate Batteries | 79.01% | High |
| Dow Chemical | 78.78% | High |
| Cheerios | 70.37% | High |
| Snicker's Extreme | 65.43% | Fairly High |
| Bass Pro Shops | 64.20% | Fairly High |
| Axalta | 58.02% | Fairly High |
| TaxSlayer.com | 53.09% | Near Middle |
| Pedigree Dog Food | 51.85% | Near Middle |
| American Ethanol | 48.15% | Near Middle |

These numbers are very revealing about the demographic who completed the survey. The three most-recognized sponsors included everyday purchase products (M&M's and Skittles) and

one auto-related product (Nationwide Insurance). Clearly sweet treats stick in the fans' minds, and likely frequent purchases reinforce the memory.

Only one sponsor had less than 50% recognition, American Ethanol, perhaps a surprise due to pushing ethanol use at NASCAR. This product did a decent amount of connecting at 48.15%, but it's got a long way to go to make the connection that M&M's has made.

At least half of the participants recognized the other products, and many more for many products. Clearly the partnerships do stick in race fans' heads. And almost no one was fooled by the false products. When fans watch NASCAR, they actively remember the sponsorships they see, which hopefully translates to more money. That's the goal of sponsorship.

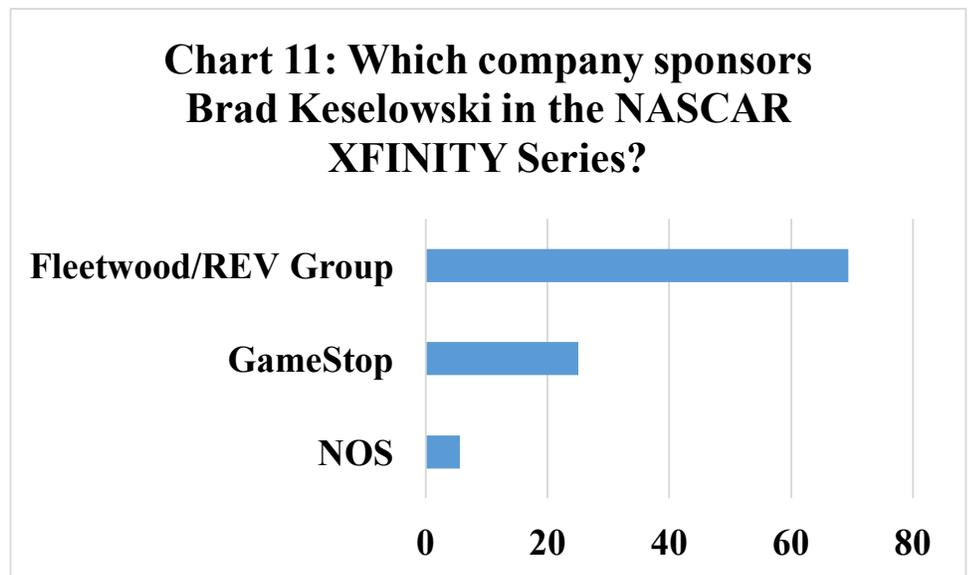
The next series of questions in the driver-sponsorship identification section related to the NASCAR XFINITY Series. Thirty-six participants who said they primarily follow NASCAR answered these questions.

It's important to note that in this section that there were two incorrect choices and one correct choice listed for each question. However, each of the drivers mentioned can have multiple primary and associate sponsors that alternate per race each week. For example, Armour Vienna Sausage could be the sponsor for the Kevin Harvick in the NASCAR XFINITY Series at Talladega, but Busch Beer could be sponsoring him at Martinsville. The researcher simply chose one of the many possible correct sponsors per question to see how well the participants knew the driver.

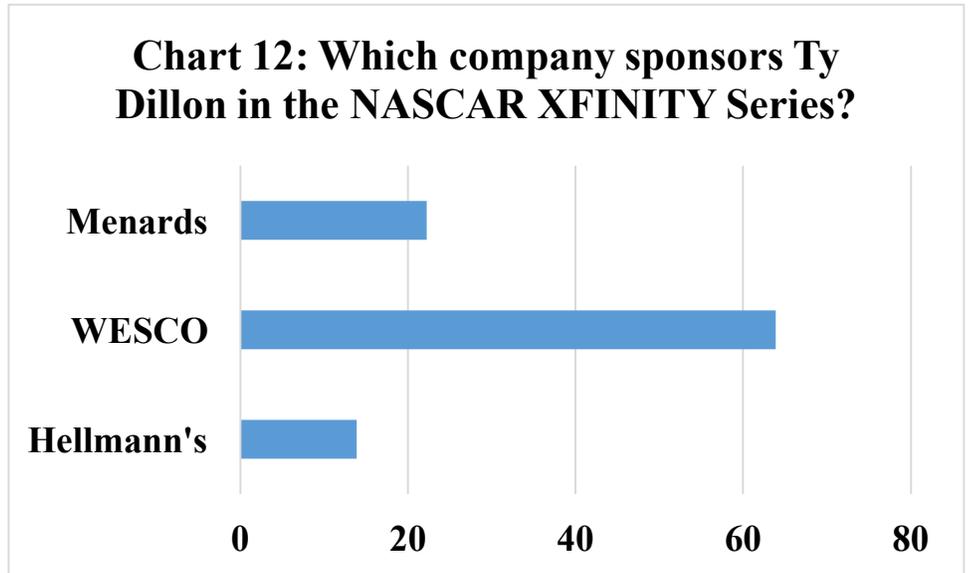
The first question asked which company sponsors Brad Keselowski in the NASCAR XFINITY Series. Participants were asked to choose the correct answer from a list of three choices, where one choice was the correct sponsor and the other two were incorrect. As Chart 11 demonstrates, the majority of participants chose Fleetwood/REV Group, which was the correct choice.

Twenty-five, or 69.44% of participants, correctly answered this question. 9, or 25.00% of the participants answered GameStop and 2, or 5.56% of participants, guessed NOS, both of which are incorrect.

Interestingly, one-fourth of respondents incorrectly guessed GameStop. Only one incorrect answer in the Sprint Cup Series questions broke into the double digits, when 11.1% guessed incorrectly that Pepsi sponsored Dale Earnhardt Jr. Perhaps both brands could capitalize on advertising and sponsorship with NASCAR. GameStop is linked to personal gaming, likely an everyday product used by some fans.



The next question in this section asked which company sponsors Ty Dillon in the NASCAR XFINITY series. Once again participants were asked to choose the correct answer from a list of three choices, where one choice was the correct sponsor and the other two were incorrect. As Chart 12 demonstrates, the majority of participants chose WESCO, which was the correct choice.



Twenty-three, or 63.89% of participants, correctly answered this question. 8, or 22.22% of participants, chose Menards and only 5, or 13.89% of participants, chose Hellmann's, both of which are incorrect.

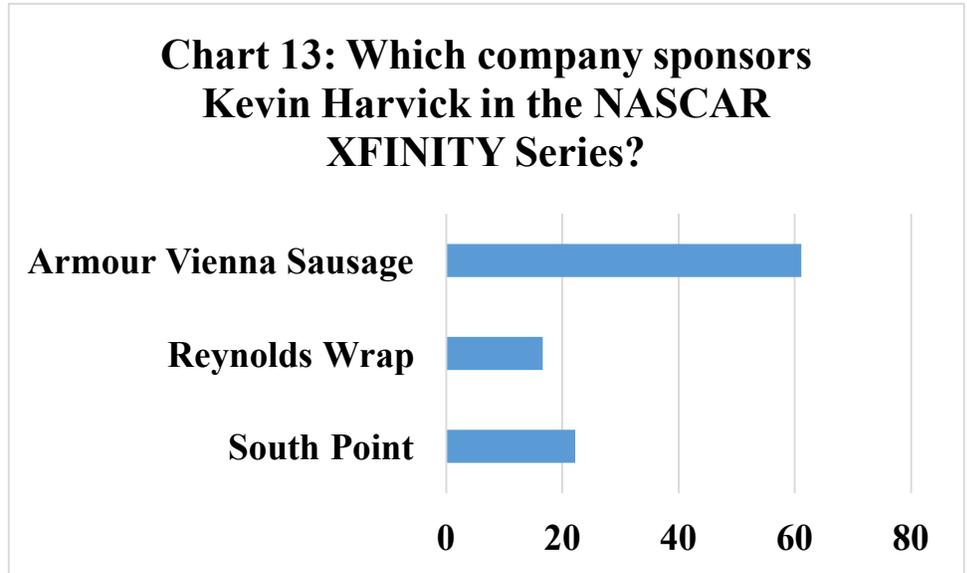
Once again, unlike the Sprint Cup Series, the incorrect guesses drew a higher percentage. Nearly 25% of respondents picked Menards, while nearly 14% picked Hellmann's. When forced to guess, fans opted for home improvement/DIY giant Menards and an everyday item like Hellmann's, which are brands they most likely knew from daily life. It is also possible that Menard's stumped people because it is the major sponsor for Paul Menard, Ty Dillon's teammate at Richard Childress Racing. A fan who doesn't follow Ty Dillon closely could have easily gotten the two confused.

WESCO, the correct sponsor, is an electrical distributor that specializes in lighting, distribution equipment, control and automation equipment, and a host of other products and services ("Products," WESCO, 2016). It's interesting that this company, which is more expensive and not considered an everyday purchase, gained higher recognition. It's clear that Ty Dillon has a loyal group of fans who easily picked this sponsor, despite it being a service that they may infrequently or never take advantage of.

For the final question in this section, participants were asked which company sponsors Kevin Harvick in the NASCAR XFINITY Series. Of the three choices given, there was one

correct answer and two incorrect options. Chart 13 demonstrates that once again the majority of participants chose the correct answer, which is Armour Vienna Sausage.

61.11% of participants correctly answered this question, which was 22 of the 36 participants. 8, or 22.22% of participants, chose South Point and 6, or 16.67% of participants, chose Reynolds Wrap, both of which are incorrect.



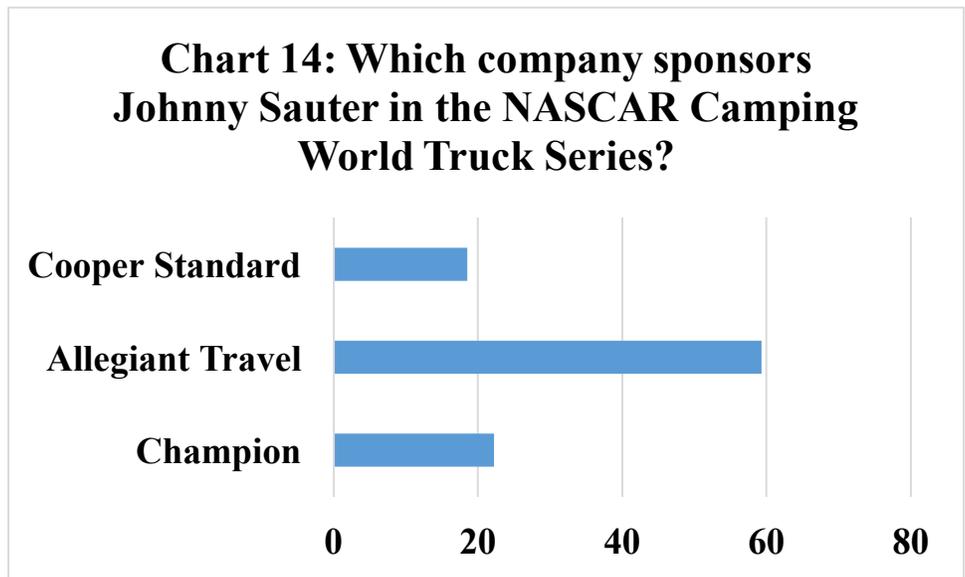
In this scenario there are two everyday buys, Armour Vienna Sausage and Reynolds Wrap. It is interesting that the lower numbers in the Sprint Cup Series were around 60%, but here the highest knowledge of the correct sponsors was in the 60-70% range. Clearly here the products clearly weren't as closely associated with driver sponsorships; this is perhaps due to rotating sponsorship. Still, given the changing sponsorships, it's impressive that 60% to 70% still know the sponsor. However, to reach the high recognition seen in the Sprint Cup Series, sponsors need to make their presence clearer and more stable.

South Point was thrown in as a teaser option as it is the primary sponsor for Richard Childress Racing driver Brendan Gaughan. Kevin Harvick is a former Richard Childress Racing driver, so it's possible that fans might have gotten the sponsorships confused. Reynolds Wrap was also an incorrect option as it is the primary sponsor for JD Motorsports driver Eric McClure.

The last section of driver-sponsorship questions related to the Camping World Truck Series. There were 27 respondents to these questions.

The first question asked which company sponsors Johnny Sauter in the NASCAR Camping World Truck Series. As was the case with the XFINITY Series questions, there was one correct choice and two incorrect choices. The majority of participants correctly chose Allegiant Travel, as shown in Chart 14.

59.26% of participants correctly answered this question,



which was 16 of the 27 participants. 6, or 22.22% of participants, chose Champion and 5, or 18.52% of participants, chose Cooper Standard, both of which were incorrect.

Interestingly, the majority of fans chose Allegiant Travel, which is not an inexpensive, everyday purchase. However, for those who want to make the most of their travel it would make sense to use Allegiant for booking flights, hotel rooms, and car rentals. It could be that frequent NASCAR race attendees use the service to arrange transportation to and from races. Also worth noting is that 12 of the 16 fans who correctly answered Allegiant Travel have been following the sport for more than 10 years, so there is obviously a strong connection for what can be categorized as “very high” fans who have kept up with the sport for a long time.

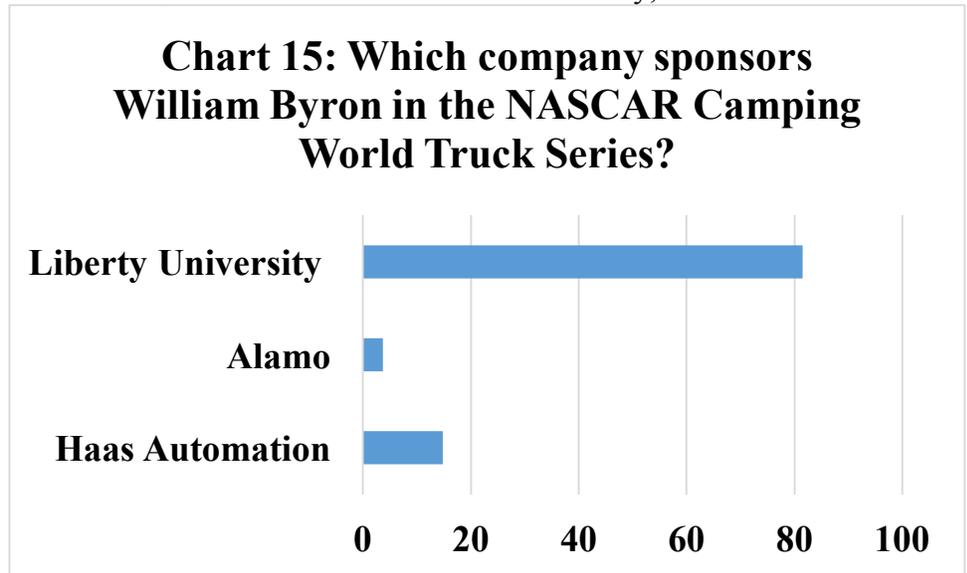
Champion, an athletic apparel line, received almost a quarter of responses. It's interesting that this many people incorrectly associated this company with Johnny Sauter and the NASCAR Camping World Truck Series. Just like 22.22% of participants wrongly chose Menards as a sponsor for Ty Dillon in the XFINITY Series questions, 22.22% of people thought Champion was the correct sponsor for Johnny Sauter for some reason. Champion might need to get in the game with NASCAR.

The same goes for Cooper Standard, another incorrect answer, which was chosen by 5 of the 27 participants. Cooper Standard is a leader in the automotive services industry, which obviously struck a chord with racing fans (“Products and innovations,” 2016). It would make sense for a company like this to get involved with advertising and sponsorship if this many people thought it already was.

The second question asked which company sponsors William Byron in the NASCAR Camping World Truck Series. With one correct choice and two incorrect choices, the majority of fans correctly chose Liberty University. Results are shown in Chart 15.

81.48% of participants correctly answered this questions, which equates to 22 of the 27 participants. 4, or 14.81% of the participants, chose Haas Automation and 1, or 3.70% of participants, chose Alamo, both of which are incorrect.

It is interesting that the majority of respondents correctly answered this question. Furthermore, it's fascinating that this is the first learning institution and intellectual-based service serving as a sponsor. Liberty University “is the largest private, nonprofit university in the nation, the largest university in Virginia, and the largest Christian university in the world” (“Sponsors,” 2016). This highly conservative school whose mission “is to train *Champions for Christ*” obviously has a strong connection to NASCAR fans (“Sponsors,” 2016). It's not particularly surprising that of the 22 participants who correctly answered this question, 9 of them were in the 25-34 age category, which could very well consist of young adults who have either just graduated from college or been out of college for a few years. Nevertheless, it's clear that

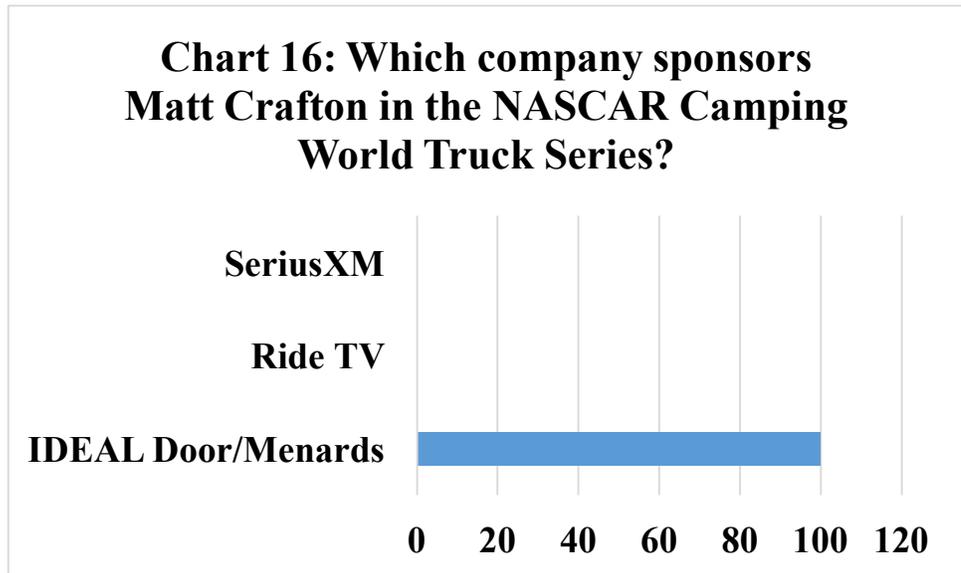


the younger crowd is familiar with Liberty University and William Byron.

Haas Automation is the parent company to Haas Motorsports, a successful race team in NASCAR, and was listed as an incorrect answer choice to throw people off (“Haas Automation Motorsports,” 2016). Surprisingly, nearly 15% of participants chose it, which indicates that the Haas name holds a special weight in this sport. Clearly fans connected with it and falsely assumed that since it was a car company it might be Byron’s sponsor.

Only one participant out of the 27 chose Alamo, which was also incorrect. Alamo is a car rental company, and while it is probably a valuable service to many, it clearly doesn’t hold a lot of weight in the NASCAR fan circle, since only one person thought it was a sponsor (“About Alamo Rent A Car,” 2014). All in all, Liberty University clearly appears to be doing its job in promoting its sponsorship with William Byron well.

The final question in this section asked which company sponsors Matt Crafton in the NASCAR Camping World Truck Series. With one correct choice and two incorrect choices, all of the participants correctly chose IDEAL Door/Menards. Results are shown in Chart 16.



All 27

participants answered this question correctly. None of the participants chose Ride TV or SeriusXM, which were the two incorrect options.

Interestingly, along with M&M’s, this was the only 100% response from participants. Unlike M&M’s, which are inexpensive and frequently purchased, an IDEAL Door is a rare purchase and rather expensive. Menards, however, may be a place fans shop fairly frequently, especially those do-it-yourselfers. “NASCAR dads,” one would imagine, might well be do-it-yourselfers as blue collar workers who need to make ends meet by repairing and improving their home themselves. It’s fascinating, though, that the two 100% responses were such different products, and yet the high recognition bodes well for IDEAL Door/Menards.

The survey concluded with questions about purchasing products based on NASCAR sponsorship. These questions were meant to determine how much money fans are spending on NASCAR-sponsored products and services, which would indicate their recognition of driver sponsorship.

The first question in this section asked fans if they purchase products or services

specifically because the company has a sponsorship in NASCAR. 78 participants answered this question. Of those 78 participants, 34, or 43.59% of participants, responded yes and 44, or 56.41% of participants, responded no. The results are shown in Chart 17.

Even though fans clearly recognized sponsors' brands, not all fans based purchasing decisions on such

sponsorship. It's interesting, however, that more than 40% did see NASCAR sponsorship as a reason to buy products. Surely such a statistic would please sponsors. The 43.59% of participants who bought products or services based on sponsorship are truly hard-core, devoted fans and any sponsor's dream.

For the 56.41% of participants who responded no, the survey ended there. The 43.59% of participants who responded yes, which was 34 participants, were then asked to list which products or services they had purchased. Of those 34 participants who responded yes, 31 actually listed the products or services they have purchased.

The most popular products and services of the responses given included Miller Lite, Interstate Batteries, M&M's, and Bass Pro Shops. Overall the major responses fell under the categories of drinks, food/snacks, auto, and shopping/misc. It was especially interesting to read through the comments and see the ties that participants had to certain drivers, which then influenced their shopping.

For instance, one participant responded, "Busch Beer (Kevin Harvick), Kobalt Tools (Jimmie Johnson) and Mobil 1 (Tony Stewart)." This participant had listed Kevin Harvick and Tony Stewart as two of his favorite drivers in a previous question. Another participant responded that he shops at Lowes over Home Depot, which makes sense because he had listed Jimmie Johnson as one of his favorite drivers in a previous question.

The products listed tallied as follows in Chart 18.

