The Proceedings of GREAT Day 2012

SUNY Geneseo

ACKNOWLEDGMENTS

The State University of New York at Geneseo

The Proceedings of the Sixth Annual GREAT Day

Geneseo Recognizing Excellence, Achievement & Talent is a college-wide symposium celebrating the creative and scholarly endeavors of our students. In addition to recognizing the achievements of our students, the purpose of GREAT Day is to help foster academic excellence, encourage professional development, and build connections within the community.

The GREAT Day Website: http://great.geneseo.edu

Table of Contents

A Philosophical Dialogue Over Our Shared Reality

Matthew McNeill

I anxiously hurried my way towards Welles 131, as I was late for the very meeting I had worked so hard to organize. Today's date had been circled on my calendar for over two months in advance of this event, and in my excitement I couldn't fall asleep. I finally passed out at around 6 a.m. and subsequently slept through my alarm going off at 9 a.m. As I shuffled my way to the classroom, both excited and scared, it was obvious the impending discussion could drastically change my agnostic outlook on the reality I currently found myself in. When I finally reached the door and thrust it open, I stood stupefied in the doorway as the faces of Carlo Filice, Muhammad, and Shankara stared back at me.

"It's about time you got here. I was beginning to run out of funny jokes, and I've got a million of them," Filice said slyly.

"That's what you think. Thank Allah we can finally get started, I was starting to lose it after being subjected to such humorless material," added a frustrated Muhammad.

I collected myself as best I could and addressed the matter at hand. "Thanks again everyone for coming. I have called you three here today, as you know, for a discussion on what the true purpose of life is. As a young student who currently struggles to make sense of the world around me, I have adhered to a form of loose agnosticism to keep me sane and grounded. Having said this, I was hoping to glean something from this discussion today. Seeing as how you gentlemen are some of the most knowledgeable individuals available on Eastern Philosophy, I am eager to hear what you have to say on this topic."

"I'll kick things off," Muhammad said confidently. "You see, Allah created the Earth and everything on it. I know because He told me all this himself. There is a clear distinction between 'creator' and 'creature'. We are all unique individuals who have our own soul, and how we act on Earth determines whether or not our soul earns the right to live with Him in paradise. For those who fail to submit to His will, it is said that God has 'set a seal upon their hearts and ears; their sight is dimmed and grievous punishment

awaits them' (Dawood 11). To avoid this fate, certain specific rituals must be followed…"

"I'm not so sure," interrupted Shankara. "I would contend that the world is merely Maya, a dream-like illusion that is our ordinary perspective. Afterall, 'our senses may deceive us; our memory may be an illusion… The objects of knowledge may be open to doubt, but the doubter himself cannot be doubted' (Radhakrishnan and Moore 506). The doubter within each of us, therefore, is one with Brahman, the one universal, supreme spirit who is without qualities or attributes. Duality is thus a false perspective, with no separation between creator and creature."

"You're not far off, Shankara," interjected Filice. "I agree with you that our experiences take place within an imagined world dreamed up by a 'super-deity' of sorts. The idea of an overly personalistic and jealous God has always seemed like a silly notion to me…"

Suddenly, before Filice could finish his argument, the clouds outside parted and a ray of sunlight shone through the windows of Welles 131. It made its way across the room until it nearly blinded Muhammad. A trance came over him and he began talking in a voice that was not his own.

"How dare you mere mortals question me?" bellowed Allah in a clearly agitated tone. "I am not a jealous God! I don't care that I wasn't invited to this little meeting… but at least have a little decency and stop acting like I don't exist. Jeez."

"Get out of here, Allah. You weren't invited for a reason," I replied. "I knew you wouldn't be able to handle the criticism without putting up a fuss. Apparently you didn't even need an invite to do that though. This is the type of behavior Filice was referring to. How can we take you seriously if you act more like a child than a supreme being?"

Allah let out a noticeable grunt but perhaps came to terms with what I had just said. He quickly departed, returning the skies to their cloudy state. He made sure, however, to let his unhappiness be known to us all with a hard rainfall.

"Yeah, uh, sorry about that. Hopefully we don't get another Great Flood right? Uh…where were we?" Muhammad offered uneasily.

"Let's get back to our discussion on the purpose of life," I added hopefully.

"Ah yes, well as I mentioned earlier, a Muslim should submit to God's will in hopes of gaining entrance into Paradise. To be reunited with the heavenly Father is the ultimate goal in life," answered Muhammad. "A Muslim woman specifically should be obedient to her husband and be veiled in public because 'God has made the one superior to the other' (Dawood 64). Violence is excusable if necessary in defending the will of God, whether it be beating a disobedient wife or taking control of Mecca through military means. A Muslim must also pray five times a day facing Mecca, fast at certain times of the year, make a pilgrimage to Mecca, and engage in almsgiving."

"I agree that reaching the creator should be the chief goal in an individual's life," stated Shankara. "An individual must do his duty according to what caste he finds himself in and try to better himself in his current life so as to move up in society through reincarnation. In this way he can escape the endless cycle of rebirth and finally be reunited with Brahman outside the dream-like Maya world. Similar to your notions of 'paradise' and 'hell,' karmic actions determine whether you move up or down this hierarchy of lives. I have serious issues with many of your practices though. There seems to be very little attention paid to the well being of animals, and here we differ greatly, as they are spiritual creatures just like humans that demand similar respect. Our women, unlike your own, hold a high role in society, as females 'must be honored and adorned by their fathers, brothers, husbands, and brothers-in-law, who desire their own welfare' (Radhakrishnan & Moore 189). Furthermore, the…"

"Hold on a second," interjected Filice. "The Hindu view of women is not quite as rosy as you make it out to be, since they can do nothing independently, even within their own houses, and likewise must be 'kept in dependence by the males of their families' (Radhakrishnan & Moore, 190). I do agree with your notion of multiple lives, however. I believe 'we are here to learn to appreciate, express, and master, a variety of values" (Filice 133). These values manifest themselves in, and are experienced through, humans and animals who act as avatars of a sort for the super-deity in his dream world."

"Avatars?! Are you crazy?" an incredulous Muhammad belted out. "Shankara, can you believe this guy?"

"No problem with avatars here," Shankara replied to an angry and disappointed Muhammad.

"For me, both the full submission to Allah for Muslims and the dreary affair of trying to escape from a sufferable reality for Hindus make little sense to me. Why would a God create a world full of toil or pointlessness?" asked Filice.

"That seems like a very legitimate question," I replied, noticing that both Muhammad and Shankara had become virtually silent. "The Arabic world was once at the forefront of achievements in such fields as mathematics. The Muslims' lack of emphasis on human experience in this world has greatly diminished personal development in some sense."

"I believe," continued Filice, "that the sexes are equally sacred as they house within them a piece of the super-deity I mentioned before. Animals are sacred for the same reason, something I would agree with Shankara on. The caste system, however, needs to go. It can arguably be seen as contradictory and holds people back in terms…"

At that moment another uninvited guest glided into the classroom. A chariot being driven by what looked to be a prince crashed through the window. It was Krishna, Vishnu's avatar, who appeared to be very unsupportive of the current discussion.

"Hello Krishna," said Shankara in a welcoming greeting.

"Who the heck is this guy?" asked Muhammad.

"It's Krishna, avatar of the god Vishnu," Shankara answered matter-of-factly. "He's yet another part of Brahman. Have you been paying attention to any of our discussion here today?"

"Acceptance of avatars? Worshiping of a multitude of gods? Or are they one God? I'm at a loss for words…" replied a defeated Muhammad.

"The caste system has been around since ancient times," explained Krishna. "It has persevered into modern times. It gives everyone a job to do and coincides nicely with reincarnation, in that you get to move upwards before finally reaching enlightenment at the top of the social pyramid. What's not to like?"

"Well for one," argued Filice, "it's not fair for a system to be in place where people are kept locked in place. All humans, as sacred beings and part of a larger deity, should be shown the same amount of respect and given equal chances of advancement regardless of where they start from in society. Adopting 'some form of the mixed economic system is the way to respect everyone when it comes to issues of wealth and

poverty,' and implementing 'some form of democracy will treat everyone with respect, by giving everyone a voice in governing'" (Filice 93).

"Excuse me, but you weren't invited Krishna," I said. "You and Allah should hang out, I have it on good authority he's having a lousy day too."

"No time unfortunately," replied Krishna, who began to float away on the chariot. "There's some guy who refuses to fight for his kingdom despite being a great general. His name is Arjuna, I believe. I'm his charioteer and now, it appears, also his motivational speaker. I've obviously already overstayed my welcome anyway."

"Isn't it contradictory to tell someone to quell their desires and lecture about the immorality of killing, only to force a young warrior to kill his family in return for a kingdom he no longer desires because he must 'do his duty'?" I asked in a confident tone. "Arjuna seems to be in quite an undesirable and confusing predicament." Krishna had no answer as he continued to drift away and was soon long gone. I think even he realized there might not be a satisfactory answer to my question.

"As much as it pains me to say this, I see and understand where you are coming from. I refuse to accept it, but I can tolerate it," stated a now much calmer Muhammad.

"I agree with much of what you have said as well, and I feel many of our viewpoints are the same," added Shankara. "I will, however, point out that the widespread implementation of a mixed market economy will produce byproducts like pollution that can and will harm both animals and nature itself."

"That," agreed Filice, "is perhaps the biggest flaw in my personal rationalization of things, and it is not a simple problem to solve."

"Even so," I added, "you have coherently and compellingly explained what you believe to be the purpose of the life experience, with few glaring potholes. That is better than most can claim. No offense intended, Muhammad or Shankara."

"It's fine," replied Muhammad.

"I'm not offended," added Shankara tiredly.

"Well, I think this discussion has come to its conclusion," I said, feeling even more worn out than Shankara. "Thank you so much guys for agreeing to do this for me. I appreciate it and I feel that I've come away with a lot from this experience."