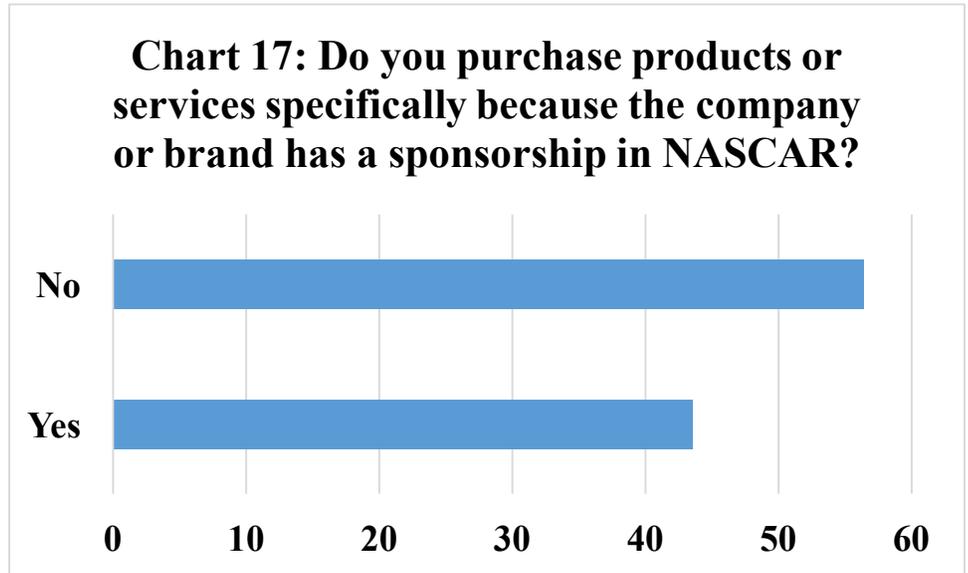


Chart 18: If your answer was yes to the last question, which products or services have you purchased?

| Product | Category | Frequency |
|-------------------------|----------------|-----------|
| Coke | Drinks | 1 |
| Mountain Dew | Drinks | 4 |
| Pepsi | Drinks | 1 |
| 5 Hour Energy | Drinks | 1 |
| Coke Zero | Drinks | 1 |
| Coor's Lite | Drinks | 3 |
| Miller Lite | Drinks | 5 |
| Budweiser | Drinks | 3 |
| Dr. Pepper | Drinks | 1 |
| Busch Beer | Drinks | 1 |
| Monster Energy | Drinks | 1 |
| Subway | Food/Snacks | 2 |
| McDonald's | Food/Snacks | 1 |
| Jimmy John's | Food/Snacks | 2 |
| General Mills Cereals | Food/Snacks | 3 |
| Skittle's | Food/Snacks | 1 |
| Cheez-It | Food/Snacks | 1 |
| Bullseye Barbecue Sauce | Food/Snacks | 1 |
| Hooter's | Food/Snacks | 1 |
| Snickers | Food/Snacks | 1 |
| Smithfield Bacon | Food/Snacks | 1 |
| Nature's Bakery | Food/Snacks | 2 |
| M&M's | Food/Snacks | 7 |
| Shell | Auto | 2 |
| Interstate Batteries | Auto | 3 |
| Mobil | Auto | 1 |
| Peak Motor Oil | Auto | 1 |
| Napa Auto Parts | Auto | 1 |
| Pennzoil Oil | Auto | 2 |
| Nationwide Insurance | Auto | 1 |
| Goodyear Tires | Auto | 1 |
| Chevrolet | Auto | 1 |
| Quaker State | Auto | 1 |
| Energizer | Shopping/Misc. | 1 |
| Tide | Shopping/Misc. | 3 |
| Bass Pro Shops | Shopping/Misc. | 4 |
| Target | Shopping/Misc. | 2 |
| Kobalt Tools | Shopping/Misc. | 2 |
| Great Clips | Shopping/Misc. | 1 |
| Diecasts | Shopping/Misc. | 2 |
| Dollar General | Shopping/Misc. | 2 |
| Menard's | Shopping/Misc. | 2 |
| Sprint | Shopping/Misc. | 1 |
| Home Depot | Shopping/Misc. | 2 |
| Lowe's | Shopping/Misc. | 3 |
| FedEx | Shopping/Misc. | 1 |

As evidenced by the numbers above, there were some leaders within each category. Miller Lite was the clear leader for drinks, which seems to be a NASCAR fan staple. The cheap cost and easy access of Miller Lite is most likely what makes it a frequent purchase above the rest. Mountain Dew was a close second place top finisher in the drinks category, which could also be attributed to the low cost and convenience. It also doesn't hurt that Mountain Dew is the sponsor for Dale Earnhardt Jr., one of the most recognized drivers in the sport. Roughly 90% of fans recognized Mountain Dew as a sponsor.

M&M's came out on top once again as the leader in the food/snacks category. When 100% of fans recognized the tasty treat as a sponsor for Kyle Busch in the Sprint Cup Series, it's clear that this company is doing its job well when it comes to reaching its audience. The candy is cheap, accessible, and enjoyed by many NASCAR fans. Kyle Busch's car also has an eye-catching paint scheme that feature the M&M's logo and it is fairly easy to recognize on the track.

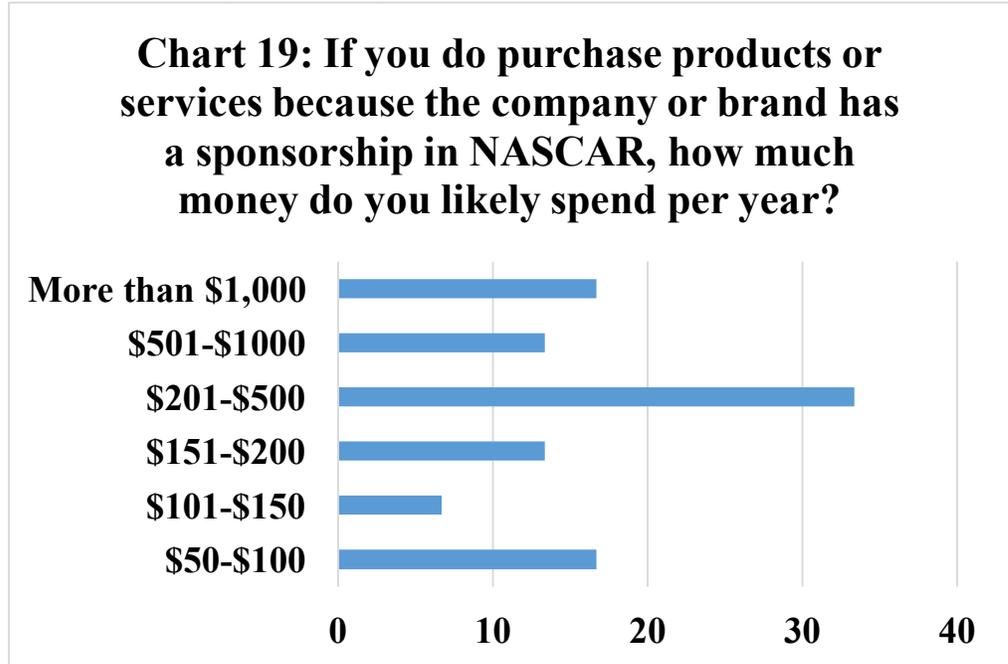
Interstate Batteries was most popular in the auto section. As mentioned previously Interstate Batteries is not an everyday purchase, but nearly 80% of fans recognized the company as the sponsor for Kyle Busch in the Sprint Cup Series, so clearly the marketing efforts have proven to be effective.

Bass Pro Shops was the leader in the shopping/misc. category. 64.20% of fans recognized the hunting/fishing/camping/outdoor giant as a sponsor for Austin Dillon in the Sprint Cup Series. While this percentage may not be as high as some of the others, that's still a large number of people who associate the brand with a well-known driver. Clearly the recognition is paying off for many in the store.

A key takeaway from these results is that fans who purchase products with a NASCAR sponsorship attached do so for that specific purpose; primarily out of loyalty to both the driver and sport, and additionally because there's a level of product satisfaction. One fan captured it best with this statement: "Too many to list. Mostly auto parts/products. All things equal, I choose to support the brands that choose to sponsor the sport over brands that don't."

The final question of the survey asked how much money these fans are likely to spend on NASCAR sponsors per year. Thirty participants answered this question. Results are shown in Chart 19.

Five, or 16.67% of participants, would spend \$50-\$100 per year; 2, or 6.67% of participants, would spend \$101-\$150 per year; 4, or 13.33% of participants, would



spend \$151-\$200 a year; 10, or 33.33% of participants, would spend \$201-\$500 per year; 4, or 13.33% of participants, would spend \$501-\$1,000 per year and 5, or 16.67% of participants, would spend over \$1,000 per year.

The results were interesting for several reasons. First, there was no discrepancy between the participants who spend over \$1,000 and participants who spend between \$50 and \$100. The percentage of participants in both of those categories is 16.67%, which equates to 5 of the 30 participants who answered this question. Neither high end nor low end products were the bigger sellers, although admittedly fans could factor in buying many M&M's or Mountain Dew's to just one big expense, such as a new door or tuition for a class at Liberty University. In other words, low-end purchases might be in very high quantities but lower totals.

Second, an equal number of participants spent between \$151-\$200 and \$501-\$1,000. The percentage for both is 13.33%, which equals 4 participants. In general, these participants were likely buying more than just packages of M&M's or boxes of Cheerios. It seems likely that when fans had an infrequent large expense, they bought based on NASCAR sponsorship.

Finally, it's easy to tell that the majority of participants fell right in the middle by spending between \$201 and \$500. That is a total of 33.33% participants, or 10 of the 30 participants who answered this question. Perhaps a mix of many low-price purchases and the occasional high-priced purchase could total that amount yearly.

Conclusion

The researcher concluded from this survey that there was such high knowledge of sponsorship that it was hard to suggest that there was a "low knowledge" group of fans when it came to sponsorship. High knowledge was pervasive. There was no difference in knowledge of sponsorship across a broad range of variables, including: gender, amount of time following NASCAR, races attended, races watched and social media following.

These findings bring good news to NASCAR, demonstrating that the sponsorship efforts are effectively reaching a multitude of followers across various demographics and backgrounds. In other words, men are not more likely to know more sponsorships than women; a fan who's followed the sport for less than five years knows the sponsors just as well as a fan who has followed the sport for more than ten years, and so on. At the end of the day, a fan is a fan is a fan.

An examination of the evidence suggests a difference between age and knowledge of Sprint Cup sponsorships, with younger fans seeming to have a slight edge over older fans in their ability to recognize sponsorships in the Sprint Cup Series.

Three participants, or 27.3%, of the 19-24 year olds, answered all of the Sprint Cup questions correctly. The 25-34 and 35-44 age groups tied in second place for the number of participants to correctly answer all of the Sprint Cup questions. Both of these groups had 2 participants who knew all of the sponsors for Sprint Cup. For the 25-34 age group, the 2 participants accounted to 9.5% of respondents in that age category and the for the 35-44 age group the 2 participants represented 15.4% of the age bracket.

There were no 100% correct answers in the 45-54 or 65+ age groups, and there was one participant who got all of the correct answers in the 55-64 age group. This indicates that as a whole, younger fans have a slightly better recognition of Sprint Cup sponsorships than older fans. A variety of factors could play into this, including social media use, time spent on the internet, and television viewing. However, statistical testing would be necessary in order to obtain a more conclusive result.

There also seemed to be a difference between fans who follow the Camping World Truck

Series and knowledge of the XFINITY Series. Twelve participants, or 44.45% of participants who primarily follow the Camping World Truck series, correctly answered all of the XFINITY Series Sponsorship questions.

This seems to indicate that fans follow their favorite drivers through each level of NASCAR, meaning that if a fan keeps up with a driver in the Camping World Truck Series, that same fan is going to continue following the driver once he advances to the XFINITY Series. Furthermore, the sponsorship is effective because fans are able to recognize sponsors across different series of the sport. However, statistical testing would be necessary in order to obtain a more conclusive result.

This paper reaffirmed the notion that NASCAR fans are some of the most loyal people in the world of sports, and their recognition of sponsorship is extremely high across the board. The fact that there is no real differentiation in knowledge of sponsorships across multiple variables implies that NASCAR has an effective marketing strategy and takes full advantage of the sponsorship opportunities among the three series.

Of course there is always room for improvement, and as the data demonstrated there were plenty of participants who wrongly associated certain brands with certain drivers. These companies have a prime opportunity to jump in the driver's seat and take advantage of the countless benefits that sponsorship could have on their bottom line.

NASCAR is a multi-million-dollar industry that has come a long way since the bootlegging moonshiner days. The results of this paper indicate that fans in this sport are very committed and pay attention to the sponsorship. Effective marketing on behalf of the sponsors and NASCAR drivers is crucial to the sport's livelihood, and it appears to be on the right track.

Bibliography

- About Alamo Rent A Car. (2014). *Alamo*. Retrieved December 7, 2016, from <http://aboutus.alamo.com/>
- American Ethanol—Powering the NASCAR fan experience. (2016). *American Ethanol Racing*. Retrieved November 30, 2016, from <http://www.americanethanolracing.com/nascar>
- Cadwallader, S., Boyd, T., & Thomas, A. (2012). Affinity and beyond: Leveraging NASCAR's integrated supply network to drive fan equity. *Journal of Marketing Channels*, 19(3), 193–211. doi:10.1080/1046669x.2012.686861.
- Customers. (n.d.). *Axalta Coatings Systems*. Retrieved December 3, 2016, from http://www.axaltacs.com/corporate/en_US/products-and-customers/customers.html
- Dodds, M., & DeGaris, L. (2011). Using mobile marketing to engage NASCAR fans and increase sales. *CH*, 7(1), 63–75. doi:10.4127/ch.2011.0055.
- Explaining the difference in sponsors. (2015, January 5). *NASCAR*. Retrieved November 30, 2016, from http://www.nascar.com/en_us/sprint-cup-series/nascar-nation/nascar-edu/nascar-basic/sponsors-levels-placement-cars.html
- Gluck, Jeff. (2016, July 1). Fans explained why they stopped attending NASCAR races. *USA TODAY*. Retrieved November 18, 2016, from USA TODAY, <https://www.usatoday.com/story/sports/nascar/2016/07/01/nascar-fans-attendance-cost/86587042/>
- Haas Automation Motorsports. (2016). *Haas Automation Inc.* Retrieved December 7, 2016, from <http://motorsports.haascnc.com/>
- Klara, R. (2013, January 28). Your brand here. *Adweek*. Retrieved November 30, 2016, from <http://www.adweek.com/news/advertising-branding/your-brand-here-146795>
- Menzer, J. (2001). *The wildest ride: A history of NASCAR (or, how a bunch of good ol' boys built a billion-dollar industry out of wrecking cars)*. New York: Simon & Schuster.
- NASCAR announces 2016 Sprint Cup Series schedule. (2015, October 26). *NASCAR*. Retrieved December 7, 2016, from http://www.nascar.com/en_us/news-media/articles/2015/10/26/nascar-2016-schedule-sprint-cup-series-daytona-500-homestead.html
- NASCAR's different series. (2015, January 5). *NASCAR*. Retrieved November 30, 2016, from http://www.nascar.com/en_us/sprint-cup-series/nascar-nation/nascar-edu/nascar-basic/nascar-national-series-home-tracks.html
- Products. (2016). *The Dow Chemical Company*. Retrieved December 3, 2016, from <http://www.dowautomotive.com/products/>
- Products. (2016). *WESCO*. Retrieved December 3, 2016, from <http://www.wesco.com/products/>
- Products and innovations. (2016). *Cooper Standard*. Retrieved December 6, 2016, from <http://www.cooperstandard.com/products-and-innovations>
- Rotthoff, K. W., Depken, C. A., & Groothuis, P. A. (2014). Influences on sponsorship deals in NASCAR: Indirect evidence from time on camera. *SSRN Electronic Journal*, 46(19), 2277–2289. doi:10.2139/ssrn.1886985.
- Sponsors. (2016). *William Byron Racing*. Retrieved November 18, 2016, from <http://www.williambyron.com/sponsors/>
- Top global brands. (n.d.). *PepsiCo*. Retrieved December 7, 2016, from <http://www.pepsico.com/Brands/BrandExplorer - top-global-brands>

Vavrus, M. D. (2007). The politics of NASCAR dads: Branded media paternity. *Critical Studies in Media Communication*, 24(3), 245–261. doi:10.1080/07393180701520942.

Weaver, M. (2016, July 18). Danica Patrick is rocking it in NASCAR Sprint Cup Series, but who's got next? *Autoweek*. Retrieved November 18, 2016, from Autoweek, <http://autoweek.com/article/nascar-sprint-cup/women-racing-whos-got-next>

Recruitment Script

Dear NASCAR Fan,

My name is _____ and I'm a _____ major at Samford University. For my _____, I'm examining fan sponsorship recognition in the NASCAR industry. Because you are NASCAR fans yourselves, I'm inviting you to participate in this research study by completing the following survey. The survey will require approximately 10-15 minutes to complete.

You must be 19 years of age or older to take the survey. Your answers are completely anonymous, and of course, participation is strictly voluntary. If you choose to participate in this project, please answer all questions as honestly as possible and complete the survey within the next seven days. Thank you for taking the time to help me. Your participation is greatly appreciated. Feel free to forward the survey to other NASCAR fans you know. Please contact me if you have any questions or concerns.

Sincerely,

NASCAR Survey

1. What is your age?
 2. What is your gender?
 - Female
 - Male
 3. How long have you been following NASCAR?
 - Less than one year
 - 1-5 years
 - 6-10 years
 - More than 10 years
 4. What NASCAR series do you primarily follow? Check all that apply.
 - Sprint Cup Series
 - XFINITY Series
 - Camping World Truck Series
 5. How many NASCAR races have you attended in the last 12 months?
 - Zero
 - 1-5 races
 - 6-10 races
 - 11-15 races
 - More than 15 races
 6. How many NASCAR races have you watched on television in the last 12 months?
 - Zero
 - 1-5 races
 - 6-10 races
 - 11-15 races
 - More than 15 races
 7. Do you follow NASCAR related shows/personalities on social media? (e.g. Groups like FOX SPORTS: NASCAR, NASCAR on NBC, former drivers and current drivers are often active on sites such as Twitter, Facebook, etc.). If yes, list up to your top three favorite NASCAR-related social media sites that you follow in the space provided.
 - Yes. My 1, 2 or 3 favorite NASCAR-related social media are:

 - No
- For questions 8-17, answer the sponsors as you know them without looking them up.*

| Favorite Driver(s) | Sponsor(s) |
|--------------------|------------|
| | |
| | |
| | |

9. Which of the following companies are primary sponsors for Kyle Busch in the NASCAR Sprint Cup Series? Check all that apply.
 - M&M's
 - Skittles
 - Dollar General
 - Pedigree Dog Food
 - Snickers Extreme
 - Interstate Batteries
 - Little Debbie
10. Which of the following companies are primary sponsors for Dale Earnhardt Jr. in the NASCAR Sprint Cup Series? Check all that apply.
 - State Farm
 - Nationwide Insurance
 - Pepsi
 - Mountain Dew
 - TurboTax
 - TaxSlayer.com
 - Axalta Coatings Systems
11. Which of the following companies are primary sponsors for Austin Dillon in the NASCAR Sprint Cup Series? Check all that apply.
 - Lucky Charms
 - Cheerios
 - Bass Pro Shops
 - Dick's Sporting Goods
 - American Ethanol
 - Dow Chemical
 - Caterpillar
12. Which company sponsors Brad Keselowski in the NASCAR XFINITY Series?
 - NOS
 - GameStop
 - Fleetwood/REV Group
13. Which company sponsors Ty Dillon in the NASCAR XFINITY Series?
 - Hellmann's
 - WESCO
 - Menard's
14. Which company sponsors Kevin Harvick in the NASCAR XFINITY Series?
 - South Point
 - Reynolds Wrap
 - Armour Vienna Sausage
15. Which company sponsors Johnny Sauter in the NASCAR Camping World Truck Series?
 - Champion
 - Allegiant Travel
 - Cooper Standard
16. Which company sponsors William Byron in the NASCAR Camping World Truck Series?

- Haas Automation
 - Alamo
 - Liberty University
17. Which company sponsors Matt Crafton in the NASCAR Camping World Truck Series?
- Ideal Door/Menards
 - Ride TV
 - SiriusXM
18. Do you purchase products or services specifically because the company or brand has a sponsorship in NASCAR?
- Yes
 - No
19. If your answer was yes to the last question, which products or services have you purchased? Please list:
20. If you do purchase products or services because the company or brand has a sponsorship in NASCAR, how much money do you likely spend per year?
- Not Applicable
 - Less than \$50
 - \$50-\$100
 - \$101-\$150
 - \$151-\$200
 - \$201-\$500
 - \$501-\$1,000
 - Over \$1,000

The Effects of Educational Attainment on Enrollment in Government Assistance Programs

Alex Matthews

Abstract

This report examines the effect of educational attainment on enrollment in government programs including public assistance, free and reduced lunch, the Supplemental Nutrition Assistance Program, and rent assistance. Using the Current Population Survey March 2015 Supplement, also known as the CPS Annual Social and Economic Supplement (ASEC), the research attempts to find trends between government assistance enrollment and three different educational attainment levels. The three levels are those who drop out of high school, those who graduate high school, and those who receive any form of post-high school education. Additionally, the research analyzes the effects of various demographic and socio-economic identifiers on government program enrollment. These include age, race, ethnicity, marital status (divorced, widowed, never married), income, and the number of people living in the household. This report aims to bring insight into the populace of current government program enrollees, while identifying potential implications for policy going forward. The findings present interesting opportunities for further study regarding those living on government assistance.

Dropping out of high school not only impacts the individual but society as a whole, both socially and economically. According to a 2007 report by Center for Labor Market Studies, high school dropouts earn \$9,200 less per year than those who graduate. The Center for Labor Market Studies (2007) also claims this discrepancy between high school dropouts and high school graduates equates to \$375,000 over a lifetime, and about \$1 million when comparing high school dropouts to college graduates. Also, a 2007 study by the Center for Benefit-Cost Studies of Education estimates that society could save over \$200,000 in prison and other costs for every potential dropout who could complete high school. Beyond earnings and prison expenses, there is a plethora of government welfare programs that are funded by taxpayers. Researching education's impact on government assistance participation can lead to a more complete understanding of the American population, the country's educational shortcomings and strengths, as well as how schooling interacts with government aid.

I will consider which factors correlate with enrollment in the following four government assistance programs individually: public assistance, free or reduced lunch, food stamps, and rent assistance. Additionally, I will examine which factors predict the likelihood that an individual will be enrolled in any of the aforementioned programs. I am primarily interested in how varying levels of educational attainment predict enrollment in government assistance programs. The aim is to garner a well-rounded understanding of the interaction between educational attainment and various socioeconomic factors in America.

In summary, the report will center around two principal questions:

- 1) Can educational attainment predict the enrollment in government assistance programs?
- 2) Is it possible to predict which individuals will be enrolled in specific government assistance programs based on educational attainment, age, race, ethnicity, marital status, income, and number of persons in the household?

Relevant Literature and Research

There is a significant amount of research regarding educational attainment, household factors, and poverty indicators circulating today in the fields of socioeconomics, labor and public economics, as well as education. A 2015 report by Carlson and McChesney examined sustainability of income in relation to educational attainment over time. The authors used peer-reviewed literature, scholarly research, and media articles from 1991 to 2010 to gather data for the study. The research identified a positive correlation between education levels and salaries. Additionally, Carlson and McChesney (2015) found that earnings declined over time for those who attained less than a bachelor's degree. Comparing those who received a bachelor's degree and those who did not, the analysis showed that there is a less volatile change in annual earnings for females than males when looking at varying ranges of educational attainment. Furthermore, they conclude that the buying power and standard of living in America is declining for those who attain less than a bachelor's degree. While the focus of my paper is not longitudinal, Carlson and McChesney's (2015) study over time is vital to understanding the research that is currently circulating.

Secondly, a 2009 piece by Dubow, Boxer, and Huesmann attempts to predict individuals' educational and work success at age 48 from variables assessed during their middle childhood and late adolescence. The authors found that parents' educational level when the child was eight years old was a strong indicator of educational and occupational success for that same child forty years later. The method behind this study is intriguing. The data come from the Columbia County Longitudinal Study, which began in 1960 with 856 third graders from New York. These

children were interviewed with their parents at age 8, then again at nineteen, thirty and forty-eight. Dubow, Boxer, and Huesmann (2009) propose that both parent education and family interaction patterns during childhood are linked directly to a child's academic and future success. This study, while also longitudinal, is still vital to understanding the scope of the impact that parent education has on socioeconomic factors in the future. Perhaps there is a similar trend in family lineage regarding enrollment in government programs.

In a more concise study, Rank and Hirschl (2005) analyze the probability and duration of food stamp use relating to gender, race and education. Using data from 1968 to 1997 from the Panel Study of Income Dynamics, the authors concluded that slightly over half of all Americans between age 20 and 65 receive food stamps at some point. In this paper, I cross-sectionally examine the relationship between education and programs such as food stamps. The numbers in my study will differ since I am looking at an exact point in time, rather than food stamp use over the course of several decades. Nonetheless, the implications of the research of both Rank and Hirschl's (2005) work and this paper are similar, giving insight into the food stamp situation in America.

A 2005 report by Waldfogel, Garfinkel, and Kelly attempts to find a correlation between improved education and public assistance enrollment reduction. The authors utilize the March 2003 Current Population Survey to construct a sample of single mothers. They then estimated logit models for Temporary Assistance for Needy Families (TANF) participation and used the models to determine the effect of being a high school dropout or high school graduate. These effects are then used to determine how many fewer mothers would participate in TANF with a high school diploma, which allows for an estimate of total government savings. The study found that there is potential for substantial savings in public assistance costs through improved education. In fact, the total amount that the authors claimed ranges from \$7.9 to \$10.8 billion. The authors admit that these calculations are cruder than they would prefer, as some of the assumptions are likely to lead to underestimates, while others are likely to lead to overestimates. Nonetheless, the paper is congruent with the scope of my study. Educational attainment can be directly related to government assistance enrollment, and ultimately, government expenditure or saving.

Lastly, Hout (2012) analyzes the effect of education on socioeconomic factors including divorce rate, health, social contributions and income. Drawing on 110 previous works, Hout (2012) summarizes the findings, identifies congruencies and discrepancies, and attempts to draw a coherent conclusion given the research of others. Hout's (2012) most prominent conclusion is that education "makes life better" (p. 394). He acknowledges that those with a higher level of education make more money, live healthier lives, have fewer divorces, and contribute more to communities than those with lower educational attainment. The author also delves into the dynamics of education. He argues there are specific skills that are much harder to acquire outside of the classroom, such as how to make sense of Shakespeare, create a reasoned argument, or learn to see a task from beginning to end. Furthermore, Hout (2012) proffers that educational opportunity is unequal, but concedes it would be even more unequal without schools.

Overall, there is a plethora of research on the socioeconomic impact of educational attainment on families and households; however, there is not much that specifically tries to sum up a variety of poverty indicators in conjunction with educational attainment. The strength of measuring fewer variables is that it could produce more telling results; however, the weakness is that some of these studies fail to measure the interaction of multiple household factors with educational attainment. My study will supplement the conversation of education and

socioeconomic factors by finding reliable correlations that can aid in the understanding of the relationship between educational attainment and use of government assistance. Being able to evaluate the characteristics of individuals who will use government welfare programs such as food stamps, free or reduced lunches for children, and reduced rent is vital for policymakers. An improved understanding of the socioeconomic characteristics of American households will allow decision-makers to identify potential solutions to the issues the country faces.

Methods

Theory and Coefficient Predictions

The purpose of this research is to gain an understanding of the characteristics of Americans who are enrolled in government assistance programs. I examine these data through the lens of educational attainment, namely whether individuals are high school dropouts, high school graduates, and recipients of post-high school education. The variables included in the analysis are educational attainment (only high school dropouts and post-high school education recipients, as high school graduates are used as a base group for the analysis), age, race (black and other races, as white is the base group for analysis) and ethnicity (Hispanic). Additionally, marital status variables are included and split up into the following categories: divorced/separated, widowed, and never married. “Married” serves as our control group. The final two variables for measurement are the natural log of household income and the number of people in the household. In total, there are eleven variables included in the analysis to determine the likelihood of enrollment in government assistance programs.

According to a 2015 report from the U.S. Census Bureau, 21.3 percent of people in the U.S. participated in “major means-tested government assistance programs each month in 2012.” The findings of this study reported that 26.4 percent of the sample population participated in four government assistance programs, including public assistance, free or reduced lunch, food stamps, and rent assistance. Using theory, intuition, and third-party statistical findings, I will hypothesize the signs of the coefficients for the likelihood of enrollment in any or multiple government assistance programs.

High School Dropouts: I hypothesize the coefficient for high school dropouts will be positive. According to the U.S. Census Bureau (2015), 37.3 percent of high school dropouts were enrolled in government assistance in 2012. Additionally, lower education levels generally lead to lower income. Carlson and McChesney (2014) argue that the buying power and standard of living in America is declining for those who obtain less than a bachelor’s degree. Thus, I predict that:

(H1) Being a high school dropout increases the likelihood of enrollment in government programs relative to high school graduates.

Post-High School Education: I hypothesize those who received post-high school education will be less likely to receive government assistance. Hout (2012) argues that education “makes life better” because individuals are earning more money, living healthier lives, having fewer divorces, and contributing more to communities. Intuitively, those with higher educational attainment generally earn more, thus reducing their need to be enrolled in government programs. Thus, I predict that:

(H2) Receiving post-high school education decreases the likelihood of enrollment in government assistance programs relative to high school graduates.

Age: According to the U.S. Census Bureau (2015), as age increased, the likelihood of enrollment in government programs decreased. For instance, 16.6 percent of people between 18 and 64 received assistance, compared to 12.6 percent of people 65 and older. Furthermore, as

people grow older, they generally earn more money and have more stability in their lives. Thus, I predict that:

(H3) An increase in age will decrease the likelihood of enrollment in government assistance programs.

Black, Other Races: According to the U.S. Census Bureau (2015), the groups of people who were most likely to receive government assistance were African-Americans (41.6 percent), followed by Hispanics (36.4 percent), then Asians or Pacific Islanders (17.8 percent). The report stated that 13.2 percent of non-Hispanic whites enrolled in government programs. This hierarchy of past data leads me to predict that:

(H4) Being black increases the likelihood of enrollment in government assistance programs relative to being white.

(H5) Being any race other than white or black increases the likelihood of enrollment in government assistance programs relative to being white.

Hispanic: As previously stated, the U.S. Census Bureau (2015) found that Hispanics were the second most likely people group to receive government assistance, behind African-Americans. Therefore, I predict that:

(H6) Being Hispanic increases the likelihood of enrollment in government assistance programs relative to being non-Hispanic.

Marital Status: The U.S. Census Bureau (2015) reports that people in female-householder families had the highest rates of participation in government programs, reporting 50 percent. Married-couple families and male-householder families were 14.7 percent and 29.5 percent respectively. Intuitively, this means that divorce, separation, death of a spouse, or singleness may increase the need for government assistance since a single person is responsible for supporting themselves and the others in the household. Having a single source of income could mean heavier reliance on government aid. Thus, I predict that:

(H7) Being divorced or separated will increase the chance an individual is enrolled in a government assistance program relative to those who are married.

(H8) Being widowed will increase the chance an individual is enrolled in a government assistance program relative to those who are married.

(H9) Never being married will increase the chance an individual is enrolled in a government assistance program relative to those who are married.

Income: Intuition suggests the more money that someone earns, the less likely they are to need any type of government assistance or aid. Theory would imply that increased income would reduce the chance of enrollment in any of the five identified government aid programs (public assistance, free or reduced lunch, food stamps, and rent assistance). Thus, I predict an inverse relationship between income and aid enrollment, specifically that:

(H10) An increase in income will decrease the likelihood of enrollment in any aid program.

Number of Persons in Household: Intuitively, more people living in a household places more economic strain on owners of the household to provide for those under the roof. Desai's (1995) report proposes that the effect of family size on children's well being depends on the extent to which parents bear the cost of rearing children. Therefore, as family size increases, burden increases, making families with more children increasingly likely to be enrolled in a government assistance program, such as free or reduced lunches. Thus I predict that:

(H11) As the number of people in the household increases, the likelihood of enrollment in a government aid program increases.

These predictions will remain the same with each government program that is analyzed: public assistance, free or reduced lunch, food stamps, and rent assistance.

Regression Model

A single regression will be required to analyze government programs all together, as well as separate regressions for each individual government assistance program, as the variables that are included will remain the same across each regression. The hypotheses for each variable will be examined for its impact on the likelihood of receiving government assistance. *Y* is enrollment in government programs as well as one of five binary outcomes including enrollment in public assistance, free or reduced lunch, food stamps or rent assistance. The overarching regression model is as follows:

$$Y = \beta_0 + \beta_1\text{HSDROPOUT} + \beta_2\text{POSTHSEDU} + \beta_3\text{AGE} + \beta_4\text{BLACK} + \beta_5\text{OTHERRACE} + \beta_6\text{HISPANIC} + \beta_7\text{DIVSEP} + \beta_8\text{WIDOW} + \beta_9\text{NEVERMARRIED} + \beta_{10}\text{LN(INCOME)} + \beta_{11}\text{PERSONS} + \epsilon,$$

where *Y* = GOVTPROGRAM (any of the four programs);
 PUBLIC ASSISTANCE;
 FREE / REDUCED LUNCH;
 FOOD STAMPS;
 RENT ASSISTANCE

Therefore, the dependent variable can be substituted for any of the individual programs or can be analyzed for the population as an overall estimate of those who receive any form of government aid or assistance. Coefficient predictions are in the table below (Table 1):

Table 1 – Coefficient Predictions

| Variable | HS Drop out | Post- HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income | Persons in HH |
|----------------|-------------|--------------|-----|-------|------------|----------|--------|-------|---------------|--------|---------------|
| Coef. Predict. | + | - | - | + | + | + | + | + | + | - | + |

The above coefficient predictions will remain the same for the following five regressions. The only part of the regression that will change is the dependent variable, as it will be substituted to analyze each individual government assistance program based on the 11 independent variables. These will include public assistance, free or reduced lunch, food stamps, and rent assistance. The base cases for these regression models include high school graduates, those who identify as white, and married individuals.

Data Description

This paper will examine the effects of educational attainment on enrollment in government assistance programs with a sample of Americans ages twenty-five and older. The data come from the Current Population Survey March 2015 Supplement, which is also known as the CPS Annual Social and Economic Supplement (ASEC). The U.S. Census Bureau creates and distributes the ASEC annually for the Bureau of Labor Statistics. The ASEC provides annual estimates based on a survey of more than 75,000 households. It covers the social and economic

characteristics of each person in the interviewed household. The Bureau of Labor Statistics also uses the ASEC to derive poverty status from income and household composition variables.

To understand the research at hand and the impact of educational attainment on socioeconomic factors, it is vital to anchor the results to the sample used. The data were limited to people twenty-five years and older to observe the effects of educational attainment levels. Intuitively, anyone under twenty-five was excluded for the reasoning that twenty-five years is ample time for someone to receive a high school degree or GED, and this age still covers a sample that graduated high school and went on to college. The descriptive statistics look at four groups: (1) Full Sample, twenty-five years and older; (2) High School Dropouts, twenty-five years and older; (3) High School Graduates, twenty-five years and older; (4) Any Post-High School Education, twenty-five years and older. Table 2 individually breaks down the descriptive statistics of the educational groups. High school dropouts are coded as anyone who finished their educational career between 9th and 12th grade, with no high school degree. High school graduates are those who obtained a secondary education diploma with no further education. Those with any post-high school education include anyone who attended college but did not obtain a degree, as well as those who obtained associate degrees (vocational and academic), bachelor's degrees, master's degrees, professional school degrees, and doctorate degrees.

First, the demographics of the sample are examined, including age, race, ethnicity, and marital status. Individuals were coded "black" if they selected any racial category with "black" in the description. Individuals were coded as "white" if they selected any racial category with "white" in the description, barring any "black" descriptors. Individuals were coded "other" if they did not fit the previous two constraints. The full sample is 78.9 percent white, 12.9 percent black and 8.2 percent other races. Individuals were coded "married" if they selected either "married – civilian spouse present," "married – Armed Forces spouse present," or "married – spouse absent." Individuals were coded "divorced/separated" if they selected "divorced" or "separated." Individuals were also coded as "widowed" and "never married" depending on their personal identification. Furthermore, individuals who identified as "Hispanic" were included under the ethnicity statistic.

With the second part of the analysis, the focus shifts to household information regarding socioeconomic and familial factors. The descriptive statistics include household income (in thousands), persons in household, and number of own children under eighteen years old. Also, the analysis examines the prevalence of receiving public assistance as well as public assistance income in dollars. The analysis also incorporates the prevalence of food stamp recipients and the prevalence of households who pay child support. The descriptive statistics provide the percentage of children who receive free or reduced lunch, as well as the average number of children per household who receive free or reduced lunch. Anyone without children was dropped so that I could narrow the scope to households with children and analyze the true number of children who receive free or reduced lunch.

The descriptive statistics for this cross-sectional study (as seen in Tables 3 and 4) are derived from a Stata analysis implementing the coding decisions. Tables 3 and 4 include basic statistics such as mean, standard error, standard deviation, and median. Each analysis was run for the four different sub-groups created among the sample population.

The discrepancies in income are worth noting. Among the full sample, average income is \$90,340, which is far above the U.S. Census Bureau's September 2014 average American income of \$51,939. Since we are limiting the sample to those above 25, we eliminate many individuals who are students, thus increasing the average income for this report. The income

value for high school dropouts drops to just over 50 percent of the full sample statistic at \$46,270. This is also \$5,000 below the Census Bureau’s average American income for 2014. Intuitively, this demonstrates the importance of educational attainment on income.

Regarding educational attainment and food stamps, the full sample (those who are 25 and older) shows a mean value of .112. In other words, 11.2 percent of the sample receives food stamps. Looking at high school dropouts, this number jumps 150 percent, as 28.8 percent of high school dropouts receive food stamps. The U.S. Census Bureau has offered that 21.3 percent of the American population receives food stamps; therefore, one is more likely to be a part of that group if he or she did not graduate from high school. The statistics show that obtaining some college education cuts participation in the food stamp program down to 6.4 percent, well below the occurrence of the average population.

For children receiving free or reduced-price lunches, it is important to note that families with incomes below 130 percent of the poverty level are eligible for free lunches, while families with incomes between 130 percent and 185 percent of the poverty level are eligible for reduced-price meals. According to the data, of the households with children, 28 percent of those households’ children receive free or reduced price lunches. That number jumps to 62.7 percent for the high school dropout sample, but decreases to 38.8 percent for high school graduates and 18.6 percent for those who receive any type of college education.

Weaknesses of the data arise from this limiting of the sample. While it may be necessary for the measurement of educational attainment on government assistance enrollment, it may eliminate an important portion of the population who are enrolled in government programs, namely young adults. As younger demographics generally earn lower incomes, the likelihood of government assistance enrollment increases, and this report may miss out on some of its effects.

Overall, the descriptive statistics are intuitive and logical. None of the data raise any red flags indicating errors in the coding of variables. Potential problems with the data set are the fact that incomes seem to be far above American averages, yet poverty-indicators such as free or reduced lunch still also seem relatively high.

TABLE 2 – Descriptive Statistics (Education Breakdown)

| Group | Mean | SE | Std. Dev. | Median | N |
|------------------------------|-------------|-----------|------------------|---------------|----------|
| High School Dropout | 0.0747 | 0.001 | 0.263 | 0 | 120882 |
| High School Grad | 0.306 | 0.001 | 0.461 | 0 | 120882 |
| Post-High School Edu. | 0.619 | 0.001 | 0.486 | 1 | 120882 |

TABLE 3 – Descriptive Statistics (Full Sample and High School Dropouts)

| <u>Variable</u> | <u>Full Sample (25+ years)</u> | | | | | <u>High School Dropouts (25+ years)</u> | | | | |
|--|--------------------------------|--------|-----------|--------|--------|---|-------|-----------|--------|-----|
| | Mean | SE | Std. Dev. | Median | N | Mean | SE | Std. Dev. | Median | |
| AGE | 49.437 | 0.045 | 15.49 | 48 | 120882 | 51.153 | 0.177 | 16.81 | 50 | 90% |
| RACE | | | | | | | | | | |
| <i>White</i> | 0.789 | 0.001 | 0.408 | 1 | 120882 | 0.742 | 0.005 | 0.437 | 1 | 90% |
| <i>Black</i> | 0.129 | 0.001 | 0.336 | 0 | 120882 | 0.182 | 0.004 | 0.386 | 0 | 90% |
| <i>Other</i> | 0.082 | 0.001 | 0.275 | 0 | 120882 | 0.076 | 0.003 | 0.265 | 0 | 90% |
| ETHNICITY | | | | | | | | | | |
| <i>Hispanic</i> | 0.141 | 0.001 | 0.348 | 0 | 120882 | 0.332 | 0.005 | 0.471 | 0 | 90% |
| MARITAL STATUS | | | | | | | | | | |
| <i>Married</i> | 0.619 | 0.001 | 0.485 | 1 | 120882 | 0.509 | 0.005 | 0.499 | 1 | 90% |
| <i>Divorced/Separated</i> | 0.14 | 0.001 | 0.347 | 0 | 120882 | 0.165 | 0.004 | 0.371 | 0 | 90% |
| <i>Widowed</i> | 0.059 | 0.001 | 0.236 | 0 | 120882 | 0.104 | 0.003 | 0.306 | 0 | 90% |
| <i>Never Married</i> | 0.181 | 0.001 | 0.385 | 0 | 120882 | 0.222 | 0.004 | 0.415 | 0 | 90% |
| HOUSEHOLD | | | | | | | | | | |
| <i>Income (in thousands)</i> | 90.34 | 0.269 | 93.38 | 69.4 | 120882 | 46.27 | 0.553 | 52.578 | 33.619 | 90% |
| <i>Natural Log of Income</i> | 4.082 | 0.003 | 1.093 | 4.24 | 120882 | 3.385 | 0.012 | 1.128 | 3.52 | 90% |
| <i>Persons in Household</i> | 3.028 | 0.005 | 1.568 | 3 | 120882 | 3.256 | 0.02 | 1.89 | 3 | 90% |
| <i>Number of own children under 18</i> | 0.694 | 0.003 | 1.076 | 0 | 120882 | 0.699 | 0.012 | 1.17 | 0 | 90% |
| GOVERNMENT AID | | | | | | | | | | |
| <i>Receives Public Assistance</i> | 0.014 | 0.0003 | 0.117 | 0 | 120882 | 0.038 | 0.002 | 0.1191 | NA | 90% |
| <i>Public Assistance Income (in dollars) [if received]</i> | 3674 | 86 | 3508 | 2664 | 1664 | 3578 | 156 | 2879 | 2864 | 36% |
| <i>Anyone Receives Food Stamps</i> | 0.106 | 0.001 | 0.307 | 0 | 120882 | 0.288 | 0.005 | 0.453 | 0 | 90% |
| <i>Any Children Receive Free/Reduced Lunch (have children 5 to 18)</i> | 0.28 | 0.002 | 0.449 | 0 | 45409 | 0.627 | 0.008 | 0.484 | 1 | 36% |
| <i>Any Children Receive Free/Reduced Lunch (total population)</i> | 0.105 | 0.001 | 0.307 | 0 | 120880 | 0.255 | 0.005 | 0.436 | 0 | 90% |
| <i>Number of children with free/reduced lunch (ages 5 to 18)</i> | 0.5 | 0.004 | 0.952 | 0 | 45409 | 1.183 | 0.02 | 1.231 | 1 | 36% |
| <i>Pays Child Support (children under 18 in household)</i> | 0.082 | 0.001 | 0.275 | 0 | 51889 | 0.082 | 0.004 | 0.275 | 0 | 39% |
| <i>Receives Government Assistance on Rent</i> | 0.011 | 0.0003 | 0.103 | 0 | 120882 | 0.024 | 0.002 | 0.153 | 0 | 90% |
| RECEIVES ANY GOVERNMENT AID | 0.204 | 0.0012 | 0.403 | 0 | 120882 | 0.449 | 0.005 | 0.497 | 0 | 90% |

TABLE 4 – Descriptive Statistics (High School Graduates and Post-High School Education)

| Variable | High School Graduates (25+ years) | | | | | Any Post-High School Education (25+ Years) | | | | |
|--|--|-----------|------------------|---------------|----------|---|-----------|------------------|---------------|----------|
| | Mean | SE | Std. Dev. | Median | N | Mean | SE | Std. Dev. | Median | N |
| AGE | 51.38 | 0.084 | 16.079 | 51 | 37015 | 48.271 | 0.055 | 14.911 | 46 | 74841 |
| RACE | | | | | | | | | | |
| <i>White</i> | 0.784 | 0.002 | 0.412 | 1 | 37015 | 0.797 | 0.001 | 0.403 | 1 | 74841 |
| <i>Black</i> | 0.149 | 0.002 | 0.356 | 0 | 37015 | 0.114 | 0.001 | 0.317 | 0 | 74841 |
| <i>Other</i> | 0.067 | 0.001 | 0.251 | 0 | 37015 | 0.089 | 0.001 | 0.286 | 0 | 74841 |
| ETHNICITY | | | | | | | | | | |
| <i>Hispanic</i> | 0.169 | 0.002 | 0.375 | 0 | 37015 | 0.104 | 0.001 | 0.305 | 0 | 74841 |
| MARITAL STATUS | | | | | | | | | | |
| <i>Married</i> | 0.58 | 0.003 | 0.494 | 1 | 37015 | 0.653 | 0.001 | 0.476 | 1 | 74841 |
| <i>Divorced/Separated</i> | 0.154 | 0.002 | 0.361 | 0 | 37015 | 0.13 | 0.001 | 0.337 | 0 | 74841 |
| <i>Widowed</i> | 0.081 | 0.001 | 0.273 | 0 | 37015 | 0.043 | 0.001 | 0.202 | 0 | 74841 |
| <i>Never Married</i> | 0.185 | 0.002 | 0.388 | 0 | 37015 | 0.174 | 0.001 | 0.379 | 0 | 74841 |
| HOUSEHOLD | | | | | | | | | | |
| <i>Income (in thousands)</i> | 66.89 | 0.339 | 65.117 | 52.418 | 37015 | 107.26 | 0.381 | 104.216 | 85.006 | 74841 |
| <i>Natural Log of Income</i> | 3.815 | 0.006 | 1.05 | 3.959 | 37015 | 4.298 | 0.004 | 1.044 | 4.442 | 74841 |
| <i>Persons in Household</i> | 3.034 | 0.009 | 1.645 | 3 | 37015 | 2.99 | 0.005 | 1.483 | 3 | 74841 |
| <i>Number of own children under 18</i> | 0.587 | 0.005 | 1.032 | 0 | 37015 | 0.747 | 0.004 | 1.08 | 0 | 74841 |
| GOVERNMENT AID | | | | | | | | | | |
| <i>Receives Public Assistance</i> | 0.018 | 0.001 | 0.134 | NA | 37015 | 0.009 | 0.0003 | 0.092 | NA | 74841 |
| <i>Public Assistance Income (in dollars) [if received]</i> | 3942 | 140 | 3638 | 3000 | 680 | 3441 | 144 | 3653 | 2440 | 643 |
| <i>Anyone Receives Food Stamps</i> | 0.146 | 0.002 | 0.353 | 0 | 37015 | 0.064 | 0.001 | 0.245 | 0 | 74841 |
| <i>Any Children Receive Free/Reduced Lunch (have children 5 to 18)</i> | 0.388 | 0.004 | 0.487 | 0 | 13227 | 0.186 | 0.002 | 0.389 | 0 | 28510 |
| <i>Any Children Receive Free/Reduced Lunch (total population)</i> | 0.139 | 0.002 | 0.346 | 0 | 37015 | 0.071 | 0.001 | 0.257 | 0 | 74841 |
| <i>Number of children with free/reduced lunch (ages 5 to 18)</i> | 0.694 | 0.009 | 1.061 | 0 | 13227 | 0.323 | 0.005 | 0.786 | 0 | 28510 |
| <i>Pays Child Support (children under 18 in household)</i> | 0.097 | 0.002 | 0.296 | 0 | 14591 | 0.076 | 0.001 | 0.265 | 0 | 33306 |
| <i>Receives Government Assistance on Rent</i> | 0.014 | 0.001 | 0.119 | 0 | 37015 | 0.007 | 0.0003 | 0.086 | 0 | 74841 |
| RECEIVES ANY GOVERNMENT AID | 0.264 | 0.002 | 0.441 | 0 | 37015 | 0.144 | 0.001 | 0.351 | 0 | 74841 |

Results and Analysis

The goal of this research is to examine the following questions:

- 1) Can educational attainment predict the enrollment in government assistance programs?
- 2) Is it possible to predict which individuals will be enrolled in specific government assistance programs based on educational attainment, age, race, ethnicity, marital status, income, and number of persons in the household?