We all exchanged pleasantries, and Welles 131 was soon empty once more. As I slowly made my way back to my apartment, I reflected on what I had just experienced. While I agreed most with what Filice had to say on the purpose of life, it was clear that he took some key ideas from Shankara before him while also having the benefit of a much more modern time period where ancient scripture does not dominate and constrict like it once did. Shankara was in a similar situation, with a more modern outlook on things compared to Muhammad, whose religion is based on one of the oldest monotheistic religions in human history.

What Filice posits is somewhat revolutionary and much more pleasing in coherent terms than other, usually more dated answers to this question. While I believe what he offers to fellow philosophers is very possible, the agnostic within me is still, naturally, somewhat noncommittal and skeptical. Before this discussion, I didn't believe that we could know for sure exactly what the dynamic is that swirls around our reality on a daily basis, and I still largely feel the same way. I would, however, fully agree with Filice's point that this life should be about having fun and trying to get as much value out of it as possible while we are here. As long as you can do these things, all the other specifics of our existence and purpose in this life can be pushed to the background.

Works Cited

Dawood, N. J. *The Koran*. London: The Penguin Group, 2003. Print.

Radhakrishnan, Sarvepalli, and Charles A. Moore. *A Sourcebook in Indian Philosophy*. New Jersey: Princeton University Press, 1957. Print.

Filice, Carlo. *The Purpose of Life*. Maryland: University Press of America, 2011. Print.

Corruption and Conversion: The Artist and the Dandy in *Brideshead Revisited*

Cailin Kowalewski

In *Brideshead Revisited*, Evelyn Waugh presents the dandy as an explorer of fulfillment through the experience of transient beauty. According to Christopher Lane's article *The Drama of the Impostor: Dandyism and Its Double*, a dandy's complete sense of self relies on a comprehension of the beautiful that is derived from instantly dissolving impressions, creating a search for the fleeting, the impressionable, and the alarming. To reach "art" and beauty in the form of experience, the dandy must navigate the mundane world of cultural norms that surrounds him so that he can subvert them, while remaining necessarily separated from them. He isolates himself from the commonplace, insufficient habits of culture, and seeks out the unique.

Crucial to this process is the preservation of his sense of creative autonomy. To do this, the dandy establishes degrees of separation from reality and creates his own sub-spheres: lenses of behavior through which he can interpret social codes. In *Brideshead*, Waugh's primary dandy figure, Sebastian, struggles in a constant effort to dissociate himself from reality, whether physically or mentally. However, the dandy's establishment of *sur*realities is directly challenged by the existence of the artist, who attempts to "'fix' each occasion or 'scene'" of the art that is the dandy's behavior "before it passe[s] without recognition or judgment" (Lane 34); and by what John Edward Hardy calls "the convert-artist" (167), who moves away from secular art in service of itself, and towards art that is governed by its expression of religious traits. In this context, Waugh presents the threat to dandy figures as a re-contextualization; a re-imposition of cultural ennui that perverts the immediacy of the dandy's "life" art, that destroys the levels of mediation protecting the dandy from mainstream contextualization. Both Waugh and his protagonist Charles Ryder do this by trying to make Sebastian's life tangible and palpable to others, imposing reason and morality on chaotic beauty. Through this, Waugh exposes an interesting relationship between the passionately disengaged and the culturally captive.

Sebastian Flyte's identity as a dandy is easily identified as a search for a kind of fleeting artistic realization, uninhibited by the external "truths" of the society he principally rejects. In *The Picturesque Prison*, Jeffrey Heath characterizes Waugh's dandy, Sebastian Flyte, as

"resolutely adolescent, and his intense desire for solitude leads him into self-love… he refuses responsibility and conformity" (Heath 177). However, Sebastian's behaviors do not necessarily reflect an urge to be alone, but rather "let alone"—that is, free to pursue happiness through art and its interpretation from a lens of the surreal. In his discussion of the dandy aesthetic, Christopher Lane stresses this importance of interpretation in art so that it is not merely seen, but "realized," so that the dandy himself embodies transient artistic beauty via internally significant "impressions" (Lane 31). These transitory images, as Lane conveys,

> …can be only fleetingly sustained before they "melt under our feet" (R 189), leaving no trace of their demise other than the demand for a repetition and intensification of their effect... This paradox tied the subject to an impossible urge for self-completion and an increasing dependence on the image to define—but not arrest—its limit and vanishing-point (33).

With this perspective in mind, Sebastian Flyte's identity as a dandy is easily identified as a search for a kind of fleeting artistic realization, uninhibited by the external "truths" of the society he principally rejects.

To separate himself from social realities, Sebastian immerses himself in the unusual and the startling. This practice "allows the dandy to irritate, shock, and deride, yet at the same time to captivate, seduce, and thus dominate the society he rejects" (Rossbach 277). Waugh primarily displays this in Sebastian's regression into immaturity and alcoholism, his isolation in the haven of Brideshead estate, and his relationship with the spiritual elements of Catholicism. From the first occasion that he is presented, Sebastian deliberately removes himself from the mature world. He refers to his mother as "Mummy," claims a sole affection for his old "Nanny Hawkins," and feigns a terrible helplessness, evident in his telegram to Charles over a sprained ankle relating, "gravely injured" (74). Sebastian's behavior effectively separates him from his peers at his university as an outsider and an oddity. Charles notes, "he was the most conspicuous man of his year by reason of… his eccentricities of behavior which seemed to know no bounds… I was struck less by his looks than by the fact that he was carrying a large Teddy-bear" (28). Sebastian creates a realm for himself as a child in a world of adults: an immaturity only increasingly surreal when continued on the grounds of Brideshead estate.

Brideshead itself is posited as a detached oasis, a locus Waugh creates for the embodiment of past aristocratic glory that is isolated, protected, hidden. It is a "new and secret landscape" (37), with darkened halls, shuttered windows, gilt mirrors, and sheeted furniture (38).

The building's very stones are deceptive, taken and relocated from a castle that had previously borne the name. Its eclectic, lavish, and ornate style captivates Charles, who says, "This was my conversion to the baroque. Here under that high and insolent dome, under those tricky ceilings…I felt a whole new system of nerves alive within me, as though the water that spurted and bubbled among [the fountain's] stones was indeed a life-giving spring" (82). In this oasis, Sebastian and Charles are immersed in their own illusions, and the artistic grandeur that is Brideshead facilitates Sebastian's efforts to abandon the conventional. He says, "If it could only be like this always—always summer, always alone, the fruit always ripe" (79), and upon their first meeting, Sebastian is drunk and looks at Charles with "unseeing eyes" (29). By bringing Charles away from their life at Oxford, it is clear that he means to include him in his mirage of indulgence and youth— to drink wine, eat strawberries, and smoke Turkish cigarettes under the shade of an elm tree (23). Already, he seems to exist in more of a fairy world than that of a collegiate, and thrives at Brideshead where he is able to easily uphold his own fantasies.

However, it is in this realm that Waugh presents the dandy's ultimate undoing, and in a manner that argues that the dandy's embodiment of art through lifestyle is ultimately overpowered by the logic-governed culture it reinterprets. While it is rather brash foreshadowing by Waugh, the placement of a skull in Charles's room with the inscription "Et in Arcadia Ego" ("Even in Arcadia I am there") does not just foreshadow Sebastian's decline in the Arcadian world of Brideshead, but actually parallels the fact that Charles brings this destruction with him. Waugh introduces Charles as an unwittingly antagonistic force of normalcy, corrupting both Sebastian's "religion" and "art."

To understand the effect Charles has on Sebastian, it is useful to analyze Waugh's construction of the pair's religious beliefs and practices. Brideshead estate is essentially dichotomized by two different perspectives on Catholicism; a conflict that Heath describes as "between the will of man and the will of God" (161). Perhaps more accurately, we may clarify that this conflict is between the opposing means and definitions of achieving divine grace. Characteristically, Sebastian's concept of Catholicism is grounded in a belief in the unseen. He holds faith in traditions and beliefs not logically proven, and speaking of his family's practices, says, "happiness doesn't seem to have much to do with it, and that's all I want" (89). While Sebastian "always heard his mass," his emphasis on happiness and dandy aesthetic suggest that his fixation is upon the impalpable, mystical qualities of faith and the captivating image it

creates, rather than its moral consequences. Heath eloquently summarizes this as "the operation of grace through the inauthentic" –a search for that "self-completion" through the inspiring beauty of divine mystery (168). At the beginning of Chapter 4, Waugh expresses this through Charles in a discussion of what he calls "the languor of Youth," suggesting, "perhaps the Beatific Vision itself has some remote kinship with this lowly experience; I, at any rate, found myself very near heaven, during those languid days at Brideshead" (79). Sebastian's contrived Arcadian world is clearly paralleled with a state of grace, in which the dandy can nourish his creation of a fleetingly beautiful existence.

However, it is clear that Waugh inherently rejects this unrealistic lifestyle and utilization of religion, drowning Sebastian in his own self-indulgence and misery and proving that art cannot exist for itself or for the sake of beauty and happiness alone. "Waugh emphatically rejects Keats' pagan heaven of art," Heath writes. "In Waugh's view, it is not true that 'a thing of beauty is a joy for ever'; a thing of beauty is a joy only if it is a way-place on the road to God" (174). Without this necessary service of divine morality, Sebastian's exploration of beauty is profane. According to Waugh's portrayal, "Sebastian is pernicious when loved as an end in himself but valuable as an avenue to the divine" (Heath 178). Yet when he is forced into a matrix of religious consequence and made part of its system, he is forced into a common sort of reality and consequently deprived of his creative ability of dissociation.