The regression coefficients will be evaluated using the following criteria:

- 1) Is the absolute value of the t-statistic greater than 1.645? (This value is generated from a t-score chart evaluated for a one-sided test with 120,880 degrees of freedom at a .05 confidence level).
- 2) Does the t-statistic have the sign (+/-) implied by the alternative hypothesis?

Refer to the Table 5 on the next page for coefficient values, standard error, and the t-statistics that correspond to the five different government program enrollment regressions. Table 6 indicates the accuracy of the predicted values as compared to the actual values for each regression equation. Every regression predicts correct outcomes over 80 percent of the time, ranging from 82.2 percent to 98.9 percent depending on the dependent variable that is being measured.

In Appendix II, there are all the correlations for each variable in each regression. I used the value of 0.8 to as a threshold measure for multicollinearity between the variables. Of the six regressions, there was no evidence of simple multicollinearity in the correlation output. Every value is well below 0.8. Furthermore, in Appendix III there is the Variance Inflation Factor output to further examine the possibility of multicollinearity among the variables. The cut-off value for these data was 5. Each output falls far below 5, with the highest around 1.5, once again indicating no evidence of multicollinearity between variables.

TABLE 5 – Regression Results

| Variable | All Government Programs | | Public Assistance | | Free or Reduced Lunch | | Food Stamps | | Rent Assistance | |
|----------------------|-------------------------|--------------|----------------------|--------------|-----------------------|--------------|----------------------|--------------|----------------------|--------------|
| | Coef. (Std. Err.) | Reject H_0 | Coef. (Std. Err.) | Reject H_0 | Coef. (Std. Err.) | Reject H_0 | Coef. (Std. Err.) | Reject H_0 | Coef. (Std. Err.) | Reject H_0 |
| | [t-stat] | | [t-stat] | | [t-stat] | | [t-stat] | | [t-stat] | |
| HS Dropout | 0.108 (0.004) | | 0.012 (0.001) | | 0.058 (0.003) | | 0.089 (0.003) | | 0.004 (0.001) | |
| | [27.76] | Yes | [8.98] | Yes | [17.82] | Yes | [26.86] | Yes | [3.12] | Yes |
| Post-HS Edu | -0.051 (0.002) | | -0.004 (0.0007) | | -0.031 (0.001) | | -0.033 (0.002) | | -0.001 (0.001) | |
| | [-23.75] | Yes | [-5.36] | Yes | [-17.23] | Yes | [-17.97] | Yes | [-1.94] | Yes |
| Age | -0.006 (0.0003) | | -0.0001 (0.0001) | | -0.007 (0.0003) | | -0.003 (0.0003) | | -0.0003 (0.0001) | |
| | [-19.27] | Yes | [-0.088] | No | [-24.57] | Yes | [-9.51] | Yes | [-3.16] | Yes |
| Black | 0.09 (0.003) | | 0.004 (0.001) | | 0.045 (0.002) | | 0.059 (0.002) | | 0.014 (0.001) | |
| | [30.82] | Yes | [4.12] | Yes | [18.34] | Yes | [23.71] | Yes | [15.59] | Yes |
| Other Race | 0.006 (0.003) | | 0.001 (0.001) | | -0.003 (0.003) | | -0.006 (0.003) | | 0.005 (0.001) | |
| | [1.59] | No | [0.71] | No | [-0.89] | No | [-2.04] | No | [4.83] | Yes |
| Hispanic | 0.064 (0.003) | | -0.001 (0.001) | | 0.072 (0.002) | | -0.002 (0.002) | | -0.0005 (0.0009) | |
| | [22.47] | Yes | [-0.69] | No | [30.61] | Yes | [-0.67] | No | [-0.53] | No |
| DivSep | 0.134 (0.003) | | 0.015 (0.001) | | 0.077 (0.002) | | 0.101 (0.002) | | 0.008 (0.001) | |
| | [45.95] | Yes | [14.58] | Yes | [31.99] | Yes | [41.10] | Yes | [9.30] | Yes |
| Widow | 0.061 (0.004) | | 0.006 (0.001) | | 0.035 (0.004) | | 0.037 (0.004) | | 0.003 (.001) | |
| | [14.03] | Yes | [3.97] | Yes | [9.61] | Yes | [10.08] | Yes | [2.20] | Yes |
| Never Married | 0.084 (0.003) | | 0.021 (0.001) | | 0.022 (0.002) | | 0.089 (0.002) | | 0.012 (0.001) | |
| | [30.06] | Yes | [21.64] | Yes | [9.67] | Yes | [37.61] | Yes | [14.08] | Yes |
| Income | -0.11 (0.001) | | -0.009 (0.0003) | | -0.060 (0.001) | | -0.088 (0.001) | | -0.011 (0.0003) | |
| | [-114.34] | Yes | [-28.03] | Yes | [-75.28] | Yes | [-107.31] | Yes | [-35.31] | Yes |
| Persons in HH | 0.078 (0.001) | | 0.009 (0.0002) | | 0.074 (0.001) | | 0.041 (0.001) | | 0.0001 (0.0002) | |
| | [112.13] | Yes | [34.97] | Yes | [128.08] | Yes | [68.43] | Yes | [0.65] | No |
| Constant | 0.422 (0.007) | | 0.022 (0.002) | | .0183 (0.001) | | 0.345 (0.006) | | 0.05 (0.002) | |
| | [64.20] | Yes | [9.74] | Yes | [33.50] | Yes | [61.93] | Yes | [25.41] | Yes |
| N | | 120882 | | 120882 | | 120882 | | 120882 | | 120882 |
| R² | | 0.2521 | | 0.0257 | | 0.214 | | 0.1847 | | 0.023 |

TABLE 6 – Predicted vs. Actual Outcome Accuracy (where 1 = successful prediction)

| Dependent Variable | Mean | N |
|---------------------------|-------------|----------|
| Any Govt. Program | 0.822 | 120882 |
| Public Assistance | 0.986 | 120882 |
| Free/Reduced Lunch | 0.895 | 120882 |
| Food Stamps | 0.894 | 120882 |
| Rent Assistance | 0.989 | 120882 |

Results -- Regression 1 -- All Government Aid

The mean government program enrollment for the total sample is 0.204. In this regression, I analyze how the 11 independent variables influence the likelihood that an individual is enrolled in any of the four government assistance programs included in this study (public assistance, free or reduced lunch, food stamps, and rent assistance.) For this regression, all null hypotheses are rejected, besides other race. Table 5 shows that for all variables, except other race, have an absolute value of the t-statistic that is greater than 1.645 and a t-stat sign that is implied by the alternative hypothesis. However, with other race, while the sign is congruent with that of the alternative hypothesis, the t-stat is insignificant ($1.59 < 1.645$). Thus, there is a failure to reject the null hypothesis that $\beta \leq 0$ for other race.

The coefficient for high school dropout is 0.108, over half of the sample mean for anyone over the age of 25. Therefore, a high school dropout has a 53 percent higher chance of enrolling in a government assistance program (holding all else constant). Post-high school education accounted for a 5.1 percentage point decrease in the likelihood of government aid enrollment. As one gets older, the likelihood of government aid enrollment decreases; however, the regression shows that this number does not carry a large impact with a coefficient value of -0.006. Being black accounts for 44 percent increase in the likelihood of government aid enrollment. Being Hispanic has the economic implication of a 31 percent increase in the likelihood of government aid enrollment. Furthermore, being divorced or separated seems to have a significant outcome on whether an individual is enrolled in government aid. With a coefficient value of 0.134, this indicates being divorced or separated accounts for a 13.4 percentage point increase in receiving government assistance. In other words, this accounts for a 66 percent increase in the likelihood of being enrolled in a government program. The widow coefficient is 0.061, which translates to a 6.1 percentage point increase of receiving government assistance and accounts for a 25 percent increase in likelihood of enrollment. Never getting married increases the likelihood of government aid enrollment by 32 percent. Furthermore, as income increases by 1 percent, the likelihood of enrolling in a government decreases by 11 percentage points. Comparing this finding with the population mean of enrollment in government programs (20.4 percent), increasing income by 1 percent indicates a 50 percent decrease in the chance of enrolling in a government aid program. Lastly, increasing the number of people in the household by one leads to a 7.8 percentage point increase in the likelihood of enrolling in any government aid program.

Results -- Regression 2 -- Public Assistance

This regression uses the same variables as the previous regression, but looks solely at enrollment in the U.S. government public assistance income program, as opposed to all four government programs. The mean public assistance program enrollment for the total sample is 0.014. In this regression, I fail to reject three null hypotheses, namely age, other race, and Hispanic. Age has the correct hypothesized sign, however its t-statistic is too small

($0.088 < 1.645$), leading me to fail to reject the null. The same is also true with other race and Hispanic. All other variables pass both tests of having the sign of the alternative hypotheses and a t-statistic over 1.645.

Due to the fact that a much smaller portion of the population partakes in public assistance than in any government program all together, the coefficient values are smaller than our previous regression. The coefficient for high school dropout is 0.012, making for an 86 percent increase in the likelihood of receiving public assistance. Having any education after high school decreases the likelihood of receiving public assistance by nearly 30 percent, holding all else constant when comparing to the population mean. The coefficient value is -0.004. According to Table 5, the coefficient for the black variable is 0.004, which also indicates a 30 percent change in the likelihood of receiving public assistance; however, this coefficient is positive, therefore it relates to an increase in public assistance enrollment chances. Once again, being divorced or separated has a very significant impact on the outcome. In this regression, being divorced or separated doubles the likelihood of receiving public assistance from the government, as the coefficient is 0.015. The effect is even larger if one has never been married. A person who has never been married has 150 percent higher chance of receiving public assistance than married individuals, holding all other variables constant. Furthermore, a 1 percent increase in income indicates a 0.9 percentage point decrease in the likelihood of receiving public assistance income when compared to the left-out group of high school graduates, which does not seem very significant.

Results -- Regression 3 -- Free or Reduced Lunch

10.5 percent of the total sample utilizes free and reduced lunch. For this regression regarding children on free or reduced lunch, there is a failure to reject one of the null hypotheses. The hypothesis belongs to the other race coefficient and it is rejected due to an incongruent sign and an insufficient t-statistic ($0.89 < 1.645$). The other ten variables are statistically significant and have coefficient signs that align with the alternative hypotheses.

When compared to high school graduates, the results show that high school dropouts have a 5.8 percentage point higher chance of having children on free or reduced lunch. Comparing with the population mean (.105), this equates to a 55 percent increase in the chance of enrolling children in the free or reduced lunch program. Receiving any education after high school, comparing to those who complete high school only, leads to a 30 percent decrease in the likelihood of having children on free or reduced lunch. Being divorced or separated increases the chances by 73 percent. This is an interesting statistic. Broken homes can lead to financial pressure, especially on families with children, so perhaps it is apparent that more divorced or separated households would utilize free and reduced lunches to ease some of the financial constraints. A 1 percent increase in income leads to a 6 percentage point decrease in the likelihood of utilizing free or reduced lunches when compared to the high school graduate sample (the left out group).

Results -- Regression 4 -- Food Stamps

10.6 percent of the total population utilizes food stamps. For this regression, analyzing food stamp recipients, I fail to reject two of the eleven null hypotheses regarding the variables' coefficient values. The two null hypotheses that are not rejected belong to the other race variable and Hispanic variable. Other race has a significant t-statistic, but a sign that contradicts the alternative hypothesis. The alternative hypothesis predicts that being any race other than white or black would increase the chance that an individual partakes in food stamps; however, the regression shows a negative sign. For Hispanic, both an insignificant t-statistic and a contradictory coefficient sign lead to a failure to reject the null hypothesis. The alternative

hypothesis, under the same intuition for other race, assumes that being of Hispanic descent leads to an increase in the likelihood of receiving food stamps. The regression analysis shows a negative sign, thus a failure to reject the null.

According to the regression results, being black increases the likelihood of receiving food stamps by over 50 percent. The coefficient value for black is 0.059, and the mean for receiving food stamps over the whole population is 0.106. Being divorced or separated again has a significant impact on receiving this specific government aid. With a coefficient of 0.101, being divorced or separated nearly doubles the likelihood of receiving food stamps. A 1 percent increase in income leads to an 8.8 percent reduction in receiving food stamps when compared to the left-out group of high school graduates.

Regression 5 -- Rent Assistance

The mean rent assistance program enrollment for the total sample is 0.011. For this regression regarding rent assistance, there is a failure to reject two of the null hypotheses, namely Hispanic and persons in household. The Hispanic coefficient has a sign that is opposite the alternative hypothesis and an insignificant t-statistic ($0.53 < 1.645$). The persons in household coefficient has a positive sign, which aligns with the alternative hypothesis, but does not meet the t-statistic requirement ($0.65 < 1.645$).

According to the regression, being black more than doubles the likelihood of utilizing rent assistance (coefficient = .014). In fact, it equates to a 127 percent chance increase. Additionally, holding all else constant, being divorced or separated increases the likelihood of receiving rent payment assistance by over 70 percent. The coefficient for high school dropouts is 0.004. While that number may seem insignificant, it carries noteworthy economic implications. Compared to the population mean, being a high school dropout increases the likelihood of receiving rent assistance by 36.

Logit Output – All Regressions

In running logistical regressions for all regression models, I found similar results to the Ordinary Least Squares models that were previously used. The output for the logit models can be found in Appendix III. The logit coefficients were the same sign for all variables that were statistically significant and met the criteria for a rejection of the null hypothesis. Hence, the OLS method and Logit method are qualitatively similar, giving confidence in the validity of the results.

Conclusions

Re-examine the two research questions at hand:

- 1) Can educational attainment predict the enrollment in government assistance programs?
- 2) Is it possible to predict which individuals will be enrolled in specific government assistance programs based on educational attainment, age, race, ethnicity, marital status, income, and number of persons in the household?

Statistically, the answer to the first question is “yes.” The regression analysis showed that the coefficients for high school dropout and post-high school education recipients were significant for each regression. Therefore, educational attainment provides insight into whether an individual will be a part of a government assistance program.

For the second question, the answer is more complicated. In looking through the results section, there are plenty of variables in each regression that were not statistically significant in certain situations. Therefore, we cannot fully answer the second research question. In examining the hypotheses I failed to reject, I notice that for nearly every regression (except rent assistance),

the variable for those of other races was determined statistically insignificant. This is intriguing. Most of the population was either white or black, so perhaps this statistical insignificance was derived from having a very small portion of individuals who were races other than black or white or the notion that they do not look different from white individuals in terms of the variables that I measured. The smaller other race population could have made it difficult for the regression to produce a statistically significant figure.

For the most part, the findings were in line with other works as well as intuition. There were no variables that produced any significant “novel” insight; however, some of the results showed far more impact on the likelihood of receiving government programs than I would have predicted. For example, being divorced or separated had a noticeable influence on the likelihood of enrollment in each of the five government programs. In fact, in each regression when compared to the population mean, being divorced or separated increased the likelihood of enrolling in a program by at least 60 percent. For the most part, lower educational attainment as a whole leads to an increased likelihood of enrollment in government assistance programs. While this result is expected, the results allow us to compare which characteristics of an individual is more likely to influence government aid participation.

Income is probably the most intuitive variable to predict when it comes to government assistance, and in each regression that income was deemed statistically significant, it mattered. However, I predicted it would be more noticeable than the regression results showed. Limiting the population to people over 25 makes sense for measuring education, but could have hindered some of the effects for people on government aid that are under 25. In fact, that could be the biggest drawback to this work. Cutting out those that are under 25 (possible full-time students) significantly reduces the sample population and could influence the way some of the other variable coefficients are generated. However, dropping those under 25 allowed me to take an in-depth look at how education is influencing government aid enrollment.

After completing this report, there are recommendations that I have for future research. Mainly, a longitudinal study that captures an understanding of how long people stay on any form of government assistance can be beneficial to recognizing the effects of education and the various other demographic and economic variables included in this study. Furthermore, and perhaps slightly less related to the scope of this report, I recommend another longitudinal study that determines how educational attainment of household owners and enrollment in government aid programs affects children. It would be interesting to know how many of these children whose families receive government assistance now will receive that same assistance in the future. Being able to track these data would be beneficial for policymakers, influential members of the U.S. government, and would allow the American taxpayers to understand more about where the funding is going.

Bibliography

- Carlson, R.H., & McChesney, C.S. (2015). Income Sustainability through Education Attainment. *Journal of Education and Training Studies*, 3(1), 108-115.
- Center for Labor Market Studies (2007). Left behind in America: The nation's dropout crisis. Retrieved from <http://chicagowritingservice.com/content/wpcontent/uploads/2008/09/clmsexecreport2.pdf>.
- Desai, S. (1995). When Are Children from Large Families Disadvantaged? Evidence from Cross-National Analyses. *Population Studies*, 49(2), 195-210. Retrieved from <http://www.jstor.org/stable/2175152>
- Dubow, E.F., Boxer, P., & Huessman, L.R. (2009). Long-term Effects of Parents' Education on Children's Educational and Occupational Success: Mediation by Family Interactions, Child Aggression and Teenage Aspirations. *Merrill-Palmer Quarterly (1982-)*. (3), 224.
- Hout, M. (2012). Social and Economic Returns to College Education in the United States. *Annual Review of Sociology*, 38(1), 379-400.
- Rank, M.R. & Hirschl, T.A. (2005). Research Article: Likelihood of Using Food Stamps during the Adulthood Years. *Journal of Nutrition Education and Behavior*, 37137-146.
- Waldfogel, J., Garfinkel, I., & Kelly, B. (2005, October 24). Public Assistance Programs: How Much Could Be Saved with Improved Education? *Columbia University School of Social Work*. doi:http://devweb.tc.columbia.edu/manager/symposium/Files/79_waldfogel_paper.ed.pdf
- 21.3% of US Participates in Government Assistance Programs Each Month. (2015, May 28). Retrieved October 15, 2016, from <http://www.census.gov/newsroom/press-releases/2015/cb15-97.html>

Appendix I – Correlations

Note: n for all Logistical Regressions is 120882

| Any Government Program | | | | | | | | | | | | |
|------------------------|-------------|------------|-------------|--------|--------|------------|----------|--------|--------|---------------|--------------|---------------|
| | Govt. Prog. | HS Dropout | Post-HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income (log) | Persons in HH |
| Govt. Prog. | 1 | | | | | | | | | | | |
| HS Dropout | 0.173 | 1 | | | | | | | | | | |
| Post-HS Edu | -0.189 | -0.362 | 1 | | | | | | | | | |
| Age | -0.149 | 0.030 | -0.094 | 1 | | | | | | | | |
| Black | 0.116 | 0.044 | -0.060 | -0.004 | 1 | | | | | | | |
| Other Race | -0.002 | -0.006 | 0.037 | -0.037 | -0.115 | 1 | | | | | | |
| Hispanic | 0.138 | 0.156 | -0.135 | -0.134 | -0.111 | -0.076 | 1 | | | | | |
| DivSep | 0.099 | 0.020 | -0.036 | 0.059 | 0.053 | -0.042 | 0.006 | 1 | | | | |
| Widow | -0.015 | 0.054 | -0.089 | 0.338 | 0.024 | -0.016 | -0.045 | -0.101 | 1 | | | |
| Never Married | 0.106 | 0.030 | -0.023 | -0.310 | 0.160 | 0.014 | 0.058 | -0.190 | -0.118 | 1 | | |
| Income (log) | -0.275 | -0.181 | 0.252 | -0.107 | -0.153 | 0.036 | -0.067 | -0.166 | -0.154 | -0.133 | 1 | |
| Persons in HH | 0.229 | 0.041 | -0.025 | -0.351 | -0.066 | 0.082 | 0.134 | -0.161 | -0.180 | -0.093 | 0.259 | 1 |

| Public Assistance | | | | | | | | | | | | |
|-------------------|---------------|------------|-------------|--------|--------|------------|----------|--------|--------|---------------|--------------|---------------|
| | Public Assist | HS Dropout | Post-HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income (log) | Persons in HH |
| Public Assist | 1 | | | | | | | | | | | |
| HS Dropout | 0.059 | 1 | | | | | | | | | | |
| Post-HS Edu | -0.057 | -0.362 | 1 | | | | | | | | | |
| Age | -0.046 | 0.030 | -0.094 | 1 | | | | | | | | |
| Black | 0.034 | 0.044 | -0.060 | -0.004 | 1 | | | | | | | |
| Other Race | 0.005 | -0.006 | 0.037 | -0.037 | -0.115 | 1 | | | | | | |
| Hispanic | 0.029 | 0.156 | -0.135 | -0.134 | -0.111 | -0.076 | 1 | | | | | |
| DivSep | 0.028 | 0.020 | -0.036 | 0.059 | 0.053 | -0.042 | 0.006 | 1 | | | | |
| Widow | -0.005 | 0.054 | -0.089 | 0.338 | 0.024 | -0.016 | -0.045 | -0.101 | 1 | | | |
| Never Married | 0.066 | 0.030 | -0.023 | -0.310 | 0.160 | 0.014 | 0.058 | -0.190 | -0.118 | 1 | | |
| Income (log) | -0.088 | -0.181 | 0.252 | -0.107 | -0.153 | 0.036 | -0.067 | -0.166 | -0.154 | -0.133 | 1 | |
| Persons in HH | 0.078 | 0.041 | -0.025 | -0.351 | -0.066 | 0.082 | 0.134 | -0.161 | -0.180 | -0.093 | 0.259 | 1 |

| Free / Reduced Lunch | | | | | | | | | | | | |
|----------------------|--------------|------------|-------------|--------|--------|------------|----------|--------|--------|---------------|--------------|---------------|
| | Free/Reduced | HS Dropout | Post-HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income (log) | Persons in HH |
| Free/Reduced | 1 | | | | | | | | | | | |
| HS Dropout | 0.138 | 1 | | | | | | | | | | |
| Post-HS Edu | -0.143 | -0.362 | 1 | | | | | | | | | |
| Age | -0.185 | 0.030 | -0.094 | 1 | | | | | | | | |
| Black | 0.063 | 0.044 | -0.060 | -0.004 | 1 | | | | | | | |
| Other Race | 0.006 | -0.006 | 0.037 | -0.037 | -0.115 | 1 | | | | | | |
| Hispanic | 0.167 | 0.156 | -0.135 | -0.134 | -0.111 | -0.076 | 1 | | | | | |
| DivSep | 0.055 | 0.020 | -0.036 | 0.059 | 0.053 | -0.042 | 0.006 | 1 | | | | |
| Widow | -0.042 | 0.054 | -0.089 | 0.338 | 0.024 | -0.016 | -0.045 | -0.101 | 1 | | | |
| Never Married | 0.040 | 0.030 | -0.023 | -0.310 | 0.160 | 0.014 | 0.058 | -0.190 | -0.118 | 1 | | |
| Income (log) | -0.165 | -0.181 | 0.252 | -0.107 | -0.153 | 0.036 | -0.067 | -0.166 | -0.154 | -0.133 | 1 | |
| Persons in HH | 0.340 | 0.041 | -0.025 | -0.351 | -0.066 | 0.082 | 0.134 | -0.161 | -0.180 | -0.093 | 0.259 | 1 |

Appendix I (continued) – Correlations

Note: n for all Logistical Regressions is 120882

| Food Stamp | | | | | | | | | | | | |
|----------------------|------------|------------|-------------|--------|--------|------------|----------|--------|--------|---------------|--------------|---------------|
| | Food Stamp | HS Dropout | Post-HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income (log) | Persons in HH |
| Food Stamp | 1 | | | | | | | | | | | |
| HS Dropout | 0.169 | 1 | | | | | | | | | | |
| Post-HS Edu | -0.173 | -0.362 | 1 | | | | | | | | | |
| Age | -0.080 | 0.030 | -0.094 | 1 | | | | | | | | |
| Black | 0.131 | 0.044 | -0.060 | -0.004 | 1 | | | | | | | |
| Other Race | -0.012 | -0.006 | 0.037 | -0.037 | -0.115 | 1 | | | | | | |
| Hispanic | 0.069 | 0.156 | -0.135 | -0.134 | -0.111 | -0.076 | 1 | | | | | |
| DivSep | 0.114 | 0.020 | -0.036 | 0.059 | 0.053 | -0.042 | 0.006 | 1 | | | | |
| Widow | 0.015 | 0.054 | -0.089 | 0.338 | 0.024 | -0.016 | -0.045 | -0.101 | 1 | | | |
| Never Married | 0.132 | 0.030 | -0.023 | -0.310 | 0.160 | 0.014 | 0.058 | -0.190 | -0.118 | 1 | | |
| Income (log) | -0.330 | -0.181 | 0.252 | -0.107 | -0.153 | 0.036 | -0.067 | -0.166 | -0.154 | -0.133 | 1 | |
| Persons in HH | 0.102 | 0.041 | -0.025 | -0.351 | -0.066 | 0.082 | 0.134 | -0.161 | -0.180 | -0.093 | 0.259 | 1 |

| Rent Assistance | | | | | | | | | | | | |
|----------------------|-------------|------------|-------------|--------|--------|------------|----------|--------|--------|---------------|--------------|---------------|
| | Rent Assist | HS Dropout | Post-HS Edu | Age | Black | Other Race | Hispanic | DivSep | Widow | Never Married | Income (log) | Persons in HH |
| Govt. Prog. | 1 | | | | | | | | | | | |
| HS Dropout | 0.036 | 1 | | | | | | | | | | |
| Post-HS Edu | -0.042 | -0.362 | 1 | | | | | | | | | |
| Age | -0.009 | 0.030 | -0.094 | 1 | | | | | | | | |
| Black | 0.072 | 0.044 | -0.060 | -0.004 | 1 | | | | | | | |
| Other Race | 0.004 | -0.006 | 0.037 | -0.037 | -0.115 | 1 | | | | | | |
| Hispanic | 0.006 | 0.156 | -0.135 | -0.134 | -0.111 | -0.076 | 1 | | | | | |
| DivSep | 0.039 | 0.020 | -0.036 | 0.059 | 0.053 | -0.042 | 0.006 | 1 | | | | |
| Widow | 0.014 | 0.054 | -0.089 | 0.338 | 0.024 | -0.016 | -0.045 | -0.101 | 1 | | | |
| Never Married | 0.066 | 0.030 | -0.023 | -0.310 | 0.160 | 0.014 | 0.058 | -0.190 | -0.118 | 1 | | |
| Income (log) | -0.132 | -0.181 | 0.252 | -0.107 | -0.153 | 0.036 | -0.067 | -0.166 | -0.154 | -0.133 | 1 | |
| Persons in HH | -0.035 | 0.041 | -0.025 | -0.351 | -0.066 | 0.082 | 0.134 | -0.161 | -0.180 | -0.093 | 0.259 | 1 |

*Note: None of the correlations in each regression are close to 0.8, which is the rough value to determine collinearity for this paper. Therefore, there is no evidence of simple multicollinearity among the variables.