Thus, it is not so much the existence of this kind of religion that harms Sebastian, but rather the constant imposition of morality and the mundane that shatters his carefully constructed haven of surreal beauty. Waugh primarily develops this concept in contrast to Charles's relationship with religion. Lady Marchmain seeks to convert Charles' alliance via Christianity, and by the end of the novel it seems that this change has been effectively completed in Charles's melodramatic prayer for salvation on behalf of the dying Lord Marchmain. One of the final images of the novel is of "a small red flame… burning anew among the old stones" of the Brideshead chapel (351), a flame apparently lit in Charles's saved soul. But Waugh does not present Charles as a convert-hero, a fact clarified if only in that Charles's prayers are not at all convinced or convinc*ing* ("Oh God, if there is a God, forgive him his sins, if there is such a thing as sin,") and that his more earnest prayer is only for an acknowledgment of ritual: "I suddenly felt the longing for a sign, if only of courtesy, if only for the sake of the woman I loved, who knelt in front of me, praying, I knew, for a sign (338). Charles is not a true Christian, because he

lacks the characteristic that Sebastian, ironically, has: a faith in the unseen. Charles feels a flame of religious passion, but his belief is unavoidably restricted by the "beaten" physicality that grounds him to reality. When Lord Brideshead lies dying, the thought of offering him final sacraments is "nonsense" (335).

In this surprisingly complex portrayal, it is apparent that Charles is *not* the religious hero/protagonist that we may wish him to be. It is tempting to ignore his essential betrayal of Sebastian's profligate lifestyle for his family's resolute theology, but in this shoddy "conversion," Waugh clearly posits him as a threat to both the fragile Sebastian and precariously surviving Brideshead, indicting him as a conduit of the mundane and commonplace perspectives of the outside world.

When they first meet, Charles is as captivated by Sebastian, Brideshead, and the "divine" beauty they share; but as Charles becomes more invested in the dynamics of the Flyte family, less of a participant in Sebastian's way of life, and more of an observer, Waugh reveals the true nature of his protagonist as a corrupter of art. Anthony Blanche introduces our first definition of the artist, saying, "Artists are not exquisite… but the Artist is an eternal type, solid, purposeful, observant— and, beneath it all, p-p-passionate, eh, Charles?" (52). But when Charles, our proclaimed "artist," creates, his works do not seem to be the product of an intense emotional experience or even a deliberate attempt at expression. They are merely recordings of things he sees, lacking originality and creativity. He says of his sketch of the fountain that only "by some odd chance, for the thing was far beyond me, I brought it off and by judicious omissions and some stylish tricks, produced a very passable echo" (81). Blanche laments that Charles' paintings are not the daring, passionate, "unhealthy" creations he had hoped for, but corrupted by "the great English blight" of "charm" (273). Charles' depictions of the exotic are "simple, creamy English charm, playing tigers" (273). His landscapes are correctly romantic, but his attempt at a pastoral scene with living figures is a flop. As a passive observer and recorder of beauty, he is the antithesis to Sebastian, whose styles of art, worship, and being are evaluated in terms of a divine happiness— that "beatific vision" of beauty through the inauthentic, the crafted, and the passionately surreal.

Charles's very presence then serves as a deliberate threat to the Flyte tradition and the dandy aesthetic. He is a constant reminder to Sebastian of the common world he is trying to flee, and it is in this context that his skeptical, uninspired view of life and the divine is most

significant. Waugh plainly associates his quaint "charm" with his lack of true faith, in that both are characterized by his emotional dispassion and his preoccupation with placing the exotic in the realm of the familiar: "'I have left behind illusion,' I said to myself. 'Henceforth I live in a world of three dimensions—with the aid of my five senses'" (169). His relationship with Sebastian is a temporary exploration of the surreal, which eventually emerges as a fixation upon something to be contextualized, normalized, and rationalized: a dispassion that is equally as profane as Sebastian's amorality.

This dispassion makes *Brideshead* an autobiographically suggestive exploration of the "convert-artist" through Charles, and of Waugh's similar role himself as a writer. In the novel, the artist's job is to capture the fleeting and the impressionable that so vivifies the nature of true art itself. Because of this, we see Brideshead, the Flytes, and Sebastian all in the brilliant color of Charles' memories. Yet by the end of the novel, this noble, aristocratic beauty and all its embodiments (Sebastian, Julia, and the estate) have been effectively destroyed, and the artist is clearly to blame. Charles is the partner to his art, and the force that tethers it to reality, sobering up as Sebastian downs a drink, presenting marriage to Julia as she grapples with her state of "sin," and returning to the crumbling oasis of Brideshead as part of a modern military force. Waugh's message seems to be that this kind of otherworldly beauty is divine in its transience, though it may not be wholly moral, and that the artist's attempts to capture it are ironically its destruction.

It is clear that Waugh does not intend Charles to progress in the novel, despite Waugh's confirmation of religion as "the most important of all the agents that form a man's character" (Phillips 55), Charles's misleading "conversion," and the faintly flickering light of his faith housed in its deplorable lantern. As an artist without inspiration, he is doomed to perpetuate that which should not be clung to, effectively destroying all that he captures. The text is impressive in that Waugh conveys this sentiment via a multitude of plot lines, and reflects it again in the novel's very style and narration.

William J. Cook, Jr., argues that "Even though [*Brideshead Revisited*] possesses both comic and tragic potential, its loss of ironic perspective through the use of first person causes it to drift inevitably into melodrama" (208). But what Cook overlooks is Waugh's use *of* Charles' voice to convey irony in itself, in that by eliminating the narrative distance the third person introduces, Waugh involves his readers in the violation of Brideshead and its art on the most

basic, essential level of the text. Waugh does not directly jab at Charles for his tragically foolish attempts to cling to art in a barbaric and modern world. Yet the text is still highly ironic in that we as readers share in Charles' "sacred" memories before we realize that they are "profane." The narrative structure of the novel is also satirical, in the sense that Charles's mental revisitation of Brideshead's decline is framed by its physical revisitation by Captain Charles Ryder, and in a manner that blurs the lines between present and past quite unsettlingly. Charles' recollections are strange in their ambiguous shifts from the Charles of the past, of *that* moment, to the Captain Charles Ryder of *this* moment. Cook sees this as a fatal error:

> The flaw is not with the "then I" or the "intermediate I," but in the proximity of the "now I" with the past… There is no emotional separation of personae… All the events are filtered through a unique consciousness, which though seeing them in the past, colors them with present emotion; this is detrimental, for the "now I" is a romantic (216).

Instead of such a seeming "flaw," this "emotional separation of personae" is deliberate, and exactly what qualifies *Brideshead* as satiric— even tragic.

Charles's downfall as an artist is his failure to accept the intangible, and the "inauthentic" that necessarily serves as inspiration to counter bland and common "charm." In his world of crass soldiery, regurgitated art, and uninspired religious motions, the isolated oasis of Brideshead is doomed to destruction. Through these plot devices and his narrative choices, Waugh also expresses his own comparable guilt as a writer, in his nostalgic preservation of that which is worthy only of its own age. Waugh's activity as an author, recording stories based on his own mundane reality *outside* the novel, comes into question as a similar condemnation of the beautiful that exists *within* the novel. And if Charles's flaw is his compulsion to recollect, then are not Waugh and his readers, recording and interpreting stories based on their own mundane realities, equally to blame? While Waugh is sure to indict Sebastian's "languid" lifestyle and rejection of Catholic morality, there is also an inevitable tenderness reserved for this delicate, eccentric youth and the "beatific vision" that he is incapable of realizing. So although the dandy may be dissolute, detached, and even deluded, Waugh presents him not as the culprit of a destruction of beauty through "languor," but rather a victim of the artist and the culture in which that artist forces him to exist.

Works Cited

Cook, Jr., William J. *Masks, Modes, and Morals: the Art of Evelyn Waugh*. Cranbury:
Associated University Presses, Inc., 1971. Print.

Hardy, John Edward. *Man in the Modern Novel*. Seattle: University of Washington Press, 1964.
Print.

Heath, Jeffrey. *The Picturesque Prison: Evelyn Waugh and His Writing*. Quebec: McGill-
Queen's University Press, 1982. Print.

Lane, Christopher. "The Drama of the Imposter: Dandyism and Its Double." *Cultural Critique* 28
(1994): 29–52. Web. 23 Nov 2011.

Phillips, Gene D. *Evelyn Waugh's Officers, Gentlemen, and Rogues: The Fact Behind His
Fiction*. Chicago: Nelson-Hall, Inc., 1975. Print.

Creating the Conversation:
When the Classroom Tells All

Dana Livingston

As students of color became middle class, the criticism of persistent American racist practices declined. In a mid-level sociology course originally designed to empower black women students' evidence of discomfort, using materials that challenged racist practices, provoked a redesign of the course structure. This allowed for in-depth understanding and acceptance of history and the present day. Giving access to the ways that racism is viewed in America, a number of texts were reviewed and discussed throughout the course that expanded upon the idea of there being only one black America. This journal will present the analysis of the interaction between the twenty-one students of the State University of New York at Geneseo's sociological course Black Women in American Society. Furthermore, their research project and the redesign of the 2011 fall semester course, which implemented a new classroom approach by creating a space in which discursive connections, brought social transformation of academic institutions.

Student Demographics

In the State University of New York at Geneseo, only a handful of courses are offered that focus on the history of African Diaspora and the effects it has had on the current racial practices of the United States. These courses intend to expose what textbooks neglect and open the minds of all students to create a safe haven for discussion and interactions. The 201 sociology course Black Women in American Society consisted of twenty two students; five African American students, three Asian students, one Hispanic student, and twelve Caucasian students; one of them the only male in the course. All the students ranged from freshman to seniors. All students came from different states and cities of origin; metropolitan cities, small rural towns, and suburban communities. The economic status of the students ranged from lower class to upper-middle class. At the start of the semester the students were asked two questions; "What does it mean to be black?" and "What does it mean to be white?" the stereotypical

answers highlighted the link between experience and consciousness, thus, paving the way for course reform.