Appendix II – VIF

| Variable | VIF | 1/VIF |
|----------------------|------------|--------------|
| Age | 1.45 | 0.6887 |
| Black | 1.09 | 0.91911 |
| DivSep | 1.15 | 0.86804 |
| Hispanic | 1.1 | 0.90978 |
| HS Dropout | 1.18 | 0.84543 |
| Income (log) | 1.25 | 0.79908 |
| Never Married | 1.31 | 0.76441 |
| Other Race | 1.03 | 0.96794 |
| Persons in HH | 1.35 | 0.73985 |
| Post-HS Edu | 1.23 | 0.81239 |
| Widow | 1.19 | 0.8397 |
| Mean VIF | 1.21 | |

***Note:** None of the VIF's are above the value of 5, which is the value used to determine multicollinearity in this paper. In fact, all values are below 1.5, therefore, there is no multicollinearity among the variables.

Appendix III – Logit Models

Note: n for all Logistical Regressions is 120882

| Any Government Program | | | |
|-------------------------------|--------------|-----------|----------------|
| | Coef. | SE | z-score |
| HS Dropout | 0.487 | 0.028 | 17.4 |
| Post-HS Edu | -0.478 | 0.018 | -26.69 |
| Age | -0.078 | 0.003 | -25.29 |
| Black | 0.579 | 0.023 | 25.57 |
| Other Race | 0.001 | 0.031 | 0.03 |
| Hispanic | 0.386 | 0.022 | 17.79 |
| DivSep | 0.965 | 0.023 | 41.43 |
| Widow | 0.537 | 0.039 | 13.81 |
| Never Married | 0.568 | 0.023 | 25.08 |
| Income (log) | -0.714 | 0.009 | -81.55 |
| Persons in HH | 0.510 | 0.006 | 86.82 |
| Constant | 0.334 | 0.055 | 6.1 |

| Free / Reduced Lunch | | | |
|-----------------------------|--------------|-----------|----------------|
| | Coef. | SE | z-score |
| HS Dropout | 0.3798 | 0.03502 | 10.85 |
| Post-HS Edu | -0.494 | 0.02413 | -20.49 |
| Age | -0.17 | 0.00457 | -37.27 |
| Black | 0.6706 | 0.03053 | 21.97 |
| Other Race | -0.004 | 0.04086 | -0.1 |
| Hispanic | 0.6812 | 0.02663 | 25.58 |
| DivSep | 0.8979 | 0.03087 | 29.09 |
| Widow | 0.4964 | 0.06472 | 7.67 |
| Never Married | 0.0077 | 0.03019 | 0.25 |
| Income (log) | -0.609 | 0.00962 | -63.32 |
| Persons in HH | 0.7048 | 0.00739 | 95.34 |
| Constant | -0.847 | 0.06753 | -12.55 |

| Public Assistance | | | |
|--------------------------|--------------|-----------|----------------|
| | Coef. | SE | z-score |
| HS Dropout | 0.397 | 0.071 | 5.61 |
| Post-HS Edu | -0.478 | 0.057 | -8.35 |
| Age | -0.059 | 0.010 | -6.08 |
| Black | 0.300 | 0.066 | 4.57 |
| Other Race | 0.093 | 0.090 | 1.03 |
| Hispanic | 0.041 | 0.066 | 0.61 |
| DivSep | 1.099 | 0.071 | 15.59 |
| Widow | 0.753 | 0.129 | 5.81 |
| Never Married | 1.163 | 0.064 | 18.11 |
| Income (log) | -0.400 | 0.015 | -26.87 |
| Persons in HH | 0.363 | 0.012 | 29.59 |
| Constant | -3.875 | 0.141 | -27.52 |

| Food Stamps | | | |
|----------------------|--------------|-----------|----------------|
| | Coef. | SE | z-score |
| HS Dropout | 0.533 | 0.032 | 16.89 |
| Post-HS Edu | -0.584 | 0.024 | -24.76 |
| Age | -0.054 | 0.004 | -13.6 |
| Black | 0.573 | 0.027 | 21.03 |
| Other Race | -0.105 | 0.042 | -2.48 |
| Hispanic | 0.060 | 0.029 | 2.09 |
| DivSep | 1.282 | 0.029 | 44.36 |
| Widow | 0.868 | 0.047 | 18.47 |
| Never Married | 1.035 | 0.028 | 36.48 |
| Income (log) | -0.821 | 0.010 | -82.23 |
| Persons in HH | 0.403 | 0.007 | 59.54 |
| Constant | -0.244 | 0.066 | -3.68 |

| Rent Assistance | | | |
|------------------------|--------------|-----------|----------------|
| | Coef. | SE | z-score |
| HS Dropout | 0.212 | 0.084 | 2.51 |
| Post-HS Edu | -0.410 | 0.063 | -6.5 |
| Age | -0.031 | 0.010 | -2.96 |
| Black | 0.898 | 0.065 | 13.78 |
| Other Race | 0.601 | 0.102 | 5.9 |
| Hispanic | 0.225 | 0.082 | 2.74 |
| DivSep | 1.249 | 0.084 | 14.79 |
| Widow | 0.994 | 0.121 | 8.21 |
| Never Married | 1.322 | 0.082 | 16.13 |
| Income (log) | -0.428 | 0.015 | -29.09 |
| Persons in HH | -0.033 | 0.021 | -1.56 |
| Constant | -3.431 | 0.173 | -19.85 |

Fiber Optic Shape-sensing for Woven Materials

Jack Kawell, Drew Hackney, Guodong Guo, Luca Girolamo, Kara Peters, Mark Pankow

Abstract

This research presents a sensing technique that uses fiber Bragg gratings (FBGs) to measure the out-of-plane deformation of woven composites during quasi-static impact events. Measuring the transient deformation is important to understanding how these materials absorb energy during impacts. Currently, the primary measurement technique records the final deformation of a clay backing material after the impact. However, this method precludes important dynamic measurements during the impact event. We integrated FBGs into a woven Kevlar fabric, circularly clamped the fabric, and then applied an increasing load to the middle of the sheet. We used the FBGs as distributed strain sensors to measure in-plane strains as the fabric was deformed out-of-plane. By running these calculated strains through a mathematical algorithm, we were able to reconstruct the out-of-plane deformation of the Kevlar during the impact. We compared the out-of-plane deformation calculated from the FBGs to the deformation independently calculated from a stereo camera setup and digital image correlation. This new FBG sensing technique of reconstructing deformations in composites could eventually lead to more accurate testing techniques for applications in a variety of fields from aerospace to body armor.

1. INTRODUCTION

1.1 Current Data Collection Methods for High Speed Impacts

When testing composites in high speed impacts, understanding the dynamic forces occurring throughout the impact is important for developing materials that are both stronger and lighter. There are currently two testing methods used for composites in ballistic events. The first is to place a soft, malleable backing material behind the composite to be tested. When the front composite is impacted, the backing material is deformed by the back face deformation (BFD) of the material and retains the BFD shape so that the depth and shape of the deformation caused by the impact can be measured [1]. The problem with this technique is that it misses the important dynamic measurements that occur during the impact event and only show the final state of the composite. The second method of testing is to film impacts with high speed cameras and measure the deformation during the impact with digital image correlation (DIC) [2, 3]. This allows for the movement of the composite to be measured throughout the impact, but can only be done when no backing material is present. If the composite must be tested with solid backing materials that prohibit the line of sight necessary for high speed cameras, then this method is not useful. To resolve these issues in current data collection methods, this paper proposes the use of fiber Bragg grating (FBG) sensors to collect the dynamic, high speed data without need for line of sight.

1.2 Fiber Bragg Grating (FBG) Sensors

FBGs are a particular grating written into an optical fiber that reflect specific wavelengths of light. When a broadband spectrum of light is passed through the fiber, the reflected spectrum can be collected by an interrogator. A diagram of FBG principles is shown in Figure 1.

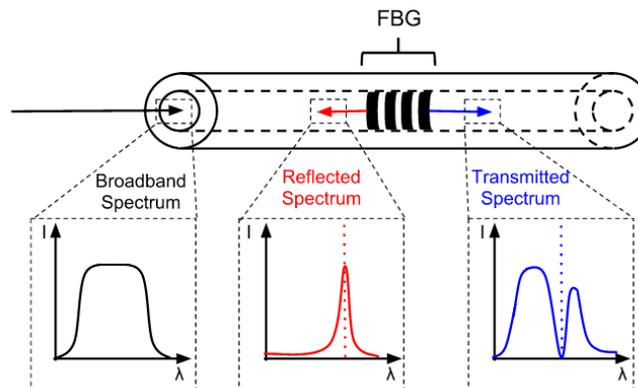


Figure 1. The reflected and transmitted spectrum response due to a fiber Bragg grating written in an optical fiber.

The wavelength of the reflected spectrum changes as a reliable function of the strain applied to the FBG as given by equation 1 [4].

$$\varepsilon = \frac{\Delta\lambda_B}{\lambda_B} \frac{1}{1 - p_e} \quad [1]$$

The variable ε is the calculated strain, $\Delta\lambda_B$ is the change in Bragg wavelength of the reflected spectrum, λ_B is the initial Bragg wavelength of the FBG, and p_e is the photoelastic coefficient of the FBG. This photoelastic coefficient is a property of the grating that couples wavelength shift to strain. By measuring the change in wavelength of the reflected spectrum, the strain across the grating can be reliably calculated. Because FBGs are written into optical fibers, their size allows them to be integrated into almost any composite without damaging the structure or changing the material properties of the host material. FBGs have been shown to maintain their sensing properties after embedding. Sensing applications have ranged from reinforced composite laminates [5], to concrete [6], to flexible textiles [7]. Also, because the data is collected through the use of reflected light, the amount of data collected during a high-speed event is limited only by the data acquisition equipment. Optical interrogators can be manufactured to collect full FBG data in the 100 kHz range, which is fast enough to capture high-speed impact events [8]. The combination of these two assets allows for FBGs to be effective in nonintrusive testing of composites that require no line of sight and can collect high speed data throughout the impact event.

This paper will describe initial experiments used to build a mathematical model to convert in-plane strain measured by FBGs embedded in Kevlar® to out-of-plane deformation. The samples will be deformed under a quasi-static load and out-of-plane deformation will be independently measured by DIC. It will also describe the initial mathematical theory used on the in-plane strain measurements taken during the impact of a Mylar membrane sample to obtain the out-of-plane deformation. Then the paper will apply this model to a Kevlar® sheet and discuss how the membrane theory model must be adapted to fit the physical properties of the Kevlar® fabric.

2. MATHEMATICAL MODEL

The in-plane strain to out-of-plane deformation model is based off an analytical solution of a circularly clamped membrane. A deformation force is placed on the center of the clamped surface of the membrane. Due to the circular geometry of the clamped membrane, a cylindrical coordinate system is used for this problem. A diagram of the coordinate system is shown in Figure 2.

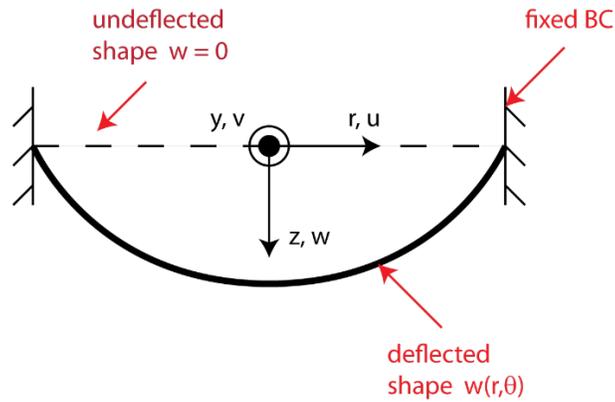


Figure 2. Coordinate system for out-of-plane deflection

The variables u , v , and w are the deformations in the r , θ , and z -directions, respectively. Out-of-plane deformation is defined as deformation in the z -direction.

The equations for strain in the radial, tangential, and z -directions for given deformations are shown in Equations 2-4.

$$\varepsilon_r = \frac{\partial u}{\partial r} + \frac{1}{2} \left[\left(\frac{\partial u}{\partial r} \right)^2 + \left(\frac{\partial v}{\partial r} \right)^2 + \left(\frac{\partial w}{\partial r} \right)^2 \right] \quad [2]$$

$$\varepsilon_\theta = \frac{u}{r} + \frac{1}{r} \frac{\partial v}{\partial \theta} + \frac{(u^2 + v^2)}{2r^2} + \frac{1}{2r^2} \left[u \frac{\partial v}{\partial \theta} - v \frac{\partial u}{\partial \theta} \right] + \frac{1}{2} \left[\left(\frac{\partial u}{\partial \theta} \right)^2 + \left(\frac{\partial v}{\partial \theta} \right)^2 + \left(\frac{\partial w}{\partial \theta} \right)^2 \right] \quad [3]$$

$$\varepsilon_z = \frac{\partial w}{\partial z} + \frac{1}{2} \left[\left(\frac{\partial u}{\partial z} \right)^2 + \left(\frac{\partial v}{\partial z} \right)^2 + \left(\frac{\partial w}{\partial z} \right)^2 \right] \quad [4]$$

Since the problem is axisymmetric, deformations are independent of tangential deformation ($v = 0$). Additionally, deformation is independent of the z -direction strain because the fabric is thin. With these simplifying features in mind, Equations 2-4 reduce to Equations 5-7.

$$\varepsilon_r = \frac{\partial u}{\partial r} + \frac{1}{2} \left[\left(\frac{\partial u}{\partial r} \right)^2 + \left(\frac{\partial w}{\partial r} \right)^2 \right] \quad [5]$$

$$\varepsilon_\theta = \frac{u}{r} + \frac{u^2}{2r^2} \quad [6]$$

$$\varepsilon_z = 0 \quad [7]$$

Due to the clamped boundary conditions, the deformation at the edge is equal to zero. The clamped conditions constrain the material, and radial deformation is small compared to out of plane deformation. For the assumption that u is much smaller than w , Equations 5-7 can be further simplified to the equations given as Equations 8-10.

$$\varepsilon_r \cong \frac{1}{2} \left(\frac{\partial w}{\partial r} \right)^2 \quad [8]$$

$$\varepsilon_\theta \cong \frac{u}{r} \quad [9]$$

$$\varepsilon_z = 0 \quad [10]$$

The radial strain, ε_r , is measured by integrated sensors. Rearranging Equation 7 gives a differential equation to solve for z-deformation, w , as a function of radial strain shown in Equation 11.

$$\frac{\partial w}{\partial r} = (\varepsilon_r)^{1/2} \quad [11]$$

The deflection, w , from Equation 11 can be assumed to be in the form of a polynomial with respect to radial position, r . The parameters of the polynomial will be determined by curve fitting. A cubic spline is chosen to be the shape of the deflection curve. The equation of the spline for $w(r)$, with its first and second derivatives are shown in Equation 12.

$$\begin{aligned} w(r) &= A_0 + A_1 r + A_2 r^2 + A_3 r^3 \\ w'(r) &= A_1 + 2A_2 r + 3A_3 r^2 \\ w''(r) &= 2A_2 + 6A_3 r \end{aligned} \quad [12]$$

In the above equations, A_i are the parameters of the polynomial.

In the experiment, the deflection can be described as having two distinct regions. Near the impactor head, deformation changes rapidly with r , and away from the impactor head, deformation changes much more slowly. Capturing both these regions with a single function may lead to inaccuracies; therefore, the radial distance is split into two segments with the dividing line at the midpoint.

Splitting the deformation into two discrete segments results in 8 unknowns (A_0 through A_3 for each segment). To solve for the 8 unknowns, two boundary conditions are used as given below:

- 1) At $r = 0$; $w' = 0$ (slope at the center of the clamped area)
- 2) At $r = a$; $w = 0$ (deformation at the clamped boundary)

Additionally, 3 continuity conditions are used. These ensure that the deformation, w , and its first and second derivatives of each segment are equal at the location where two segments meet.

- 3) At $r = \frac{1}{2}a$, $w_1 = w_2$ (deformation is equal between segment 1 (w_1) and segment 2 (w_2)).
- 4) At $r = \frac{1}{2}a$, $w_1' = w_2'$
- 5) At $r = \frac{1}{2}a$, $w_1'' = w_2''$

Finally, the strains from each of the sensors are known. The sensors are split such that there are two sensors in each segment. This results in 9 equations (2 boundary, 3 continuity, and 4 strain) to solve for 8 unknowns. The parameters are solved for using Equation 13.

$$\mathbf{A} = [\mathbf{S}^T \mathbf{S}]^{-1} \mathbf{S}^T \mathbf{B} \quad [13]$$

For the above equation, \mathbf{A} is an array consisting of the unknown parameters, \mathbf{S} is a matrix consisting of the boundary conditions, continuity conditions, and strain data, and \mathbf{B} is an array consisting of zeros for the boundary and continuity conditions, and strain data for each sensor location.

3. EXPERIMENTAL RESULTS AND DISCUSSION

The above model was developed assuming a membrane being deformed. The strain data from electric strain gages in a quasi-static impact experiment on Mylar was used to test to whether the above model is accurate. Then a similar experiment was run using FBGs woven into a Kevlar® fabric to test the model's viability on a fabric.

3.1 Quasi-static Impact Experiment Using Mylar

The sample tested was instrumented with four Micro Measurements strain gages (model EA-06-015DJ-120/L) adhered to the surface of a 0.22 mm thick Mylar sample using M-Bond AE-10, also produced by Micro Measurements. AE-10 is room temperature cure adhesive with high elongation capabilities, between 6% and 10%. The bottom surface of the sample was painted using a speckle pattern to allow for independent measurement of the out-of-plane deformation using the DIC camera setup. Images of the sample are shown in Figures 3 and 4.



Figure 3. Upper surface of the Mylar sample showing attached strain gages.

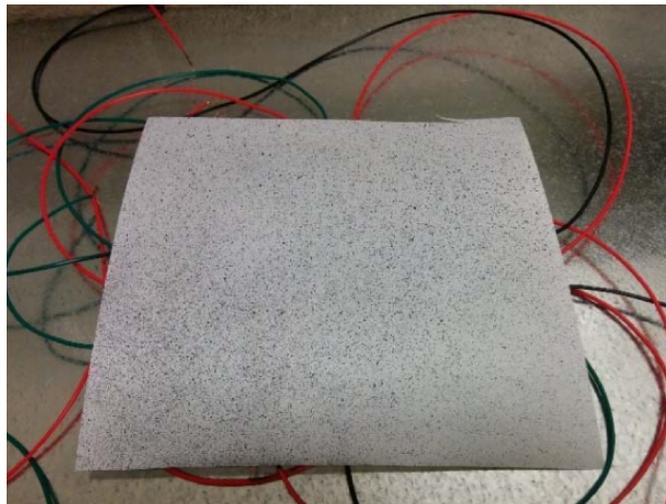


Figure 4. Lower surface of the Mylar sample showing painted on speckle pattern for DIC.

The sample was circularly clamped in a low-velocity impact tower. The radius of the test aperture is 3.8 cm. The clamps were held tightly together by 8 bolts. A mirror angled at 45° was placed below the sample to allow for the stereo camera setup to image the bottom surface of the sample. Figure 5 shows the way the experiment was set up. The cameras used for DIC were two Grasshopper2 high resolution cameras produced by PointGrey. Two LED lamps were used to illuminate the lower surface of the sample. VIC3D software was used for the 3-D DIC analysis.