Course Texts

Course (reading) material was assigned by Elaine Cleeton, Ph.D, professor of the course and an Associate Professor of Sociology at the State University of New York at Geneseo, who has been with the university since 1997. Having previously taught this course, Dr. Cleeton assigned the reading material each semester with the course focus being black women and their status in the United States. The texts paved the foundation of the study of education, law, family, and the social norms that support the majority. The required texts included Michelle Alexander's *The New Jim Crow: Mass Incarceration in the Age of Colorblindness,* Pablo Freire's *Pedagogy of the Oppressed*, Janell Hobson's *Venus in the Dark: Blackness and Beauty in Popular Culture,* Beverly Guy-Sheftall's *An Anthology of African-American Feminist Thought*, and Eugene Robinson's *Disintegration: The Splintering of Black America.* Robinson's *Disintegration* provided the groundwork to the students' research project "but I'm not a racist…" modeling Theorist Dorothy Smith's Institutional Ethnography which explores the relationship between two parties beginning with interviews and observations of everyday performance to discover America's racial stance. Robinson discusses the class division of black Americans generation gap, stating that instead of one black America there is now four (Robinson 2010):

1. Mainstream middle class

2. A large abandoned minority

3. A small transcendent elite

4. Two newly emergent groups: individuals of mixed race and
    immigrants.

Each student interviewed a friend or family member with a list of questions the class developed as a whole, using all texts that covered areas of racial emergence; socio-economic class, education, the law, and current stereotypes which were to be presented at the college's annual research symposium.

Beginning of Semester Student Interactions

New semesters equal a new group of students who come together to discuss the subject being studied, which for a course like Black Women in American Society, could be considered uncomfortable to those who are not used to talking about race; its history and present day.

Entering the "colorblind" era and making sure that the United States is "politically correct" when discussing race acts as a shelter from the uncomfortable conversation. Therefore, it was no surprise to witness the lack of verbal communication.

Learning is hard to do without emotion; consequently, the lack of emotion gave no way for the class to grow together. The classroom structure consisted of a circle, where each student faced the center. This was done with the hope that students would be more likely to interact. Each class, there were assigned readings that would be discussed; the conversation was left open to relate the text to whatever the students wanted to.

For example, on September 14th 2011, the class discussed the text of Eugene Robinson's *Disintegration: The Splintering of Black America*. Robinson believes there is no longer one black America and the problem no longer lies in the fight to "overcome"; it now lies in the individuality of reaching the American Dream. The stereotype of black people sticking together is diminishing; many are trying to deviate from the negative connotation and stereotypes that are assumed to come with black America. The attempt to get a group of students, the majority of who did not know each other, or never had to discuss subjects like this was demanding.

When considering the discussion of challenging text and the force the classroom setting put upon the students, many still did not engage in conversation. A hypothesis was developed as to why the students did not speak:

1. The students did not want to say something offensive
2. The fear of not being politically correct
3. Not being able to relate to the text
4. Believing they did not have the right to say anything.

Without discussion, the students were not able to determine if they were in a comfortable setting, one in which they could freely discuss their opinions and thoughts. "Without dialogue, there is no communication, and without communication there can be no true education" (Freire 92-93). Therefore, something had to be done to create this setting.

Questions We Asked Ourselves

Ms. Patricia Gonzalez, a counselor for the Access Opportunity Program, a program designed to assist students who were previously under-represented in their earlier levels of education with the transition to the State University of New York at Geneseo, joined us in our goal to implement a new classroom design. Before we could begin, we had to ask ourselves a

series of questions before planning a course reform; how is knowledge produced and how, then, do we engage our students in consciousness rising?

Our goal was to implement pedagogy from a dialogical perspective while having the students develop their own educational experience; by teaching and learning through each other.

Course Reform

Dialogue is key. Dialogue:

1. Is an act of love

2. Cannot exist without humility

3. Requires intense faith in humankind

4. Engages the dialoguers in critical thought.

This resulted in the use of a strategy called "concentric circles." This provides a "here and now" experience which fully engages the students with the material. The structure was intended to generate development in communication between the students in a cultural critique of racism and sexism.

"Concentric Circles" involved two circles of chairs, one circle inside the other. All the chairs of both circles face each other so each student would face the student in front of them. One student was to ask a question, comment on the answer of the other student, and then answer the question themselves.

The conversation and questions of this activity were used to create questions for the students' research project and create the comfortable setting in which students could talk among themselves, getting to know each other with the hopes of increasing consciousness and class discussion.

As a result of this activity some of the questions the students asked were used in their interviews:

1. Do you feel there is a lack of diversity at Geneseo? Is it changing?

2. Do you think race/class affected how we responded to Hurricane Katrina?

3. Do you believe that the black community still has common ground even if race is the only thing that seems to unite them? What about Blacks of mixed race or immigrants? Do cultural beliefs, such as food or religion also come into play?

These questions proved their consciousness and awareness for the issues and topics in the readings.

What the Classroom *Now* Tells

This activity proved to be successful in acknowledging consciousness and creating a comfortable environment for students to engage in conversation of challenging texts throughout the remainder of the semester. At the end of the semester we were able to freely talk with the students and all students shared personal stories, experiences, and opinions. We were able to grow together as a class.

To take a course and be able to leave with a positive academic and social experience is what true education is all about. For some students this course was the first encounter they have ever had with a diverse group of students. In lieu of this the level of engagement and eagerness to make change as a result of our activities and discussions was an amazing accomplishment. Thus, it is prevalent to share our concept, activity, and accomplishment with others so together we can break racial barriers.

By Dr. Cleeton, Ms. Gonzalez, and I sharing our research and classroom experiences at New Orleans University's Race Gender and Class Conference in New Orleans, this gave way for this research to be further examined and provide an example of what dialogue does in an academic setting.

It takes communication and faith in humankind to learn from one another and make change happen.

Works Cited

Alexander, Michelle. *The New Jim Crow: Mass Incarceration in the Age of Color Blindness.* New York: The New Press, 2010. Print.

Guy-Sheftall, Beverly. *Words of Fire: An Anthology of African-American Feminist Thought*. New York: The New Press, 1995. Print.

Freire, Paulo. *Pedagogy of the Oppressed.* New York: Continuum, 2002. Print.

Hobson, Janell. *Venus in the Dark: Blackness and Beauty in Popular Culture.* New York: Routledge, 2005. Print.

Robinson, Eugene. *Disintegration: The Splintering of Black America*. New York: Doubleday, 2010. Print.

Creating the Conversation:
When the Classroom Tells All

Elaine Cleeton, Ph.D.
Patricia Gonzalez
Dana Livingston
State University of New York at Geneseo

- "…[T]he dialogical character of education as the practice of freedom does not begin when the teacher-student meets with the student-teachers in a pedagogical situation, but rather when the former first asks herself or himself what she or he will dialogue with the latter about." (Freire, 93)

Creating the Conversation

- How is knowledge produced?
- How, then, do we engage our students in consciousness-raising?

Questions We Asked Ourselves

- "This pedagogy makes oppression and its causes objects of reflection by the oppressed, and from that reflection will come their necessary engagement in the struggle for their liberation." (Freire, 48)

Facilitating Consciousness

- Is an act of love
- Cannot exist without humility
- Requires intense faith in humankind
- Engages the dialoguers in critical thought

Dialogue

- Objectives:
  - To provide a "here and now" experience
  - To fully engage students with the material

Concentric Circles

- Class research project
- Eugene Robinson's <u>Disintegration: The Splintering of Black America</u>
- Ethnographic open-ended interviews
- Research report

Course Structure and Content

- Have you ever felt out of place because of your race?
- Where did you grow up? Were you in the majority or the minority?
- Do you feel that there is a lack of diversity at Geneseo? Is it changing?
- What have you experienced with class/race that made you uncomfortable?
- Do you think race/class affected how we responded to Hurricane Katrina in New Orleans?
- How is the mixed race perceived? Should Obama be considered the first Black president even though he is half white as well?
- Do you believe that the Black community still has common ground even if race is the only think that seems to unite them? What about Blacks of mixed race or immigrants? Do cultural beliefs, such as food or religion, also come into play?

Questions the Students Asked

- "Without dialogue, there is no communication, and without communication there can be no true education." (Freire, 92-93)

What DOES the Classroom Tell?

Thank you

Ethics of Informed Consent in Difficult Medical Decisions: Spina Bifida Diagnosed Before Birth

Max Evan Jacobson

Part 1: Scenario

The contents of this paper center on a specific, hypothetical scenario that may occur in the medical profession. Consider a young man and his wife, who is pregnant with their first child. In the third trimester, a routine ultrasound shows a myelomeningocele on the fetus. A myelomeningocele is a defect formed late in the first month of pregnancy. As the spinal cord grows, it develops into a sac outside of the spinal column, which is known as a myelomeningocele. The result of this defect is that the nerves beginning below the myelomeningocele are unable to communicate properly with the rest of the spinal cord and the brain (Wolraich 3). The defect is known as Spina Bifida after birth. Upon discovering the myelomeningocele, the physician schedules a follow-up appointment with the couple for the next day. He or she will then be responsible for informing them of the myelomeningocele and providing information about it. In addition, the physician will lay out two options of how to proceed: either allowing the pregnancy to come to term or abortion.

Part 2: The Ethics of Communication

After the interactions between the physician and the parents, the parents will make a decision of whether to abort the fetus or to let the pregnancy come to term. Because the parents must make this choice, the physician must obtain informed consent. Informed consent is simply an act made freely and rationally by a patient based on accurate information (Wright 97). Obtaining informed consent from the parents is an important responsibility for the physician. If they are poorly informed, coerced into a poor decision by the physician, or not allowed to decide for themselves, then the physician disrespects the right of the parents to make their own medical decision (Beauchamp 118). Therefore, informed consent to the course of action chosen by the parents and physician is essential if autonomy is to be upheld.

Informed consent has seven elements, each of which the physician has an ethical responsibility to uphold in any medical decision-making process (Beauchamp 120). The next section will examine each individually. I will pay more attention to some over others, as several are most relevant to the specific example in this paper.

1: *Competence*

In order for a patient to give informed consent, the patient must be able to understand the information given and be able to make a decision (Beauchamp 120). This could be a problem when dealing with patients that are mentally unstable or disabled. Also, patients that are unable to communicate and understand relevant information could be considered incompetent (Beauchamp 114).