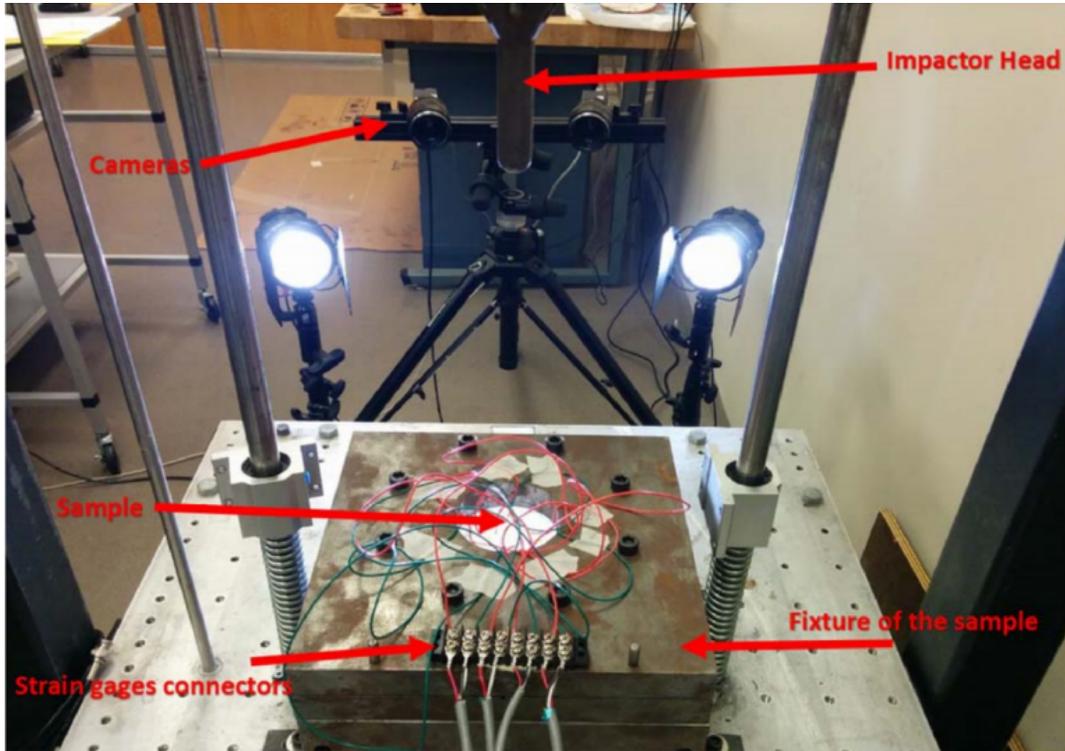


Figure 5. Mylar test experimental setup.

Six separate samples were deformed by slowly lowering the impactor crosshead onto the top surface of the samples. The crosshead weight was 5.36 kg, and the full weight rested on the center of the sample during the test. The strain gages were wired in a quarter bridge through a model 2300 System Signal Conditioning Amplifier manufactured by Vishay Measurements Group. The voltage change measured by the four strain gages during a test is shown in Figure 6. Strain gage 1 is closest to the impact location and Strain gage 4 is located furthest away.

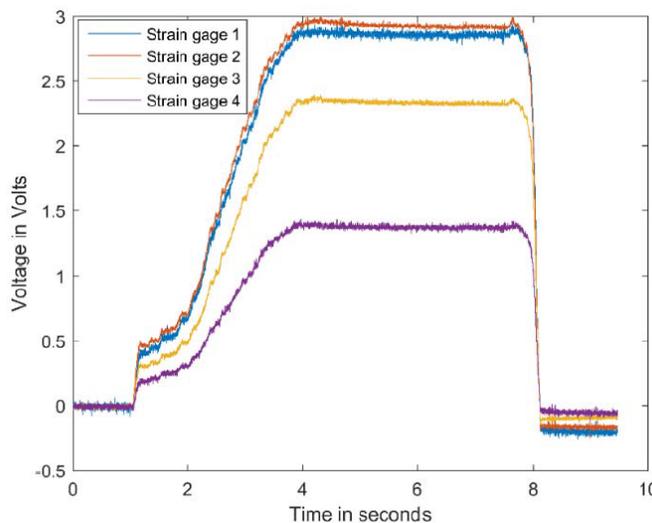


Figure 6. Strain output by the four adhered strain gages during the Mylar test.

The strain gage voltage was converted to strain using Equation 14.

$$\mu\varepsilon = \frac{4V_{out}}{(A * V_{BR} * K)10^{-6}} \quad [14]$$

For the above equations, $\mu\varepsilon$ is the strain in microstrain, V_{out} is the voltage output by the individual strain gages, and V_{BR} is the input voltage for each strain gage. In this case, $V_{BR} = 0.5$ V. The variable K is the gage factor and the amplifier gain A is defined as 2,000 for this experiment.

From the strain gage measured strains, deformation, w , was calculated from Equation 12. A plot comparing the DIC measured deformation with the deformation reconstructed from the mathematical model is shown in Figure 7.

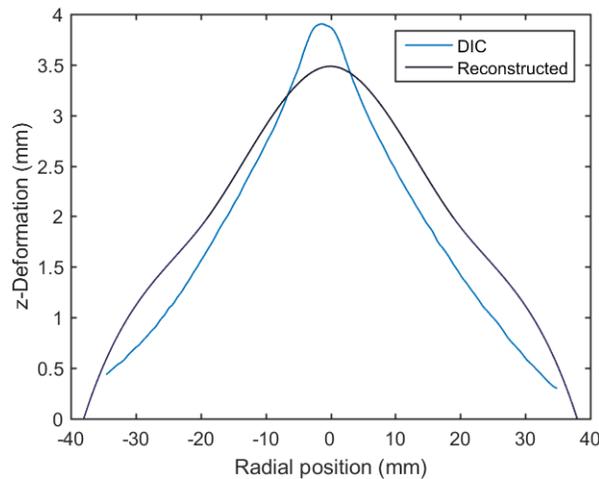


Figure 7. Comparison of the DIC measured deformation and the deformation calculated by the mathematical model for the Mylar sample.

After the Mylar test, the DIC system calculated an accurate reading of the deformation caused by the impactor load. The mathematical model from above was used on the strain data from the strain gages to find calculated values of the deformation to compare for accuracy with the deformations measured by the DIC standard. As shown in Figure 7, the reconstructed deformations calculated from the strain values collected by the gages matched well with those obtained by the DIC system. The reconstruction stays within ~ 0.5 mm of the DIC standard throughout the deformation curve. Though the maximum deformation is missed in the reconstruction, the overall shape and size of the deformation is accurately depicted by the mathematical model. This shows that the model is effective when used upon a circularly clamped membrane such as the Mylar sample.

3.2 Quasi-static Impact Experiment Using Kevlar®

The Kevlar® sample used in the experiment is a plain weave with 1 mm tow width and 0.5 mm tow thickness. The strain gages were replaced with four os1100 FBGs produced by Micron Optics interrogated by a Micron Optics Hyperion interrogator. As seen in Figure 8, the four FBGs were placed at 0.75, 1.25, 1.75, and 2.25 cm from the impact site. The fibers were woven into the fabric using a ‘1x1’ weave where the fiber went 1 cm over and 1 cm under for a single weave. Each FBG was situated underneath the Kevlar® to measure the full strain from the curvature of the fabric in tension during the load event. Also, the fibers were not bonded to the Kevlar® so that they might slip slightly as the fabric deformed. This is because the yield strain of the fabric exceeds the yield strain of the FBGs. In a high-speed dynamic test event, the FBGs would fail before the Kevlar® fabric if they were rigidly bonded to the Kevlar®. Allowing the fibers to slip allows for strain to be released on the optical fiber.

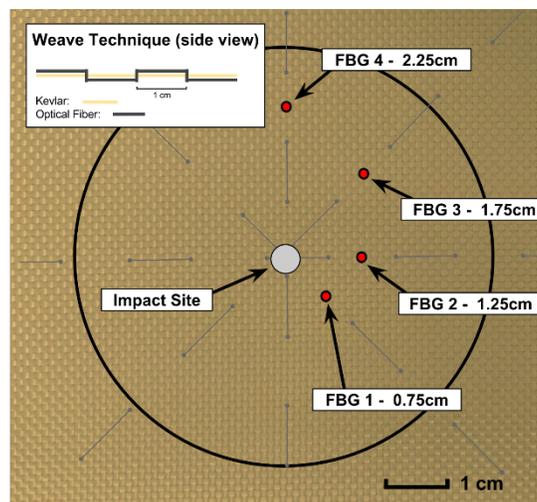


Figure 8. Diagram of the Kevlar® sample showing optical fiber weave and sensor location.

During the test, the impactor was gently lowered onto the sample as both the DIC and the FBGs were collecting data. FBG Bragg wavelengths were recorded before and after the impactor was placed on the sample. Table 1 shows the Bragg wavelengths before and after the sample was loaded, the wavelength shift, and the strain calculated from Equation 1 for each FBG.

Table 1. FBG wavelength data and calculated strains for the Kevlar® test.

| FBG | λ_{B1} (nm) | λ_{B2} (nm) | $\Delta\lambda_B$ (nm) | Strain ($\%\epsilon$) |
|-----|---------------------|---------------------|------------------------|-------------------------|
| 1 | 1564.26 | 1567.07 | 2.81 | 0.229 |
| 2 | 1552.21 | 1554.84 | 2.63 | 0.216 |
| 3 | 1544.07 | 1544.99 | 0.92 | 0.0757 |
| 4 | 1538.00 | 1540.53 | 2.54 | 0.210 |

Just like for the Mylar test, the DIC system calculated an accurate deformation for the Kevlar® test. Deformation values were then calculated from the strain data collected from the

FBGs using the mathematical model like before. Figure 9 shows this reconstructed deformation from the FBG data plotted with the DIC system standard.

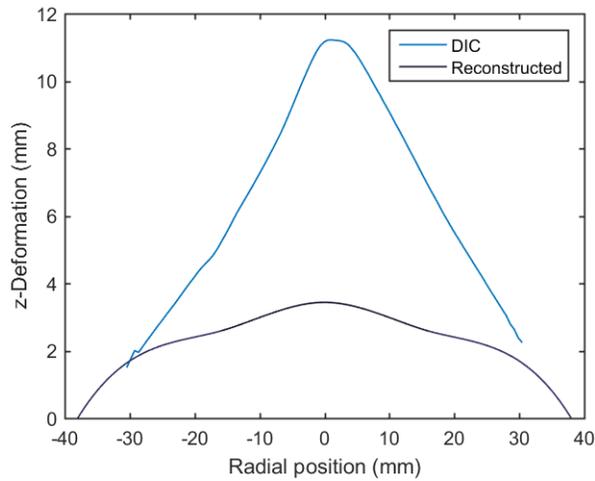


Figure 9. Comparison of the DIC measured deformation and the deformation calculated by the mathematical model for the Kevlar® sample.

Unlike with the Mylar, the model severely underestimates the deformation caused by the impact on the Kevlar®. The maximum deformation measured by the mathematical model is 3.8 mm, compared to 11.22 mm for the DIC data, a difference of 7.42 mm. However, the model reconstructed deformation has a comparable maximum deflection to the Mylar test. This disparity between the DIC standard and the model's calculated values are due to two reasons. First, the FBGs experience slip within the Kevlar® fabric during this test. This caused less strain to be transferred from the fabric to the FBGs than with the strain gages. Second, the sample was not clamped tightly enough in the mount so the clamped boundary condition assumption was violated. The first factor causes the mathematical model to underestimate the deflection. The second factor results in the sample slipping in the mount and the sample z-direction deflection being significantly greater.

4. CONCLUSIONS

The mathematical model worked very well when applied to a circularly clamped membrane such as Mylar, though there is room to improve the model to be more mechanics based. However, when applied to a fabric such as Kevlar®, the model was inaccurate in converting the in-plane strain measurements of the FBGs to correct out-of-plane deformations. Part of the inaccuracy is due to expected slip in the FBGs. The model will be improved in the future to factor for FBG slip. The primary source of error in the model is due to the poor clamping of the sample in the mount, resulting in a violated boundary condition and unexpectedly large sample deformation. The experimentation showed that FBGs show promise as non-intrusive strain sensors in composite fabric testing and the mathematical model developed in this paper serves as a valid starting point to create a more robust model for the future.

5. ACKNOWLEDGEMENTS

This material is based upon work supported by the National Science Foundation under Grant No. 1460935.

6. REFERENCES

1. Park, J.L., Chi, Y.S., Hahn, M.H., & Kang, T.J. “Kinetic Dissipation in Ballistic Tests of Soft Body Armors.” *Experimental Mechanics*, no. 52 (2012): 1239-1250.
2. Schmidt, T., Tyson, J., & Galanulis, K. “Full-field dynamic deformation and strain measurements using high-speed digital cameras.” *26th International Congress on High-Speed Photography and Photonics*, (2004): 174-183.
3. Freitas, C.J., Bigger, R.P., Scott, N., LaSala, V., & MacKiewicz, J. “Composite materials dynamic back face deflection characteristics during ballistic impact.” *Journal of Composite Materials*, no. 48 (2014): 1475-1486.
4. Melle, S.M., Liu, K., & Measures, R.M.. “Strain sensing using a fiber-optic Bragg grating.” *Conference on Fiber Optic Smart Structures and Skins IV*, (1991): 255-263.
5. Measures, Raymond M. “Fiber optic sensing for composite smart structures.” *Composites Engineering*, no. 3 (1993): 715-750.
6. Davis, M.A., Bellemore, D.G., & Kersey, A.D. “Distributed fiber Bragg grating strain sensing in reinforced concrete structural components.” *Cement and Concrete Composites*, no. 19 (1997): 45-57.
7. Artières, O. & Dortland, G. “A fiber optics textile composite sensor for geotechnical applications.” *Fourth European Workshop on Optical Fibre Sensors*, (2010).
8. Vella, T., Chadderdon, S., Selfridge, R., Schultz, S., Webb, S., Park, C., Peters, K., & Zikry, M. “Full-spectrum interrogation of fiber Bragg gratings at 100 kHz for detection of impact loading.” *Measurement Science and Technology* 21, no. 9 (2010).

Analysis of Riverbank Restoration Initiatives in a Riparian Ecosystem

Jacqueline Taylor

Abstract

Alabama's aquatically biodiverse waterways are bordered and protected by riparian vegetation that support aquatic ecosystem health through ecological benefits like nutrient cycling, water temperature regulation, and erosion control (Richardson et al., 2007). Due to their economic value and geographic location, riparian habitats are commonly overused, and less resilient to detrimental effects from the disturbances they are prone to. Exotic species tend to invade and outcompete native plant populations for resources after a disturbance has altered the plant community's structure and function. Active restoration of degraded riparian riverbank involves reintroduction of native plant species, and removal of exotic plant species, though this method is often reported with mixed, long-term success. Length of time and cost deters post-restoration management efforts, despite increased chances for full ecosystem recovery. There is extensive, ongoing research to determine the most efficient restoration techniques and methods, including analyzing how characteristics of each restoration zone affects these efforts. Two riparian sites where active restoration had been implemented one, and five years' prior, were used in this study. At each site, changes in vegetation type and abundance, between and within restored and non-restored riverbank areas were observed. It was expected that there would be more exotic plant species, than native plant species in the restored zone, and more native plant species, than exotic plant species in the non-restored zone. Results appear as representative of each site's individual physical characteristics, and time elapsed since restoration. The most significant results showed that at the site of one year since restoration, the restored riverbank had a greater mean abundance of native plant species, than exotic plant species, contrary to expectations. Conversely, at the site of five years since restoration, the restored riverbank had a greater mean abundance of exotic plant species than native plant species, in conjunction with predictions. This comparison between restored sites may represent natural successional growth processes after an active restoration. This type of restoration and its methods may act as a disturbance to the plant community, allowing exotic species to invade, and eventually outcompete native plants without continued post-restoration maintenance. Future restoration projects might consider more specific restoration methods that can meet the individual needs of each degraded ecosystem.

Introduction

Alabama is one of the most aquatically biodiverse states in the United States, and its clean water and aquatic ecosystem health relies heavily on the prosperity of the riparian vegetation that borders its waterways (Gregory et al., 1991). Riparian ecosystems provide ecological and economic benefits, and the riparian plant community is an essential reason for this. These plant communities support ecosystem functions like nutrient cycling, sediment and erosion control, water temperature regulation, and stream flow stabilization (Richardson et al., 2007). Though characteristically resilient, human-caused overuse of these ecosystems has increased their susceptibility to, and frequency of, disturbances (Richardson et al., 2007).

Disturbances are any event that disrupts ecosystem structure and function by altering the physical environment, and availability of substrate and resources (Pickett & White, 1985). These negative effects provide an ideal environment for exotic plants to invade weakened riparian habitat, and compete with native plants for space, nutrients, water, and light (Osawa et al., 2013, Richardson et al., 2007). Invasive species alone can completely degrade ecosystem function by increasing riverbank erosion, limiting available resources, altering biogeochemical pathways, and decreasing native plant populations (Richardson et al., 2007; Ruwanza et al., 2013).

Native plants are often definitive to a healthy ecosystem, as they adapt to the local environment, and provide stability for plant community biodiversity (Maunder, 1992). This, along with strong native species populations, builds ecosystem resilience (Thomas et al., 2014). When ecosystems begin to lose their functionality, due to degrading habitat from increased disturbances and over competition from invasive species, restoration initiatives are implemented.

Two types of restoration methods are used to assist a local ecosystem: passive and active. Passive restoration encourages ecosystem self-recovery through minor human-intervention, and mainly in addressing the most prominent ecosystem threat, which is the removal of invasive plants in this case (Ruwanza, 2013). Active restoration might use several methods to address the core problem more directly, like preserving the riparian habitat by removing all vegetation from the riverbank, and then replanting with native species (Ruwanza, 2013). This type of restoration has had mixed success among restoration projects (Beater et al., 2008; Bellingham et al., 2005; Ruwanza, 2013), because, while it works more quickly than passive restoration to remove invasive species, total vegetation removal causes similar plant community successional growth as that of a disturbance (Bellingham et al., 2005; Ruwanza et al., 2013).

Due to mixed success in active, and passive restoration methods, researchers have speculated as to whether leaving degraded ecosystems alone to self-restore is a better option to save time, cost, and effort. Restoration should not require taking a risk as to whether more detriments than benefits may be imposed on the degraded ecosystem (Holmes et al., 2005; 2008). This effort could instead be used for preventing ecosystem degradation in other beneficial areas.

Ecosystem health is multifaceted, requiring extensive research to fully form the development of a restoration plan that can fix major ecosystem threats, while not producing any new ones. Success of a restoration project is thus difficult to define, and difficult to generalize data for cohesion among similar future studies. Quantifying what defines “improvement” of a degraded ecosystem, or the “success” of a restoration project is most often done case-by-case. In studies of habitat restoration, success is commonly defined by the persistence and reproduction of reintroduced plant populations (Godefroid et al., 2011; Primack & Drayton, 1997; Wohl et al., 2005). Limiting factors to restoration success are due to faults in restoration design, implementation, and continued maintenance thereafter, but especially in lack of thorough

scientific knowledge of the degraded ecosystem, so as to design a promising restoration plan for it (Wohl et al., 2005).

This study assessed two sites where riverbank restoration had been initiated using active restorative methods, to determine how the plant community differed between and within the restored and non-restored zones, and with time and post-management. At both sites, all vegetation had been removed from the riverbank, native plant species were reintroduced, and rock vanes were installed to curb erosion, which was the main concern at each degraded site. The first site at the Cahaba River was restored one year prior to this study, and the second site at Shades Creek was restored five years prior. The plant community was measured by total vegetation, and plant type: native species planted at restoration, native species that self-recruited after restoration, and non-native, or exotic species, which may also be invasive. Species richness, and Shannon Diversity Indices were calculated, because these values can provide insight to plant community productivity in the wake of a disturbance (Allan et al., 2011; Pollock et al., 1998).

Analyses of the restoration site at Shades Creek in 2012 and 2014 (Bailey & Brown, 2012; Fleming, 2014) found there was an increase in invasive plant species, and low survivorship of native, self-recruited species since the restoration. Fleming additionally found that there was an increase in native, self-recruited species to the riverbank. At both the Cahaba River and Shades Creek restoration sites, it was expected that restored plots would have greater mean cover of exotic plant species, mean cover of total vegetation, species richness, and plant community diversity, than in the control plots. Within the restored plots, it was expected that vegetation type and quantity would change along the river's reach from up to downstream, and in distance from the river, due to seed dispersal methods along waterways (Johansson et al., 1996). Restored plots were expected to have greater mean cover of exotic species than of total native species. The control plots were expected to have greater mean cover of total native species than of exotic plant species. In all, it was expected that vegetation types and quantities would significantly differ between restored, and non-restored zones.

Methods

Study sites

Two study sites were used, each within the Cahaba River watershed, located in Jefferson County, Alabama (Figure. 1A). The first study site is about 920 meters by 11 meters on the East bank of the Cahaba River, near the Senior Citizen Center, and across from Trussville Middle School, in Trussville, Alabama (Table 1, Figure 1C). A restoration initiative for this bank was conducted in conjunction with the city continuing a sidewalk through the area and remaking the space into a public park. The entire bank was cleared and replanted with plugs and live stakes of native herbaceous shrub, and tree species from February through April of 2015. Post-restoration maintenance was ongoing during this study, in caring for plants, fixing bank fallouts, and planting more native species. While not directly seen, notes at conclusion of the replanting phase suggested removing prominent invasive plant species over time (Eve Brantley pers. comm.). The second study site is about 160 meters by 8 meters on the South bank of Shades Creek at the Samford University intramural fields, in Birmingham, Alabama (Table 1, Figure 1B). Shades Creek is a tributary of the Cahaba River. A restoration initiative in 2011 removed all vegetation from the South bank, and replanted with native plant species. Rock vanes were also installed in the creek to curb erosion. Since the restoration, there has been sporadic maintenance in invasive species removal, done by various Samford faculty and community.

Table 1. GPS coordinates for study sites reaches.

| Location | Cahaba River | Shades Creek |
|------------------------|------------------------------|------------------------------|
| Control upstream | 33°37'51.10"N, 86°36'9.63"W | 33°27'53.19"N, 86°47'5.44"W |
| Restoration upstream | 33°37'48.99"N, 86°36'10.47"W | 33°27'50.57"N, 86°47'11.53"W |
| Restoration midstream | 33°37'34.64"N, 86°36'7.13"W | 33°27'51.58"N, 86°47'9.11"W |
| Restoration downstream | 33°37'23.93"N, 86°36'0.51"W | 33°27'52.62"N, 86°47'6.37"W |
| Control downstream | 33°37'21.75"N, 86°35'58.73"W | 33°27'49.38"N, 86°47'14.25"W |

Data collection



Figure 1. Maps of study sites (clockwise from left) **A.** Large map shows Jefferson County and location of each study site in yellow, with blue pins. Smaller map shows location of Jefferson County in Alabama. **B.** Reach of Shades Creek. **C.** Reach of the Cahaba River.

Data was collected from September 13th through October 30th of 2016. Methods were adapted from previous restoration analyses conducted on the Shades Creek study site (Bailey & Brown, 2012; Flemming, 2014) and have been adapted to include the Cahaba River study site for this project.

Each study zone was divided into upstream, midstream, and downstream sections (Figure 1A, 1B). There were five reaches: three in the restoration zone that encompassed each section, and two adjacent control (non-restored) reaches located in the upstream and downstream sections. Each of the restored reaches had three transects laid perpendicular to the water source: spaced 10 meters apart at Shades Creek, and spaced 20 meters apart at the Cahaba River. Restored transects were at least 20 meters within the restoration zone edges upstream and downstream, while control transects were at least 20 meters away from the restoration zone. Each site had a total of 9 transects in the restored zone, and 2 transects in the control areas.

Three data collection plots were taken on each transect, designated by distance from the river. The first plot was placed 1 meter from the toe of the creek, the second plot was placed 4 meters in, and the third plot was placed 7 meters in. Each site had a total of 27 restored plots, and

6 control plots. Data collection plots were one meter by one-meter quadrats made from detachable PVC pipes.

Data collection parameters included identifying every plant within each plot, and ocularly estimating vegetation and substrate cover percentages for each plot. Substrate covers included: litter, bare ground, and rock. Categories of vegetation cover were the following plant types: total vegetation in plot; native and planted species; native and self-recruited species; and exotic species. Native, planted species were those planted at the onset of the restoration. Native, recruited species were those plants that had self-recruited to the restored zone without intentional introduction. Exotic species are non-native, potentially invasive plant species. Vegetation cover was defined as any vegetative part occupying aerial space in the plot, not necessarily attached to the ground within the plot. Cover estimations were categorized into classes (Table 2). Using this system, all percentage covers recorded within a plot should equal at least 100% cover, which can be exceeded given the volumetric component of estimating aerial coverage (CNRT). Percentages were later adjusted to values to comply with this assumption. Plants were collected and taken to the Samford University's herbarium for identification, and verified by Dr. Larry Davenport. Plants that could not be identified were removed from data analysis.