2: *Voluntariness*

A patient must want to decide for him or herself in order to meet the requirement for voluntariness (Beauchamp 120). In addition, the patient must be allowed to decide; if a patient is under the control of another's influence, then he or she does not act voluntarily (Beauchamp 132). A patient could also lose voluntariness if coerced, persuaded, or manipulated by another individual. In the specific scenario in this paper, the physician must be sure that he or she does not directly influence the parents' decision. Whatever choice that they make in the end should be accepted. In addition, the physician should be sure that the parents are not under the control of any outside influences as they make their decision.

3: *Disclosure*

Disclosure is the actual presentation of the information from the physician to the patient (Beauchamp 120). Considering that this paper will focus mostly on this step of the informed consent process, I will examine disclosure in considerable depth. Firstly, let's consider how physician bias can affect the patient's decision making process.

3a) Physician Bias

Because the decision of whether to go through with abortion or birth should be made by the parents, the physician has an obligation to leave his or her own personal interests out of the decision-making process. When the physician's professional opinions are affected by personal opinions, the physician acts with bias. For example, consider if the couple elects to abort the fetus. The physician disagrees because he or she is a devout Christian, and thus believes that abortion is immoral in such a case. Although the physician may feel this way, he or she should

not attempt to sway the parents' opinion due to this moral opposition. Doing so would create a conflict of interest, because the physician's personal interests interfere with his or her professional responsibility to respect patient autonomy. Therefore, the physician must always try to remain impartial (Beauchamp 314).

One specific bias that may occur has to do with perception of the disabled. Physicians are successful people who often happen to be in good shape and healthy. They usually have spouses and children as well. This high-achieving lifestyle may lead some physicians to devalue disabled lives. To a successful physician, a disability may seem unbearable and unimaginable; however, many disabled people live happy and normal lives despite their setbacks (Gill). To avoid making this implicit judgment, the physician should be familiar with the disabilities specific to Spina Bifida. In addition, the physician must always be objective, allowing the parents to assess the value of the child's disabled life.

The physician must not be biased in communication with the parents, but must he or she always be truthful? In order to answer that question, let's examine the nature of truthfulness and deception in the next section.

3b) Truthfulness and Deception

In general conversation, individuals expect the truth from others. People assume that their peers will not lie, and therefore they also are expected not to lie (Wright 84). In medicine this principle is considered quite important, although veracity has not always been valued by physicians in the past (Beauchamp 288). In the context of the informed consent process, physicians are obligated to give accurate information to patients (Beauchamp 289). This ensures that they have the correct details necessary to make a favorable decision.

In contrast to veracity, deception is also possible in medicine. This is any intervention that alters patient beliefs (Wright 85). Although lying is one form of deception, there are other possible ways for physicians to deceive patients. For example, a physician could withhold information or fail to disclose information. This would prevent the patient from having all of the information necessary to make an informed decision (Beauchamp 289).

There are several important reasons for veracity over deception in medicine. In general conversation between two individuals, each expects the truth from the other. Therefore, it is also true that patients expect the truth from their physicians; they enter conversations expecting to be treated fairly and respectfully. This is actually part of the reason why lying and deception by

physicians can work so well when employed; patients believe what their physicians have to say and do not anticipate that they will be deceived (Wright 94). Because truth-telling is an unwritten promise of normal conversation, this can also be held to communication between the physician and patient (Beauchamp 289).

Another reason to treat veracity seriously in medicine is that patients must also be truthful with their physicians. Good medical practice requires that accurate background information, habits, feelings, and symptoms are relayed from the patient to the physician. Without these details, the physician may not be able to make an accurate diagnosis or prognosis. If the physician has the right to the correct information from the patient, it follows that the patient also has the right to accurate details from the physician. The patient-physician relationship is built on trust; in general, neither party should violate that trust to achieve optimal care (Beauchamp 289).

Because competent patients have autonomy, the right to make their own decisions, they also have a right to the truth with which to best make those decisions. Deception directly interferes with this right, by giving the patient false information on which to make a choice (Wright 95). In order to examine whether deception can ever be acceptable in medicine, let's first examine withholding information and how this practice may be moral in certain circumstances.

3c) Reasons for Deception by Withholding Information

Limited disclosure of information can be a harmful form of deception because it may prevent patients from having the necessary details to make an informed decision. However, it is important to note that no physician can possibly know the entire truth. Medical technology does not allow him or her to look into the future and exactly determine the child's quality of life. Therefore, the account given by the physician is an approximation of what is likely to occur (Beauchamp 292). This is a specific problem when considering Spina Bifida because the condition results in extremely variable outcomes. For example, two individuals born with similar myelomeningoceles may end up living drastically different lives; one could have a family and a full-time job, while the other relies on an assisted-living home (Hetherington *et. al)*. Similar degrees of variability exist for many different aspects of Spina Bifida, and will be explored in more depth throughout the paper. Because of the nature of this condition and because physicians cannot know every detail possible, no patient consent can ever be completely informed (Wright

99). Therefore, the physician must simply give a disclosure that is appropriate enough for the patient to make a well-informed decision (Beauchamp 292).

There are several other reasons why limiting the information presented may be beneficial. For example, oftentimes there is simply too much information to tell. If parents were expected to get every piece of information possible and then use it all to formulate a decision, there would be too much to process (Beauchamp 292). The excess detail would simply prevent acceptable understanding of the condition and cause indecisiveness. In addition, much of the information is extremely complex. Too many unfamiliar, medical terms may overwhelm patients and simply cause confusion (Beauchamp 130). Diving into intense research work with patients may also have a similar effect.

The amount of information given may also depend on the willingness of the parents to listen. For example, consider a case in which the parents did not trust physicians' advice. Perhaps the couple believes that they can figure out what they should do on their own and they do not want to hear what their physician has to say. In such a case, the physician should limit disclosure in order to give a simplified account to the stubborn parents. Withholding information may also be beneficial depending on the ability of the parents to understand. Should the parents be well-educated, as they are in our scenario, a physician may wish to give them more specific details of the condition. However, this may not be the case if the parents are poorly educated. Because they will not understand much of the information, a simplified account will be most useful.

Information could also be acceptably withheld in our scenario if the parents do not want to know the details about Spina Bifida. When the physician tells them about the malformation, they say that they prefer not to hear the consequences until later. They argue that they are devout Christians and would thus never choose to abort the fetus. Therefore, rather than fret about the struggles that they will face later in life, they prefer to remain oblivious until after the child is born. Because the parents have promptly made their decision and because they do not wish to know the information, the physician has no obligation to give the details until later. However, the physician would still be responsible for informing the parents after the birth. Although this type of scenario could be possible, it is more likely that the parents would want to learn about the condition in order to at least weigh their options (Wright 86). Normally, the physician would still be obligated to inform the parents.

4: *Recommendation*

Once the information is fully presented to the patient, the physician has the opportunity to make a recommendation. In this case, the physician may indicate whether he or she feels that abortion or birth is the best alternative for the couple (Beauchamp 120). It may be best for the physician to wait until the parents have thought about the options for some time until this recommendation is made. This is because the parents will surely be stressed and vulnerable immediately after the information is made available. Should the physician quickly make a strong recommendation, it may lead the couple to agree even if this decision is not best for them. He or she could say something like, "I know this must be difficult for you to hear and process. Feel free to take some time to think this over and make sure that you are making the best choice for both you and your family. Should you need my help, I will be a phone-call or meeting away in order to give any recommendations that you may need. Please don't hesitate to ask me for help or guidance at any point over the next few days." Because this message doesn't immediately give a recommendation, it does not tread on the parents' right to make an autonomous decision. However, it also offers help for the couple if they need it. Therefore, the parents can turn to the doctor if they are struggling.

5: *Understanding*

After the physician's recommendation, it is essential that the patient understands the disclosure and recommendation; the patient must have acquired all of the information necessary to make an adequate decision (Beauchamp 120). Diagnoses, prognoses, nature of intervention, alternatives, risks, benefits, and recommendations are commonly considered necessary for acceptable understanding (Beauchamp 127).

There are several common problems that may interfere with patient understanding despite effective disclosure of information. For example, patients are often nervous and distracted during the disclosure process (Beauchamp 127). This is because the introduction of bad news causes stress, which may lead to these other problems. Even if the physician gives the best information possible to the patient, he or she may not process much of it. Therefore, the physician must be sure that a given patient is calm and collected during the disclosure process. If not, the physician could give extended time for the patient to relax or help by comforting the patient directly (Beauchamp 127).

Risks may be particularly difficult for patients to understand. This is because their comprehension of the specific risk depends strongly on how the physician explains it. Oftentimes, giving the chances in a numeric way and also a nonnumeric way, such as an analogy, may help put the risk in real-life terms (Beauchamp 128). For example, a physician could say, "Your risk of needing a revision procedure in the next year is about the same as your risk of winning the lottery." This could help the patient understand the low-risk scenario better than just saying that the circumstance is highly unlikely. Another problem with risks is that patients tend to develop disproportionate fears of them. The odds of this occurring are correlated to how the information is presented. For example, consider a patient that has a fifty percent chance of surviving from cancer after a surgery. The physician could relay this piece of information in two different ways; he or she could say that the patient has a fifty percent chance of death or a fifty percent chance of survival. By putting this detail in terms of death, the patient is more likely to be afraid of the procedure; he or she may feel unnecessarily pessimistic. However, by saying that there is a fifty percent chance of survival, the patient could feel more optimistic and empowered. Just because of the terminology, the patient may be less afraid. Therefore, physicians must be thoughtful in how they describe risks (Beauchamp 130).

A final and critical part of understanding is the opportunity for the patient to ask questions. Patients must have this chance because they have a right to know any information that they feel is important. In order to ensure that this occurs, the physician should be comforting and compassionate. Patients will not feel likely not feel comfortable inquiring if the physician seems stern and uncaring (Wright 98).

<div align="center">6: Decision</div>

After the patient understands all of the information necessary, the patient must come to a decision of which plan to accept. It is important that he or she is given sufficient opportunity to change his or her mind as well. The physician should allow this without any sort of penalty. As a result, the patient cannot become bound to a previous decision before the procedure is underway (Wright 98).

<div align="center">7: Authorization</div>

Lastly, the physician is responsible for authorizing the plan chosen by the patient. The physician must confirm the decision and make proper arrangements for the remaining course of action (Beauchamp 120).

Part 3: Suggested Preparation for Disclosure

Now that we have examined the informed consent process in detail, let's return to the specific scenario outlined at the beginning of the paper. The physician has discovered the myelomeningocele in the fetus and has told the parents that he or she would like to meet with them the next day. The physician knows that the parents are competent and will want to make the decision; thus, the first two elements of the informed consent process are fulfilled. Next, the physician is responsible for disclosing the information to the parents on the following day.

The remainder of this paper will focus on the main variables that the physician must consider when preparing to disclose information to the parents. I divided these areas into eight, broad categories as follows:

1. Physical Limitations
2. Mental Limitations
3. Childhood Experiences
4. Adolescent Experiences
5. Emotional and Sexual Difficulties
6. Professional Limitations
7. The Nature of Disability
8. The Parents' Best Interest

Within each category, there are multiple details that must be explored. This work contains many, but not all of those aspects of Spina Bifida that I explored on GREAT Day. For each aspect included in this paper, we will examine several important questions that the physician must ask when preparing for disclosure:

1. *What do we know about this*?

One of the physician's primary duties is to fully inform him or herself before addressing the parents. To them, the physician is the expert; he or she is their foremost resource on the condition and the potential options. Therefore, the physician must closely examine all aspects of the condition in preparation for the meeting. For example, he or she should closely evaluate the possible outcomes, consequences, and assess the likelihood of each.

2. *What variability exists*?

Because Spina Bifida outcomes tend to be so variable, the physician must ask this question when examining each detail. He or she should be aware of the possible range of outcomes, as well as the approximate odds of each occurring. This will help the physician draw conclusions about disclosure of this information to the parents. If a specific outcome is extremely likely, it generally should be mentioned during disclosure; it is an outcome that the parents and fetus are likely to face. However, if an outcome is extremely variable with a wide-range of outcomes, then the physician must closely consider how to best disclose it.

3. *How should this information be presented to the parents*?

Firstly, should the detail be presented at all? Some information is better withheld from the parents, as too much can be overwhelming. Therefore, the physician must consider whether the given detail is important for the parents to consider when making their decision. If it is essential for the parents to know the given detail, then the physician must decide if the information will be presented fully. Should some aspects be left out for a specific reason? Would the parents be better off not knowing one of the aspects of the condition? Once these questions have been answered, the physician needs to consider how the information is relayed to the parents. Can he or she make the details more manageable to understand? Can analogies be used? What is the best way to present a specific risk?

Also, the physician must look for opportunities to be comforting and compassionate to the couple. During disclosure, they may still be shocked and devastated from receiving their unexpected news. Part of the physician's job is to diminish anxiety and make the experience less traumatic. Not only will this make the parents less miserable, but it will also ensure a better decision. With limited stress, they will be able to think more lucidly about the costs and benefits of each option.

Let's first apply this methodology for disclosure preparation to some of the physical limitations associated with Spina Bifida.

Part 4: Disclosure of Physical Limitations

Mobility

1. *What do we know about this*?

After birth, physicians will closely examine and evaluate the child's myelomeningocele. The most important indicator of the child's future mobility will be the specific location of the

lesion on the spinal cord. Most occur in either the lumbar, sacral, or thoracic regions. The child's future ability to move his or her legs will depend on the specific vertebrae upon which the myelomeningocele is located. The movements possible are summarized below:

T-12 or above: no leg movement at all

L1: Hip flexion (bending legs upward at hips)

L2: Hip adduction (pulling legs together) + (above abilities)

L3: Knee extension (straightening legs) + (above abilities)

L4: Knee flexion (bending legs at knees), ankle dorsi-flexion (bending feet up) + (above abilities)

L5: Hip extension (straightening legs at hip joint), hip abduction (pulling legs apart) + (above abilities)

S1: Ankle plantar flexion (bending feet downward) + (above abilities)

S2: Movement of feet and toes + (above abilities)

The further the myelomeningocele is down the spinal cord, the greater the degree of leg and foot movement that is possible for the child. Should the lesion be in the sacral region or low in the lumbar region, the child will have more control of the hips and knees. However, children with this type of injury will still struggle with their ankles and feet to some degree. This is because every nerve pathway that begins below the level of the lesion will not function in the child (Wolraich 12). Follow-up studies have shown that physician predictions of future mobility and limb function are quite accurate at birth by closely observing the myelomeningocele (Laurence 153).

 Because muscles are often controlled by nerves that arise from different locations on the spinal cord, muscle imbalances are a common problem. For example, parts of the hip are controlled by L2 and others are controlled by L4 and L5. If the myelomeningocele occurs in one of these spots in the lumbar region, some muscles will be underdeveloped and immobilized while others are normal. This resulting imbalance could lead to dislocations of bones out of joints or subluxing, which is when bones are partially pulled out of joints. This commonly occurs with the head of the femur, or acetabulum, in the hip socket. In addition to dislocation, bones may also become deformed as they develop. Joints often tighten and contract too; this is common in the knee joints, causing the leg to stay straight and rigid. Another related problem is clubbed

foot, which is when the foot is permanently rotated internally at the ankle. This makes it very difficult for the child to put pressure on the feet and walk normally (Wolraich 13).

The child and parents will visit physical therapists and orthopedic physicians frequently during the early years of life. Using different strategies, they will try to help the child learn to walk as much as he or she possibly can, depending on the level of disability. In addition, they will help the child stretch and strengthen in order to counteract contractures and muscle imbalances (Wolraich 12).

These two resources will also help the child use braces. Braces are designed to help the child stand and walk while also acting to help prevent contractures, malalignments of joints, and other deformities. There are several different types of brace, depending on the level of the child's myelomeningocele and the severity of the disability (Wolraich 19). Children with very limited mobility will use a body brace; this will give them the hip stability that they need. Leg braces can be used for children with lower level lesions, as they can better control their hips (Wolraich 20). Once braces are selected, they will have to be checked with a physician every 4-6 months in order to ensure that they are still effective and properly functioning. In addition, one pair will be worn at night and a different pair will be worn during the day (Wolraich 19).

Even though braces are used to help the child walk and stand, he or she will still struggle to move long distances. Therefore, wheelchairs are useful to help overcome this challenge. Even if the child has a low-level myelomeningocele, he or she still will need a wheelchair for some occasions. However, the wheelchair will not be used as a complete substitute for walking. The child still must wear braces and try to walk in order to diminish contractures and prevent deformities. Like there are many types of braces, there also are many types of chairs. The orthopedic physicians, along with the parents, will pick one that best suits the child's needs (Wolraich 32).

2. *What variability exists*?

The location of the lesion is the most important factor in determining the future mobility of the child; there is a great deal of variability in how capable the child will be depending on where the lesion is located. Therefore, a lot depends on this diagnosis at birth. The location of the myelomeningocele will govern what braces are worn, what surgeries are performed, how bad contractures are, how often the child will be in a wheelchair, and how much control the child has over his or her lower body.

Although this variability occurs, every child still will have notably decreased mobility. Braces, crutches, and wheelchairs are all extremely likely for the child no matter where the lesion occurs. The child is destined not to walk normally and mobility will be a difficult task. The location of the myelomeningocele governs the severity of the disability, but each child will be disabled nonetheless.

3. *How should this information be presented to the parents*?

Because impaired mobility is the most severe disability caused by Spina Bifida, parents need to understand what their child will face. They must understand that the child will need braces, crutches, and wheelchairs in order to be mobile, no matter where the myelomeningocele occurs on the spinal cord. This information could be extremely critical in their decision-making process. Although their child will be disabled no matter where the lesion occurs on the spinal cord, the physician should illustrate some of the variability in degree of movement to the parents. This will help the parents understand that their child may have the ability to move some parts of the leg, and if lucky, move the hips and knees quite well. However, the physician should still stress that the ankles and feet will be problematic. Because of this, the child is highly unlikely to walk normally no matter where the lesion occurs.

In addition to explaining the outcomes, the physician should also go over the consequences and reasons for muscle imbalance. This would not be as necessary if the physician was dealing with highly uneducated parents. However, this couple is college-educated and could probably understand why the muscle problems develop. The physician should also include that this why the child will need braces; however, it is not beneficial to explain the specific types of braces and instructions for use. This information is more useful for the parents after the child is born because it should not influence their decision.

It is also important for the physician to emphasize that mobility will require a lot of work for the parents and child. Braces must be worn at all times and exercises and stretches must be done to prevent further problems. In addition, parents will need to take their child to physical therapists and orthopedic doctors frequently. Therefore, the couple should be aware that they will need to invest a lot of time into mobility training. However, the physician could also include some good news. Studies and interviews with parents have shown that parents tend to do really well with this type of training, even though it may initially seem overwhelming (Hewson 118).

The inclusion of this optimistic detail provides balance against the difficult information about limited mobility, helping the parents see the positive side of raising this type of child.

<p style="text-align:center">Incontinence</p>

1. *What do we know about this*?

   Incontinence is the inability to control the excretion of urine and the excretion of feces. Firstly, we will look at urinary incontinence. Because the nerves that control the bladder originate low on the spinal cord, children with Spina Bifida often have neurogenic bladders. This means that they are unable to control the excretion of urine (Wolraich 40). If urine release is not tightly regulated, damage to the kidneys could occur. This is because if urine is not escaping the bladder, extra urine may remain in the kidneys and put pressure on them. To avoid this problem, urination must be controlled in a child with Spina Bifida. There are several ways to accomplish this feat, including manual emptying or collection in a bag (Wolraich 38). Manual methods of control will be discussed later, as they are most relevant to the parents' daily responsibilities. Another option to control urination is a surgery called Urinary Conduit Diversion. In this procedure, the path for urine is diverted into the bowels as a conduit to the outside of the body. The urine travels from the bowels into a bag exterior to the body. Although this may avoid kidney problems, it could result in some unfavorable side effects. These include skin irritation, fever, and foul-smelling urine (Wolraich 47).