Table 2. Cover percentage classes.

| Class | Value % | Adjusted value |
|-------|---------|----------------|
| Trace | <1.0 | 0.001 |
| 1 | 1-5 | 3 |
| 2 | 5-25 | 15 |
| 3 | 25-50 | 37.5 |
| 4 | 50-75 | 62.5 |
| 5 | 75-95 | 85 |
| 6 | 95+ | 97 |

Analysis

Mean differences of plant type cover (total, planted, recruited, exotic), species richness, and Shannon Diversity in plot type (restored or control), reach (upstream to downstream), and distance from the river (1m to 7m) were compared by Analysis of Variance (ANOVA) when data was normally distributed. For some data analysis, cover values of native and planted species were combined with native and self-recruited species to measure the cover for total native species. Mean differences between plant type populations were compared with paired t-tests. When data was not normally distributed, the non-parametric Kruskal Wallace test was used. Discrimination between significantly different means relied on Tukey's Honestly Significant Difference test (HSD), unless Levene's Test for variability was significant, in which case Dunnett's T3 post hoc test was used. A result was considered significant if the P value was less than 0.050.

Results

Collected at the Shades Creek site were 70 different plant species from 27 restored plots, and 25 different plant species from control plots. Collected at the Cahaba River site were 90 plant species from 27 restored plots, and 50 plant species from the control plots.

At the Cahaba River site in both plot types, there was greater mean cover of total native plant species, than exotic plant species. There was greater mean cover of total native species than exotic species in restored plots downstream, as well as in restored plots located 7 meters from the river. Specifically, of native plants, there was greater mean cover of native, recruited species than

exotic species, which had greater mean cover than native planted species. There was greater mean cover of total vegetation and exotic plant species cover in the restored plots, than in the control plots. Mean covers of native, planted species differed spatially between reaches along the river.

At the Shades Creek site in the restored plots, there was greater mean cover of exotic plant species, than total native plant species. Specifically, of native plants, there was greater mean cover of exotic species, than native, recruited species, which had greater mean cover than native, planted species. In control plots, there was greater mean cover of total native plant species than exotic plant species. There was greater mean cover of total vegetation in the restored plots than in the control plots, specifically at 1 meter from the creek. There was greater mean cover of exotic species than total native species in plots downstream and upstream. Mean covers of exotic species differed spatially between reaches along the creek.

Species richness was weakly significant between reaches and distance from the water source at both sites, but did not significantly differ in specific comparisons of tested variables. Requirements for analysis comparing Shannon Diversity Indices among tested variables was unable to be met, and no strong differences were observed. Plant type and cover comparisons between distances from the water source were mostly null, or weakly significant.

Cahaba River in Trussville

Comparisons between plant type

Total native; exotic covers within restored plots

There was greater mean cover of total native species than exotic species in restored plots ($t_5=3.533$; $P=0.017$) (Fig. 2). Between reaches, there was greater mean cover of total native species, than exotic species at downstream plots ($t_8=3.603$; $P=0.007$), and no other significant differences between reaches. Between distances from the river, there was greater total native species cover than exotic cover at plots 7 meters from the river ($t_8=2.738$; $P=0.026$), but no other significant differences between distances from the river.

Native, planted; native, recruited; exotic covers in restored plots

There was greater mean cover of native, recruiting species than native, replanted species cover ($t_{26}=5.838$; $P=0.000$), and exotic species cover ($t_{26}=2.741$; $P=0.011$) within the restored zone (Fig. 2). There was also greater mean cover of exotic species than native, replanted species cover ($t_{26}=4.548$; $P=0.000$) (Fig. 2).

Total native; exotic covers in control plots

There was greater mean cover of total native species than exotic species ($t_{26}=4.091$; $P=0.000$) (Fig. 3).

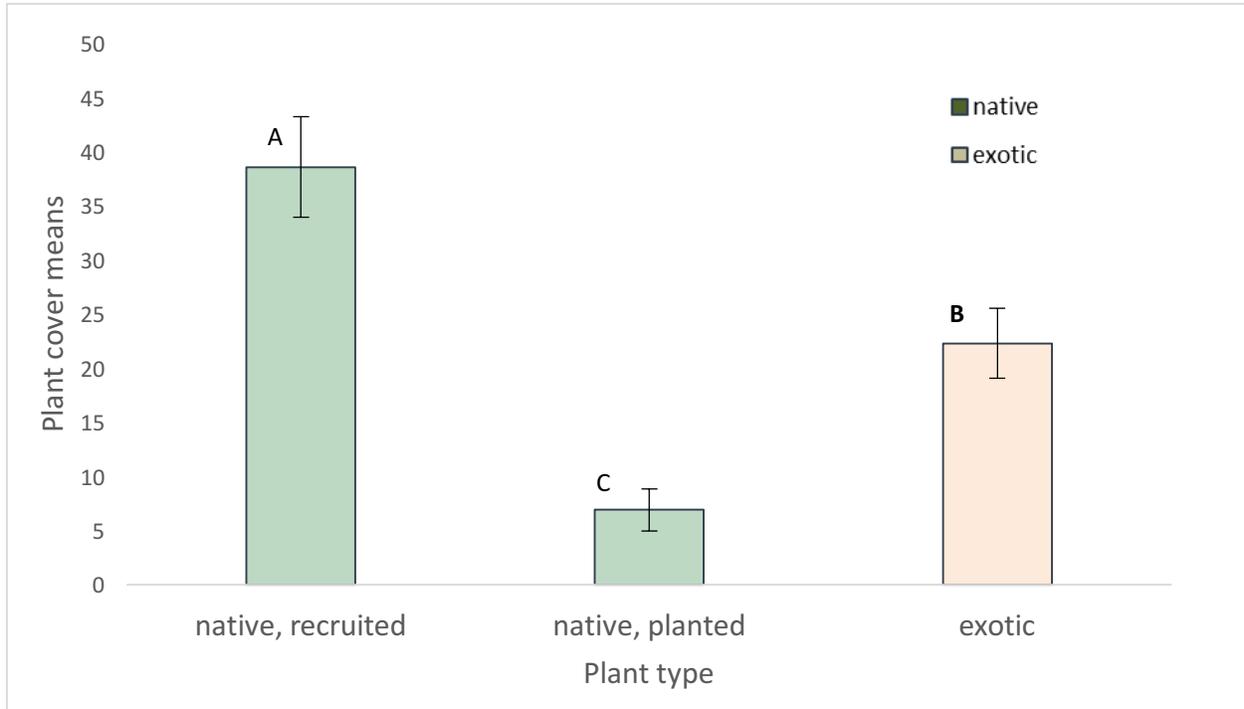


Figure 2. Comparison of plant type mean cover in restored plots at the Cahaba River. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

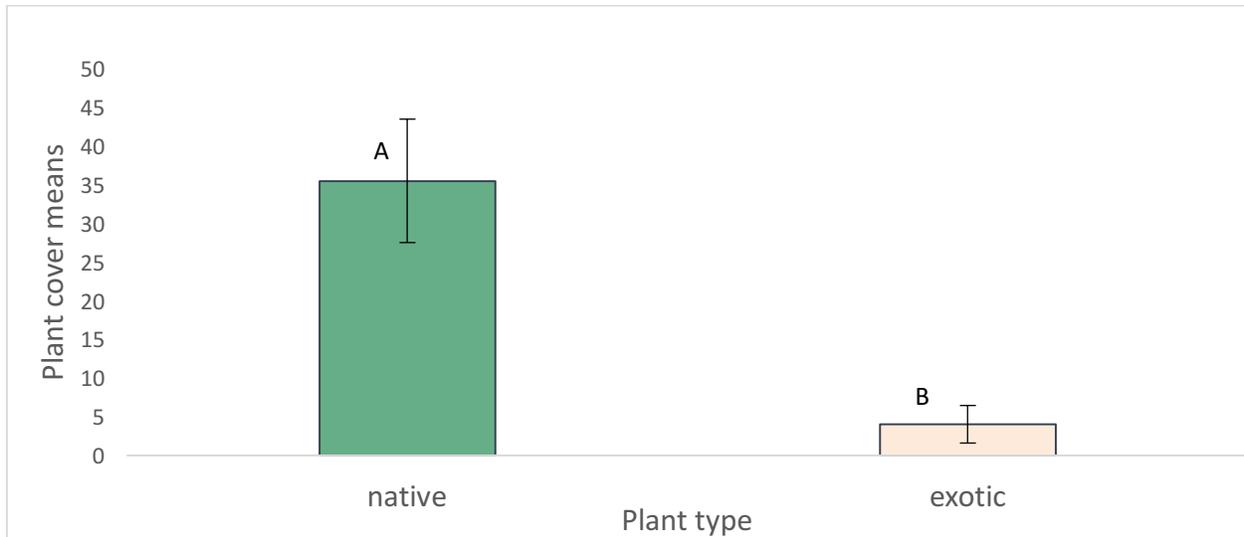


Figure 3. Comparison of mean cover of plant types in control plots at the Cahaba River. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Comparisons between plot type

Frequency

There was no significant difference in frequency of total native species, and exotic species between plot types ($P=0.661$).

Total native species

There was no significant mean difference in total native species cover between plot types ($P=0.343$); along reaches ($P=0.389$); or between distances from the river ($P=0.462$).

Exotic species

Mean cover of exotic species was significantly greater in restored plots, than in control plots ($F_{1,31}=6.834$; $P=0.014$) (Fig. 4). There was no significant mean difference in species cover along reaches ($P=0.064$), nor between distances from the river ($P=0.265$).

Total vegetation

Mean cover of total vegetation was significantly greater in restored plots, than in control plots ($F_{1,31}=6.804$; $P=0.014$) (Fig. 4). There was no significant mean difference between reaches overall ($P=0.081$), but there was a tendency for greater total vegetation cover in midstream and downstream restored plots, than in upstream control plots ($F_{4,33}=2.324$; $P=0.000$ and $P=0.024$ respectively). There was no significant mean difference between distances from the river ($P=0.172$).

Species richness

There was no significant mean difference in species richness between plot types ($P=0.566$); nor between distances from the river ($P=0.686$). There was a significant mean difference between reaches overall ($F_{4,33}=3.186$; $P=0.028$), but there were no specific trends.

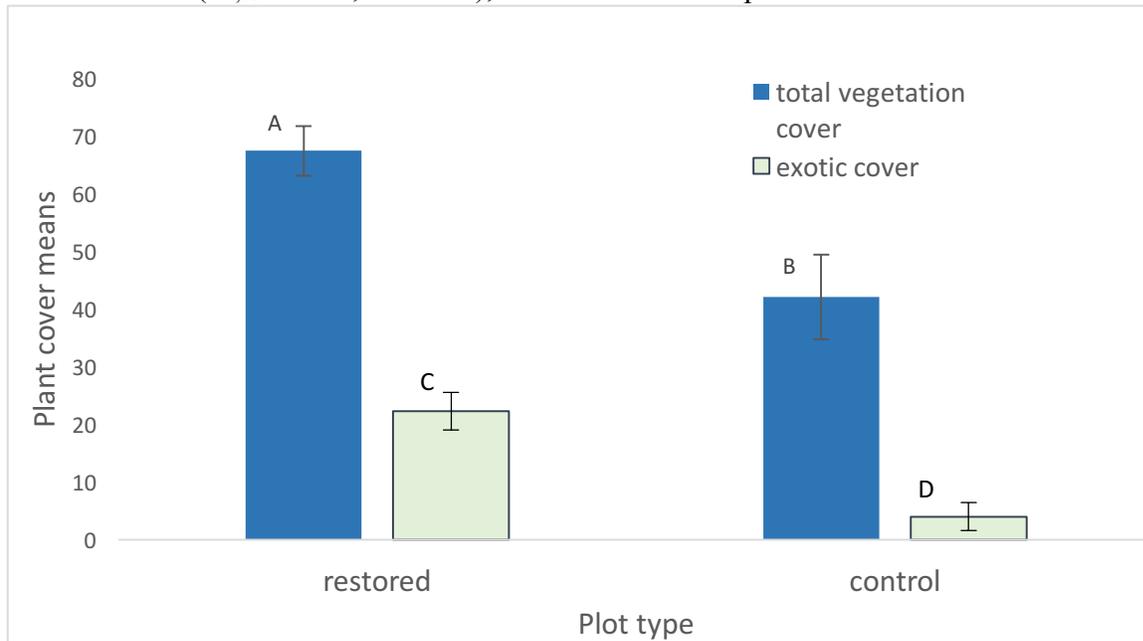


Figure 4. Comparison of mean cover for total vegetation, and exotic species between plot types at the Cahaba River. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Comparisons within restored plots

Native, planted species cover

Mean cover of native, planted species significantly differed between reaches ($F_{2,27}=4.458$; $P=0.023$) (Fig. 5). There was a trend for greater cover in midstream plots, than upstream plots ($F_{2,27}=4.458$; $P=0.018$) (Fig. 5). Mean cover differences between distances from the river could not be compared due to equal distribution across groups ($P=0.201$).

Native, recruited species cover

There was no significant mean difference in native, recruited species covers between reaches ($P=0.606$), nor between distances from the river ($P=0.253$).

Exotic species cover

There was no significant mean difference in exotic species covers between reaches ($P=0.270$), nor between distances from the river ($P=0.764$).

Total vegetation cover

There was no significant mean difference in total vegetation cover between reaches ($P=0.343$), nor between distances from the river ($P=0.778$).

Species richness

Mean species richness did significantly differ across reaches overall ($F_{2,27}=3.576$; $P=0.044$), but non-significantly between individual reach comparisons. There was no significant mean difference in species richness between distances from the river ($P=0.326$).

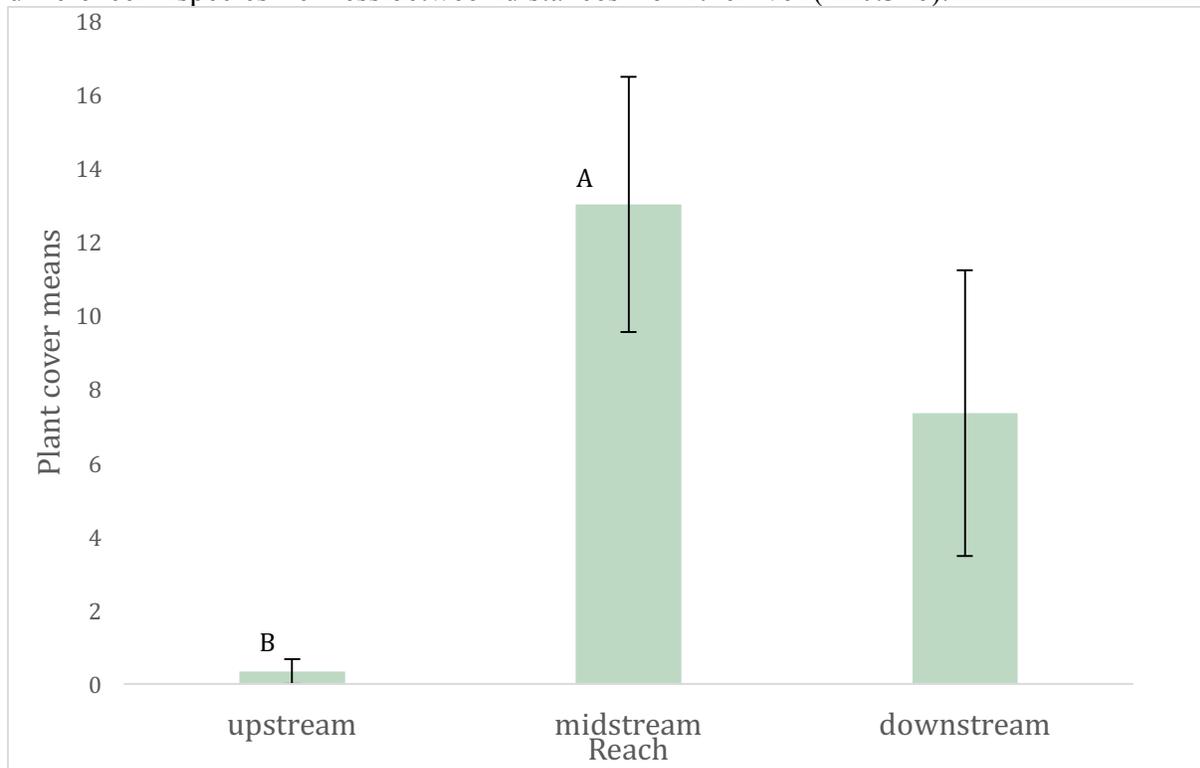


Figure 5. Comparison between of mean cover of native, planted species along the restored reach at the Cahaba River. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Shades Creek in Birmingham

Comparisons between plot type

Frequency

Frequency of total native species and exotic species were significantly different between restored and control plots ($X^2=7.306$; $P=0.007$).

Total native species cover

There was no significant mean difference in total native species covers between plot types ($P=0.405$), nor between reaches ($P=0.068$). Mean differences between distances from the creek could not be compared due to equal distributions ($P=0.933$).

Exotic species cover

There was no significant mean cover difference in exotic species between plot types ($P=0.084$), nor between distances from the creek ($P=0.389$). Mean of exotic covers did significantly differ across reaches overall ($F_{4,33}=3.244$; $P=0.026$), though there were no trends in individual comparisons.

Total vegetation cover

Mean cover of total vegetation was significantly greater in restored plots, than in control plots overall ($F_{1,31}=18.271$; $P=0.000$) (Fig. 6). Mean of total vegetation covers significantly differed across reaches overall ($F_{4,33}=4.627$; $P=0.005$), with a trend for greater cover in downstream, and upstream restored plots, than in downstream control plots ($F_{4,33}=34.627$; $P=0.023$ and $P=0.018$ respectively). Mean covers significantly differed between distances from the creek overall ($F_{5,33}=4.314$; $P=0.005$), with a trend for greater cover 1 meter from the creek in restored plots, than in 1 meter or 4 meter from the creek control plots ($F_{5,33}=4.314$; $P=0.029$ each).

Species richness

There was no significant mean difference in species richness between plot types ($P=0.848$). Other mean differences in cover could not be compared due to equal distributions across reaches ($P=0.301$), and between distances from the creek ($P=0.318$).

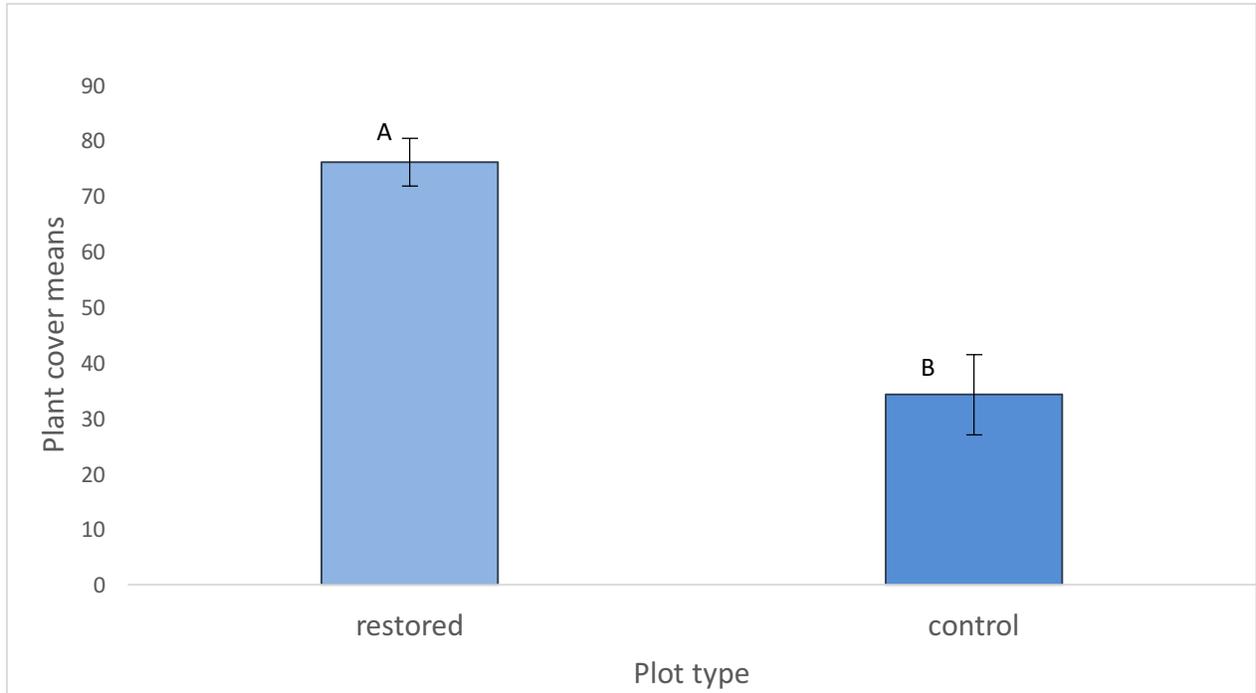


Figure 6. Comparison between plot type of mean cover of total vegetation at Shades Creek. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Comparisons within restored plots

Native, planted species cover

Mean differences in cover could not be compared due to equal distribution across reaches ($P=0.464$), and in distance from the creek ($P=0.873$).

Native, recruited species cover

There was no significant mean difference in native, recruited species covers between reaches ($P=0.130$), nor between distances from the creek ($P=0.974$).

Exotic species cover

Mean cover of exotic species did significantly differ between reaches ($F_{2,27}=4.628$; $P=0.020$), and there was a trend of greater cover in downstream plots, than midstream plots ($F_{2,27}=4.628$; $P=0.027$) (Fig. 7). There was no significant mean difference in exotic species covers between distances from the creek ($P=0.965$).

Total vegetation cover

There was no significant mean difference in total vegetation cover between reaches ($P=0.611$), nor between distances from the creek ($P=0.403$).

Species richness

Mean species richness did significantly differ between distances from the creek overall ($F_{2,27}=3.641$; $P=0.042$), but non-significantly between individual distance comparisons. There was no significant mean difference in species richness between restored reaches ($P=0.531$).

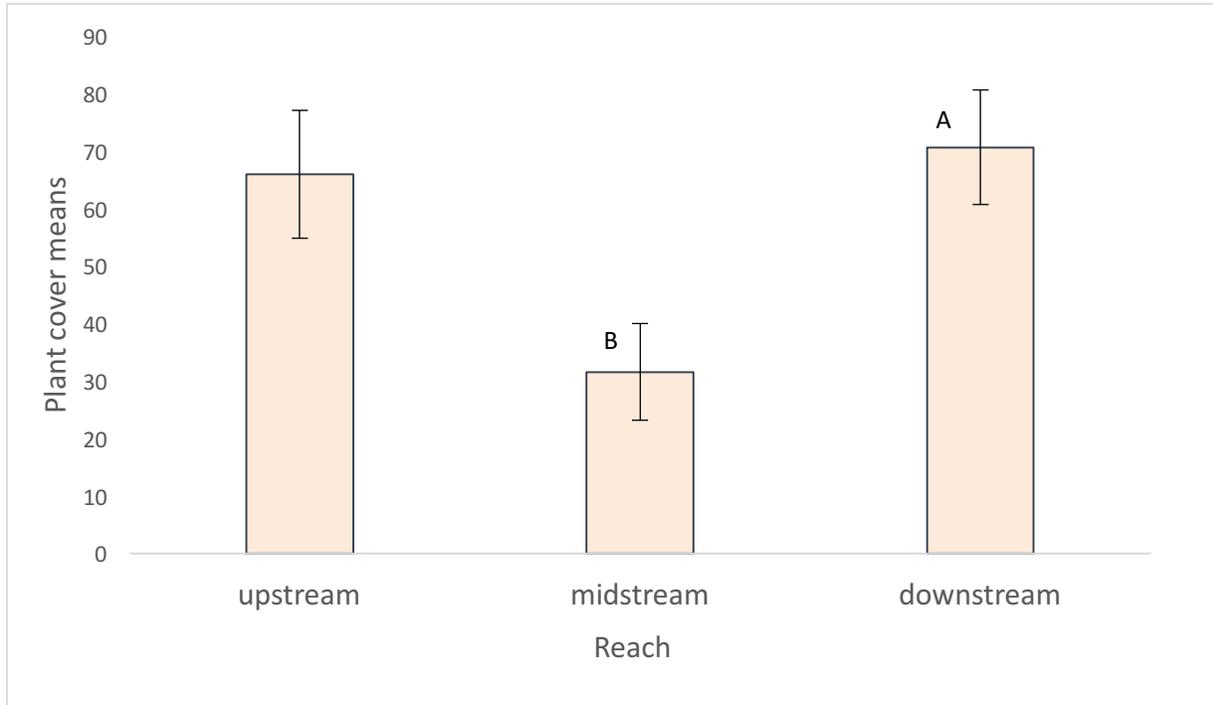


Figure 7. Comparison between reaches of mean cover of exotic species in restored plots at Shades Creek. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Comparisons between plant type

Total native; exotic covers in restored plots

There was greater mean cover of exotic species than total native species ($t_{26}=4.729$; $P=0.000$) (Fig. 8). Across reaches, there was greater mean cover of exotic species than total native species upstream ($t_8=4.424$; $P=0.002$) and downstream ($t_8=4.229$; $P=0.003$). Mean differences between plant type could not be compared between distances from the creek in restored plots due to equal variance.

Native, planted; native, recruited; exotic covers in restored plots

There was greater mean cover of exotic species than native, planted species ($t_{26}=7.278$; $P=0.000$), and native, recruited species ($t_{26}=5.583$; $P=0.000$) (Fig. 8). There was greater mean cover of native, recruited species than native, planted species ($t_{26}=2.770$; $P=0.010$) (Fig. 8).

Total native; exotic covers in control plots

In control plots, there was no significant mean difference between total native species cover, and exotic species cover ($P=0.167$).

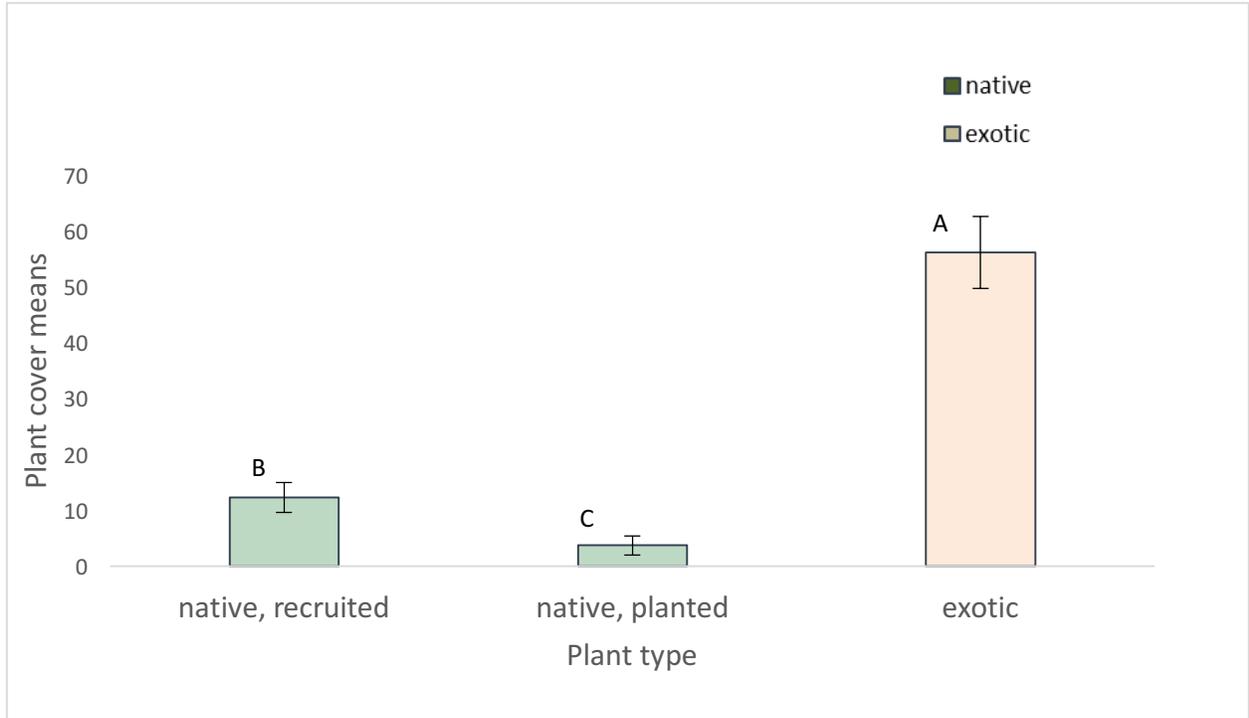


Figure 8. Comparison between reach of plant type mean cover in restored plots at Shades Creek. Letters indicate significant mean differences ($P < 0.05$). Error bars indicate standard error.

Table 3. Mean Shannon Diversity indices by reach and by distance from the river for plots at Cahaba River in Trussville. * indicates non-mean, singular values in which only one data sample was taken due to limited control plots.

| | | Distance from river | | | | | |
|-------|--------------------|---------------------|------------------|--------------------|------------------|--------------------|------------------|
| | | 1 m | | 4 m | | 7 m | |
| Reach | | Value | Std. Error | Value | Std. Error | Value | Std. Error |
| | Upstream control | *1.308716596898052 | . | *1.268870562501131 | . | *1.638546485425983 | . |
| | Upstream restored | 1.540086354769840 | .373614242085835 | 1.450585469433761 | .463081641052791 | 1.416532188998199 | .494642550011202 |
| | Midstream restored | 2.091402845974857 | .309288142492932 | 2.092886760124605 | .177096250088089 | 1.722486853734882 | .200794717156204 |

| | | | | | | |
|---------------------|---------------------------|----------------------|------------------------|----------------------|-----------------------|----------------------|
| Downstream restored | 1.7087 705457 85332 | .31014859 1874815 | 1.47230947 1326090 | .52821791 1350050 | 1.61049566 4238707 | .60664444 0968806 |
| Downstream control | *.8777 862958 80256 | . | *1.5136966 68136542 | . | *.67342669 0468197 | . |

Table 4. Mean Shannon Diversity indices by reach and by distance from the river for plots at Shades Creek in Homewood, Birmingham. * indicates non-mean, singular values in which only one data sample was taken due to limited control plots.

| | | Distance from river | | | | | |
|-------|---------------------|----------------------------|-----------------------|------------------------|----------------------|------------------------|----------------------|
| | | 1 m | | 4 m | | 7 m | |
| Reach | | Value | Std. Error | Value | Std. Error | Value | Std. Error |
| | Upstream control | *1.120 889522 198180 | . | *.80751576 9887591 | . | *.49254847 6084021 | . |
| | Upstream restored | .57770 460017 3731 | .73039373 1662313 | .893722770 455569 | .42951392 2729058 | .873778476 033315 | .24834720 7751350 |
| | Midstream restored | 1.1558 478979 46269 | .76633928 2167799 | 1.00502137 1005927 | .29413307 9576690 | .841627005 217759 | .33000805 6834396 |
| | Downstream restored | 1.3193 728750 17543 | 1.0802188 64325208 | .967320996 211661 | .64862095 8793136 | 1.16490782 5722815 | .51309171 5267505 |
| | Downstream control | *1.906 128058 555701 | . | *1.3087165 96898052 | . | *1.2279674 71740511 | . |

Discussion

In all, the results showed that there were distinct differences in vegetation type and quantity between the restored and non-restored riverbank, as well as within each of these plot types. Some initial expectations were supported by results, while other hypotheses were null, or contrary to results. These results are representative of the restoration treatments at each site, time elapsed since restoration initiation, and the variability in riverbank characteristics and environmental influences.

As expected, non-restored areas had more native than exotic plant species at both sites (Figure 3, Figure 8). Unexpectedly, this same trend was true in restored plots at the Cahaba River site (Figure 2), with greater mean cover of native plants, than exotic plants in downstream plots. This result was presumably due to the short expanse of time since reintroduction of native species one year prior. Comparatively, there were still more exotic species in the restoration zone, than in non-restored areas (Figure 4). These results appear to represent successional

processes after a disturbance, which may be a negative effect of this type of active restoration on the plant community (Beater et al., 2008; Bellingham et al., 2005; Petraitis et al., 1989).

There was a greater abundance of native self-recruited plants than exotic plants at the Cahaba River, and also Shades Creek, restored bank (Figure 2; Figure 8 respectively). This may be due to plant seeds remaining after vegetation removal during restoration, or the seeds may have dispersed to the site after the replanting. The higher abundance of native plant species that self-recruited to the site may provide valuable insight as to which native plants are first successional, and are able to establish under pressure from invasive species (Beater et al., 2008).

The higher cover of total vegetation in restored than non-restored areas represents the presence of older, established trees observed in the upstream control area, and their effect on community productivity. Trees shade the ground, occupy root space, and drop leaf litter, all of which might contribute to lack of understory growth (Belsky, 1994). This study was not informed on the condition of the riverbank prior to restoration, but results and observations suggest that there were well-established trees and a higher abundance of native species than exotic species before the restoration. Under these assumptions, the plant community may have been unproductive with low diversity and low species richness (Pollock et al., 1998). Short-term monitoring of restoration initiatives can be misleading, as ecosystems can take years for full recovery (Menges, 2008). The study done at the Cahaba River is the first one presently known, and may be helpful to future monitoring of ecosystem recovery and in continued post-restoration management.

Initial expectations for both sites were based on other restoration projects of riparian riverbanks, where lack of continued maintenance over time resulted in a regression of native species and an increase in exotic species that were often invasive (Beater et al., 2008). In conjunction, exotic plant species appeared to have eventually outcompeted native plant species for space and resources at the Shades Creek site (Figure 8). Invasive species can be quick to invade and populate available substrate immediately following a disturbance to the plant community (Beater et al., 2008).

Prior analyses of Shades Creek restoration used different methods and data analysis than this present study, and accurate comparisons could not thus be made in vegetation changes over the years since restoration. The 2012 study does suggest that there was a high abundance of exotic species in the restoration zone one-year post-restoration, as well as low survivorship of native planted species. Invasive species prone to this restoration site are *Pueraria montana* and *Humulus japonicus* vines, which are characteristically aggressive and will require extensive maintenance removal to allow native, riparian vegetation to recover. Invasive plants are known to be aggressive in introduction to a community where they compete with native plants for space, nutrients, sunlight, and water. Research shows that invasive species can take root quickly after a disturbance to a plant community, and stressed native species will often not be able to compete (Godefroid et al., 2011; Richardson et al., 2007).

At the Shades Creek site, there was greater total vegetation cover in the restoration zone than the non-restored areas (Figure 6), which contained adult trees that shaded, and covered the ground in leaf litter, in similarity to the Cahaba River site. This difference was highest between downstream plots, and plots one-meter distance from creek, in restored and non-restored areas. Indicative of these results, the downstream non-restored area was observationally noted as severely eroded, with well-established trees and little understory vegetation.

Variability in waterway characteristics, and local environmental factors will be influential and individualized for any site of riparian restoration. Broad generalizations may be made for

restoration reference in choosing where to plant native species based on these factors, so as to achieve the highest rates of survivorship and recruitment possible. Plant species chosen for replanting in degraded riparian habitat should be those better adapted to withstand common disturbances like floods, and sediment deposition, and can also remain resilient to physical damage, and aggressive invasive species (Naiman et al., 1998; Richardson et al., 2007).

Waterways are well-known as top seed dispersal pathways (Richardson et al., 2007), so it was expected that plant type and cover would differ with increasing distance from the water source. Some results weakly suggest that at the furthest distance from the water source, native plants were more abundant than exotics, which may indicate higher dispersal of invasive species at the toe of the creek. The rock vanes installed at each site may have helped or hindered dispersal to the riverbanks, by altering where seeds, water, and soil were deposited along the bank (Godefroid et al., 2011). Severe erosion, and those potential effects from rock vanes, may be the cause for a greater abundance of exotic species downstream, than midstream in Shades Creek restored plots (Figure 7). Additionally, at the Cahaba River, there was apparent low survivorship of native planted species between the midstream reach, and reaches at the restoration site's edges (Figure 5), which may indicate how plant species are affected at the edges of the restored zone.

Comparisons in species richness were weakly significant, but represent the high abundance of species at each site, and indicate the plant community productivity (Pollock et al., 1998). Untested Shannon Diversity Indices ranged from low to high at each site, though they tended to be higher at the Cahaba River site, and in the midstream section (Table 3), indicating higher diversity here among plant populations.

Severely degraded riparian habitats may be unable to recover despite restoration efforts, particularly if habitat characteristics necessary for native plants to survive have been significantly altered, or removed (Holmes et al., 2005). Disturbances to habitats are unavoidable, but can be used advantageously in disrupting negative plant community interactions. Passive restoration techniques help to regulate effects from a disturbance, allowing for more control in deciding which ecosystem dynamics are disrupted, and which can benefit (Suding et al., 2004).

Our study used data collection methods and data analyses common in other studies of plant community dynamics, and restoration. However, future studies might consider collecting data on plant abundance by measuring frequency of plant type, which can be more precise than ocular estimation that's done from oblique angles. Analyzing data by converting cover to frequency was not ideal, though there appeared to be a significance in frequency among native and exotic species between plot types at Shades Creek. It was observed at this site that plant aerial cover was greater than plant ground cover, meaning there was more available substrate than indicated by aerial cover values. At a plot in Shades Creek, it was estimated that between 75 and 95 percent of aerial vegetation cover was of invasive plant *Humulus japonicus*. Underneath its dense plant canopy, only about 1-5 percent of bare ground was occupied by plant growth, likely due to lack of sunlight and rainwater reaching the ground. Invasive species can cause erosion control this way, by blocking native plants from rooting in the soil, and causing an unstable bank underneath (Cowling et al., 1976). This difference was not found at the Cahaba River site, where these types of invasive species were not yet found, and the riverbank was still stabilized from restoration treatments.

Both of these restoration projects were implemented to enhance the local, riparian plant community and its accompanying aquatic ecosystem. The strongest factor for initiating restoration was to decrease riverbank erosion threatening land loss. Each restoration project

appeared to use the same active restoration techniques, whereby all vegetation was removed from the degraded riverbank, native species were reintroduced, and rock vanes were installed in the waterway to further curb erosion. Both sites had higher exotic species values in the restoration zone than in the non-restored plots, demonstrating successional and long-term processes that would occur after a large disturbance to the plant community (Bellingham et al., 2005). Active restoration has been described as a disturbance by some studies (Bellingham et al., 2005; Holmes et al., 2005), and its methods have produced mixed success in riparian restoration over short, and long-term monitoring (Beater et al., 2008; Richardson et al., 2007; Ruwanza et al., 2013; Suding et al., 2004).

Studies have worked to assess passive and active restoration techniques, and now more attention is turning to whether restoration should be attempted at all, particularly if considering the degree to which an ecosystem is degraded. Many restoration projects do not continue with post-restoration maintenance either, which is often fundamental to achieving some level of determined restoration success (Beater et al., 2008; Godefroid et al., 2011; Kirchner et al., 1998; Seavy et al., 2009). By not attempting to restore severely degraded ecosystems, this same time, cost, and effort could go into preventing ecosystem degradation, by protecting threatened native, riparian habitat.

It has been suggested that each site of restoration should have its own “prescription”, or a specific restoration plan that will work to meet its most prominent ecological needs, and work-around the riverbanks’ specific environmental and physical characteristics (Sweeney et al., 2004). In this study, the Cahaba River and Shades Creek sites markedly differed in length, width, surrounding development, and waterway structure and direction, which are all abiotic variables that can alter plant community dynamics (Suding et al., 2004). It appears unlikely that one restoration method could work for any variety of degraded ecosystems.

The results from this study do not suggest a failure in riparian riverbank restoration at either site, being due to that more time elapsed since the onset of restoration will allow the restored habitat, and thereby the ecosystem later, to self-regulate (Menges, 2008). Observations and results from this study should add to the growing number of analyses assessing the best methods possible for restoring degraded ecosystem function to be ecologically and economically beneficial again. Research in this area of study needs to continue extensively analyzing each part and goal of restoration initiatives (Wohl et al., 2005). Studies should look at how riparian plant communities change within and by surrounding terrestrial and aquatic ecosystems, and under what influencing abiotic factors as well. Finally, the condition of habitats prior to restoration, and afterwards, should be thoroughly assessed, to evaluate whether restoration was critically needed, and whether the paramount goal was eventually met. It remains important in our fluctuating climate to continue studying human impacts on riparian ecosystems, which are profoundly viable to healthy waterways, and productive native habitats (Seavy et al., 2009).

Bibliography

- Allan, E., Weisser, W., Weigelt, A., Roscher, C., Fischer, M., & Hillebrand, H. (2011, September 26). More diverse plant communities have higher functioning over time due to turnover in complementary dominant species. *Proceedings of the National Academy of Sciences*, 108(41), 17034-17039. doi:10.1073/pnas.1104015108
- Bailey, B., & Brown, M. (2012). Species Survival in Plant Community Restoration, Jefferson County, AL. Samford University Thesis.
- Beater, M., Garner, R., Witkowski, E. (2008, July). Impacts of clearing invasive alien plants from 1995 to 2005 on vegetation structure, invasion intensity and ground cover in a temperate to subtropical riparian ecosystem. *South African Journal of Botany*, 74(3), 495-507. Retrieved from doi: 10.1016/j.sajb.2008.01.174
- Bellingham, P. J., Peltzer, D. A., & Walker, L. R. (2005). Contrasting impacts of a native and an invasive exotic shrub on flood-plain succession. *Journal of Vegetation Science*, 16(1), 135. doi:10.1658/1100-9233(2005)016[0135:cioana]2.0.co;2
- Belsky, A. J. (1994). Influences of Trees on Savanna Productivity: Tests of Shade, Nutrients, and Tree-Grass Competition. *Ecology*, 75(4), 922-932. doi:10.2307/1939416
- Cowling, R.M., Moll, E., Campbell, B.M., 1976. The ecological status of the understorey communities of pine forests in Table Mountain. *South African Forestry Journal* 99, 13–23. doi: 10.1080/00382167.1976.9630104
- Flemming, J. (2014). Jefferson County, AL. Samford University Thesis.
- Godefroid, S., Piazza, C., Rossi, G., Buord, S., Stevens, A., Agurauja, R., Vanderborcht, T. (2011). How successful are plant species reintroductions? *Biological Conservation*, 144(2), 672-682. doi:10.1016/j.biocon.2010.10.003
- Gregory, S. V., Swanson, F. J., Mckee, W. A., & Cummins, K. W. (1991, November 8). An Ecosystem Perspective of Riparian Zones. *BioScience*, 41(8), 540-551. doi:10.2307/1311607
- Holmes, P., Esler, K., Richardson, D., & Witkowski, E. (2008). Guidelines for improved management of riparian zones invaded by alien plants in South Africa. *South African Journal of Botany*, 74(3), 538-552. doi:10.1016/j.sajb.2008.01.182
- Holmes, P.M., D.M. Richardson, K.J. Esler, E.T.F. Witkowski, S. Fourie. (2005). A decision-making framework for restoring riparian zones degraded by invasive alien plants in South Africa. *South African Journal of Science*, 101, pp. 553–564
- Johansson, M. E., Nilsson, C., & Nilsson, E. (1996, August). Do rivers function as corridors for plant dispersal? *Journal of Vegetation Science*, 7(4), 593-598. doi:10.2307/3236309
- Kirchner, J. W., Micheli, L., & Farrington, J. D. (1998, Aug.). Effects of herbaceous riparian vegetation on streambank stability. *University of California Water Resources Center*. UC Berkeley: University of California Water Resources Center. Retrieved from <https://escholarship.org/uc/item/70q7678r>
- Maunder, M. (1992, March). Plant reintroductions: An overview. *Biodiversity and Conservation*, 1, 51-61. doi:10.1007/BF00700250
- Menges, E. S. (2008). Turner Review, No. 16. Restoration demography and genetics of plants: When is a translocation successful? *Australian Journal of Botany*, 56(3), 187-196. doi:10.1071/bt07173
- Naiman, R. J., & Bilby, R. E. (1998). River Ecology and Management in the Pacific Coastal Ecoregion. *River Ecology and Management*, 1-10. doi:10.1007/978-1-4612-1652-0_1
- Osawa, T., H. Mitsuhashi, and H. Niwa. (2013). Many alien invasive plants disperse against the

- direction of stream flow in riparian areas. *Ecological Complexity*. 15: 26-32.
<http://dx.doi.org/10.1016/j.ecocom.2013.01.009>
- Petraitis, P.S., Latham, R.E., Niesenbaum, R.A., 1989. The maintenance of species diversity. *Quarterly Review of Biology* 64, 393–417.
- Pickett S.T.A., White P.S. (1985). Natural disturbance and patch dynamics: an introduction. In: Pickett S.T.A., White P.S. (eds). *The ecology of natural disturbance and patch dynamics*. Academic Press, Orlando, 3–13.
- Pollock, M. M., Naiman, R. J., & Hanley, T. A. (1998, January). Plant Species Richness in Riparian Wetlands--A Test of Biodiversity Theory. *Ecology; Ecological Society of America*, 79(1), 94-105. doi:10.2307/176867
- Primack, R., & Drayton, B. (1997, October). The experimental ecology of reintroduction. *Plant Talk*, 97, 25-28.
- Richardson, D.M., P.M. Holmes, K.J. Esler, S.M. Galatowitsch, J.C. Stromberg, S.P. Kirkman, P. Pyšek, R.J. Hobbs. (2007). Riparian vegetation: degradation, alien plant invasions, and restoration prospects. *Diversity and Distributions*, 13, pp. 126–139
- Ruwana, S., Gaertner, M., Esler, K., & Richardson, D. (2013, September). The effectiveness of active and passive restoration on recovery of indigenous vegetation in riparian zones in the Western Cape, South Africa: A preliminary assessment. *South African Journal of Botany*, 88, 132-141. doi:10.1016/j.sajb.2013.06.022
- Seavy, N. E., Gardali, T., Golet, G. H., Griggs, F. T., Howell, C. A., Kelsey, R. Weigand, J. F. (2009). Why Climate Change Makes Riparian Restoration More Important than Ever: Recommendations for Practice and Research. *Ecological Restoration*, 27(3), 330-338. doi:10.3368/er.27.3.330
- Suding, K. N., Gross, K. L., & Houseman, G. R. (2004). Alternative states and positive feedbacks in restoration ecology. *Trends in Ecology & Evolution*, 19(1), 46-53. doi:10.1016/j.tree.2003.10.005
- Sweeney, B. W., & Czapka, S. J. (2004). Riparian forest restoration: Why each site needs an ecological prescription. *Forest Ecology and Management*, 192(2-3), 361-373. doi:10.1016/j.foreco.2004.02.005
- Thomas, E., Jalonen, R., Loo, J., Boshier, D., Gallo, L., Cavers, S., Bozzano, M. (2014, December 1). Genetic considerations in ecosystem restoration using native tree species. *Forest Ecology and Management*, 333, 66-75. doi:10.1016/j.foreco.2014.07.015
- Wohl, E., Angermeier, P. L., Bledsoe, B., Kondolf, G. M., Macdonnell, L., Merritt, D. M., Tarboton, D. (2005, October 4). River restoration. *Water Resources Research*, 41(10).doi:10.1029/2005wr003985