   Bowel control will also be limited in a child with Spina Bifida. The lack of functioning nerves around the bowels causes neurogenic bowel, so the child cannot tell when he or she is full and cannot control the release of feces. The biggest problem that this causes is constipation, which is very common in children with Spina Bifida. If the constipation gets bad enough, it could lead to a bowel blockage or impaction. Should this occur, it represents a medical emergency; a child with impaction must be taken to the hospital immediately. Constipation can also lead to hemorrhoids, prolapsed, fissures or cracks around the anus, and a decrease in appetite. There are several methods to prevent these problems; they will be discussed when we address parental responsibilities later in the paper (Wolraich 64).

2. *What variability exists*?

   Although parents should expect to have an incontinent child, there are differences in severity. For example, some will not have hemorrhoids or severe constipation if the condition is

managed well. In addition, the outcomes are different depending on the treatment. A child who gets Urinary Conduit Diversion may experience different side-effects than a child who does not.

3. *How should this information be presented to the parents*?

Incontinence is one of the most severe physical consequences of Spina Bifida. Therefore, parents must be aware of this problem. The physician should explain why incontinence occurs and stress that it can be controlled. However, the main concern for the parents will be the daily care associated with incontinence management. This will be among the greatest stressors that parents face when caring for their new child. We will devote an entire section to this aspect of incontinence later in the paper.

The physician need not inform the parents of consequences of inadequate care until well-after birth. This is because such symptoms, such as constipation, can be prevented by closely managing the child's toileting. Therefore, the parents do not need to know that such an outcome can exist. Rather, they must know that the associated care is important and is very hard work. The potential consequences can be adequately addressed after the child is born.

Part 5: Disclosure of Mental Limitations

Hydrocephalus

1. *What do we know about this*?

Hydrocephalus is a condition that commonly occurs when a child is born with a myelomeningocele. Because the lower part of the brain tends to dip lower into the spine than a normal brain would, cerebrospinal fluid is less able to drain from the brain. The fluid cannot get as easily into the subarachnoid space, from where it would drain into the circulatory system. Because the fluid gets trapped in the brain and because the bones of the skull are not yet fully formed, the child will likely be born with an unusually large head circumference. In addition, the condition can result in brain damage and other serious problems if it is not treated quickly (Wolraich 6).

Therefore, a shunting procedure will be done in order to correct the hydrocephalus. It will be performed in the first few days after which the child is born. The procedure involves putting a valve into the brain that drains through a catheter into the peritoneal cavity. Incisions must be made in the brain, usually outside the occipital lobe, and also in the abdomen. In a fully-functioning shunt, excess cerebrospinal fluid is drained through the shunt and dumped into the

peritoneal cavity, where it can be reabsorbed into the circulatory system. Thus, the amount of fluid in the brain can be regulated and hydrocephalus can be controlled (Wolraich 8). Recovery from this procedure is routine and should only take a couple of days.

One major problem with shunting is the frequent need for revisions. Since physicians began performing the procedure in the 1950's, it has been one of the most highly revised surgeries performed in the United States. Even if the shunt works initially, the system tends to malfunction, requiring emergency surgery for the patient. Symptoms of shunt malfunction include nausea, headaches, vomiting, and lethargy. There are many reasons for the system to stop working appropriately, including clogging of the catheter, infection, valve failure, over-drainage of the system, and under-drainage of the system (Wolraich 8).

A prominent concern regarding hydrocephalus is that it may have lasting impact on the child's intellectual ability. Those individuals who have many complications with hydrocephalus and many revisions tend to have greater intellectual defect (Laurence 153). In contrast, those patients with very few revisions tend to be intellectually well-developed (Laurence 154). Therefore, hydrocephalus appears to the best indicator of future intelligence for a given child. One study measured that increased bouts of hydrocephalus and shunt revisions led to decreased performance IQ. Because performance IQ was closely related to functional math skills, individuals with more hydrocephalus were less able to manage money, estimate, and shop; this resulted in less independence and a lower quality of life for these patients (Hetherington *et al*).

2. *What variability exists*?

Although nearly all children with Spina Bifida will be born with hydrocephalus, this condition is one of the most variable aspects of this birth defect. Children who are shunted could have shunts that function throughout life, or they may not even need it after a certain point. However, there are other individuals who may need upwards of 10 or more revision procedures due to malfunctioning shunts. The extreme degree of variability in hydrocephalus and shunt revisions leads to high variability in the intelligence of Spina Bifida children. Those who are not afflicted by hydrocephalus throughout development will have close-to-normal intelligence; however, those who have complications may be inhibited later in life. Therefore, this variable may be the most unpredictable of all aspects we will examine in this paper.

3. *How should this information be presented to the parents*?

The high degree of variability makes information about hydrocephalus difficult to disclose. Certainly, the physician has a responsibility to inform the parents about the condition and say that it will be treated shortly after birth. In addition, a description of shunting is beneficial for a well-educated couple who can understand it. The difficulty comes when the physician describes how hydrocephalus affects intelligence. There is no way for the physician to know or predict whether or not the child will need many shunt revisions or if the child will not require any; the outcomes are simply too variable.

Therefore, the physician should emphasize to the parents that Spina Bifida is more of a physical disability than a mental disability. This is a good way to frame the discussion because the range of intelligence outcomes is extremely broad. In addition, the physical concerns associated with limited mobility and incontinence are of far more concern than the possibility of limited intelligence. The physician could emphasize to the parents that the condition does not cause mental retardation. Rather, there is simply a chance that the child will not develop intellectually as would an unaffected child. Either way, the child should be able to go through school with special accommodations. This optimistic approach indicates to the parents that the intellectual effects of Spina Bifida are not severe. Although there could be problems, they will not be as severe as the physical hindrances.

Part 6: Disclosure of Adolescent Experiences
Worrying & Depression

1. *What do we know about this*?

Teenagers with Spina Bifida express worry about many areas of life during interviews with parents and the teens themselves. We will discuss more of these specific worries as we discuss more categories throughout the paper. However, it is important to note that some Spina Bifida teens said that they worried about "everything," when asked what caused them to worry (Anderson 126).

The incidence of depression is higher in adolescents with Spina Bifida than it is in the general population. In addition, the percentage of adolescents with Spina Bifida and depression is higher than the percentage of adolescents with cerebral palsy and depression (Anderson 128). Dorner performed a study on 63 teens with myelomeningocele and differing degrees of mental handicap, finding that 66% experienced feelings of misery. These feelings were definite and

recurring, thus indicating depression. More girls than boys experienced the symptoms as well. This result was replicated by McAndrew, who additionally found that low self-esteem was extremely common in the teens studied (Anderson 117). Later studies found that 31% of girls and 15% of boys had persistent feelings of depression and/or suicidal thoughts. Mobility could be related to depression, as decreased mobility results in more social isolation for a teen (Anderson 128).

Even if a teen does not have chronic depression, he or she may have depressive symptoms. Consider Patrick, a 15 year-old interviewed by researchers. During the interview, he was happy and upbeat. However, he described that pressure sores on his backside kept him out of school for nearly an entire year, as he was confined to his bed. Patrick said that he was extremely depressed during this time period, as he was missing time at school (Anderson 11). This shows that complications and extended hospitalizations may cause bouts of depression because the teenager experiences limited social interactions.

2. *What variability exists*?

Not all children will face depression, and there is no guarantee that a given child will exhibit depressive symptoms either. Children with Spina Bifida are more likely to develop depression, but are not certain to do so. In addition, it is important to note that many adolescents are unhappy and experience depression; this symptom is not exclusive to Spina Bifida teenagers.

3. *How should this information be presented to the parents*?

Disclosure of adolescent depression is a tough subject to consider. Firstly, it is not of immense concern to the parents, as they will be most worried about the childhood experiences and adult outcomes for their child; the time in-between may not be as important to them. Secondly, it is uncertain whether or not a given adolescent will have depressive symptoms. It depends on the level of mobility and the amount of complications, among many other factors. Thirdly, depression is something that can happen to anyone; it is not exclusive to teenagers with Spina Bifida. These three factors indicate that depression should not be discussed in great detail with the parents.

Still, the increased odds of depression during adolescence are important enough to mention in a conversation about difficulties during this time of life. The physician could simply mention that some teens with Spina Bifida develop depression, and that this may be related to the degree of mobility. This information should not be a focal point for the parents, as it is far from a

certain outcome. However, they deserve to know that this is possible. This should be discussed with the social adjustment of Spina Bifida teenagers, which is our next category to examine.

Part 7: Disclosure of the Nature of Disability

Negative Aspects of Disabled Life

1. *What do we know about this*?

One major problem that disabled people face is discrimination due to their minority status. They tend to be seen as different and treated as such socially, medically, and even legally. People consider them strange and oftentimes are naturally averted to them. Doctors are less likely to adequately diagnose depression in disabled patients, simply attributing this mental disorder as a side-effect of the disability. In addition, there are many social stress factors that commonly plague disabled individuals. Some prominent ones are neglect or abuse in childhood, family dysfunction, low self-esteem, and unstable sense of identity (Gill). Even though a child may be able to adjust to disabled life over time, he or she will still be susceptible to these potential hardships.

Another problem that disabled individuals may face is that physicians are occasionally incapable of adequately working with them. In medical school, physicians often do not focus on learning about handicapped life. The instruction focuses more on the science of the disability and the resulting treatment necessary. Therefore, physicians may not be too helpful dealing with problems specific to disabled life, such as concerns about social prejudices.

Patients with Spina Bifida may have a particularly difficult time finding a hospital that specializes in all of their problems. For example, a patient will have to travel to an urologist, see a capable neurosurgeon at a different facility, and visit a speech pathologist elsewhere. This can make attaining adequate healthcare rather difficult for these patients, as a facility well-suited to dealing with all Spina Bifida problems is hard to find (Anderson 315).

2. *What variability exists*?

Not all disabled individuals will suffer from social stress and depression. However, most are likely to experience some concern with how they are perceived and treated by others.

3. *How should this information be presented to the parents*?

This information is probably best left out of disclosure. When making the decision of whether to go through with an abortion or complete the pregnancy, the parents are unlikely to

strongly consider the social stresses of being disabled. Also, the parents probably have their own preconceptions about disabled life already. Especially because they are wealthy and successful individuals, they likely could not imagine living with this type of condition. Therefore, this information would only reinforce the parents' own negative conceptions of disabled life, rather than show the more positive side. Because these issues are still important, the physician could deal with them as they arise later in the life of the child. Therefore, he or she can hold off on disclosing these concerns to the parents at this early stage.

<p align="center">Positives of Living with Disability</p>

1. *What do we know about this*?

Although disability is viewed quite negatively in America, members of the disability rights movement argue that many handicapped individuals live meaningful lives despite their setbacks. They have families and friends just like any other human beings; they do not let their disability ruin their lives. Disabled individuals argue that the actuality of living with disability is usually far from the negative public perception of it. They incorporate such devices as respirators, wheelchairs, and splints into their lives if given sufficient time to adjust. Although they may resent these things at first, they eventually become routine aspects of daily life; with time, they seem no more inhibiting than a toothbrush (Gill). In summary, disabled life may bring the same joys and benefits as would any other life.

2. *What variability exists*?

Disabled individuals tend to live meaningful and happy lives in general. Some will suffer more from depressive symptoms and other problems than others. However, many individuals with Spina Bifida will get used to their handicaps and go on to live meaningful, happy lives.

3. *How should this information be presented to the parents*?

This side of disabled life could be a great source of comfort to the parents. Because they are so successful and wealthy, they may find these details surprising and encouraging. The positive information will hopefully lead the parents not to assume that their future child will be miserable, just because he or she is limited physically. For these reasons, the physician should spend some time discussing these details with the parents. Ideally, it will allow them to understand that disability does not automatically lead to unhappiness and a poor quality of life.

Part 8: Conclusion

This type of analysis, when completed fully by a physician, ensures that disclosure is best for the specific patient. It provides a thorough examination of the variables specific to a condition and models a procedure for physicians tasked with disclosing information to potential parents. In addition, this type of analysis could be applied to other situations in which patients face difficult choices in medicine. Because there are multiple, potentially favorable options, physicians must be prepared to help facilitate the decision-making process by presenting the information in the best way possible for the patient. By also comforting the patient, the physician minimizes the effect that stress and anxiety will have on the choice.

Works Cited

Anderson, Elizabeth M., and Lynda Clarke. *Disability in Adolescence*. New York: Methuen, 1982. Print.

Beauchamp, Tom L., and James F. Childress. *Principles of Biomedical Ethics*. 6th ed. New York: Oxford University Press, 2009. Print.

Gill, C. J. "Suicide intervention for people with disabilities: a lesson in inequality." *Issues in Law & Medicine* 8.1 (1992): n. p. Web. 15 October 2011.

Hetherington et al. "Functional Outcome in Young Adults with Spina Bifida and Hydrocephalus." *Childs Nervous System* 2006(22): 117–124. Web. 17 January 2012.

Hewson, J. E. "Basic Physiotherapy of Spina Bifida." *Studies in Hydrocephalus and Spina Bifida*. Ed. Gordon Stark. London: Spastics International Medical Publications, 1976. 117–118. Print.

Laurence, K. M., R. C. Evans, R. D. Weeks, M. D. Thomas, A. K. Frazer, & B. J. Tew. "The Reliability of Prediction of Outcome in Spina Bifida." *Studies in Hydrocephalus and Spina Bifida*. Ed. Gordon Stark. London: Spastics International Medical Publications, 1976. 150–156. Print.

Snow, Jeffrey H. "Executive Processes for Children With Spina Bifida." *Children's Health Care* 28(3): 241–253. Web. 17 January 2012.

Spina Bifida Association. *The Spina Bifida Association of America*, 2011. Web. 27 March 2012.

Wolraich, Mark L., Ed. *What You Should Know About Your Child With Spina Bifida*. 2nd ed. Iowa City: University of Iowa Division of Developmental Disabilities, 1985. Print.

Wright, Richard A. *The Practice of Ethics: Human Values in Health Care*. New York: McGraw-Hill, 1987. Print.

Genesis and the New Testament in
The Faerie Queene Book I & II

Christine O'Neill

Introduction

It is impossible to quantify the collective impact that the Holy Bible[1] has had on literature since its creation thousands of years ago. A slightly less ambitious task for scholars would be tracing the influence the Bible had on Edmund Spenser's *The Faerie Queene*, a monstrously long and sophisticated poem from which many Elizabethan playwrights and poets drew heavily. In much the same way the Bible is a compendium of religious narratives, records, epistles, and laws, Spenser's *The Faerie Queene* is the result of many years of work and clearly benefitted from a great number of sources. Throughout all six of *The Faerie Queene*'s books and the two Cantos of Mutabilitie, the Bible appears to be the foremost of those sources.

At first glance, Spenser's masterpiece contains just as many (if not many more) allusions to classical Greek and Roman literature as it does to the Bible—however, the powerful sway of the Bible contributes structural components in addition to textual references. It ultimately dictates the nature and tempo of the story more than do Ovid or Homer. While the New Testament seems to have influenced the overarching themes and allegories present in Spenser's *The Faerie Queene*, specific allusions are most frequently drawn from the Old Testament. These Old Testament allusions, especially from Genesis, and New Testament themes—in regards to the Gospels, to the epistles, and to Revelation—are best exemplified in the Book of Holiness and Book of Temperance.

Book I

The first virtue Spenser tackles is holiness, exhibited through the adventures of the Redcrosse Knight against Archimago, an evil sorcerer. Multiple incidents in this first book can be traced to the first book of the Bible, Genesis. The initial enemy that Redcrosse encounters is a dragon hybrid named Error, much like the original antagonist in Genesis—that is, a serpent with

---

[1] I will use all references from the Bishop's Bible (1568) since that is the edition that Spenser (1552–1599) would have used.

legs. This dragon is meant to evoke both "the serpent,[which] was suttiller then euery beast of the fielde" (Gen 3:1) and Eve, who convinced Adam to eat the forbidden fruit, for Spenser describes it as, "Halfe like a serpent horribly displaide, / But th' other halfe did womans shape retaine" (I.i.14.7–8). The literary representations of a half-woman, half-beast, reappear in Milton's *Paradise Lost* some eighty years after Spenser's publication. Spenser brilliantly ties together a direct reference to the two characters in Genesis who were responsible for the fall of man with an image from classical mythology. God and Redcrosse both triumph over their serpentine adversaries, and end up removing a part of the foes' bodies—Redcrosse "raft her hatefull head" (I.i.24.8) while God removes the snake's legs, saying: "vpon thy belly shalt thou go" (Gen 3:14).

In addition to detailing the fall of man, Genesis has a number of narratives that involve mistaken identities—Spenser borrows this idea to introduce the primary antagonist of *The Faerie Queene*, Archimago. As Redcrosse is wandering through the woods, he comes upon a man who seems to be a "godly father" (I.i.33.9) and decides to spend the night with him. In later cantos, Duessa is introduced as the virtuous Una's malicious doppelganger, and Archimago disguises himself as Redcrosse (I.ii.11.1–2, 9). There are numerous examples in Genesis when characters put on the guise of another person to achieve desired ends. A few instances include when Jacob tricks his blind father into thinking he is Esau (Gen 28:6–40), Tamar posing as the temple prostitute (Gen 38:14–15), and Joseph's pretending not to know his brothers when he is a vizier of Egypt (Gen 42:7–8). The main contrast is that while Spenser's character dissembled with evil intent, the disguised characters of the Bible do so in accordance with God's plan. Along with the instances of mistaken identities, Spenser references a handful of other Genesis stories.

Throughout the first book, Redcrosse's adventures either contain a direct reference to, or seem to mimic narratives from Genesis. In the third canto, Sansloy's duel with Archimago has an odd connection to Jacob's faceoff with the angel at Peniel. In both cases, two men engage in combat while one hides his identity—the angel masquerading as a man, and Archimago pretending to be Redcrosse. At the end of both bouts of combat, the disguises are abandoned, and the victors each choose to let the losers go (although Jacob stipulates, "I will not let thee go, except thou blesse me" [Gen 32:26]). There is a name-change in each story—Jacob is granted the name Israel, 'He who strives with God', while 'Redcrosse' becomes Archimago. Just like with the group of mistaken identity stories, Spenser reverses the moral character of the combat. In the

Bible, two good guys fight, whereas two villains go head-to-head in *The Faerie Queene*. In borrowing the narrative structure of Jacob's fight, Spenser paints Archimago's contrivances as a mockery of righteousness, and perhaps makes a statement about his trickery and duplicity (for although Jacob is technically a Biblical hero, he uses questionable means to fulfill God's plans). Another instance of matched stories occurs with the sexual violation of Una in Canto VI.

The rape of Dinah and the rape of Una are very similar in all but one respect: Dinah is actually raped, whereas Una manages to escape with her purity. Dinah's fate is sealed when a foreign prince called Shechem beholds her beauty and "toke her, & lay with her, and forced her. And his heart laye vnto Dina the daughter of Iacob, and he loued that damsell, and spake kyndly vnto her" (Gen 34:2–3). The text indicates that, although he forces himself on her, he actually loves Dinah and treats her well afterwards, quite contrary to Sansloy's motivations. Rather, "with greedy force he gan the fort assayle" (I.VI.5.3) in "wrathfull fire" and "lustfull heat" (I.vi.3.3). Perhaps because Spenser knew that Sansloy was incapable of the love that Shechem bore, Spenser chose to have a pack of satyrs scare Sansloy away before he could finish the act. The revenge of Dinah's brothers is brutal and extreme—Simeon and Levi ask the men to circumcise themselves, and while Shechem's city is recovering, the brothers "toke eyther of them his sworde, and went into the citie boldely, and slue all that was male, And slue also Hemor and Sithem his sonne with the edge of the sworde" (Gen 34:25-26). Sansloy suffers a similarly degrading fate when Sir Satyrane, son of a satyr and a woman, engages him in a battle. Although Una flees before she can witness the end of their struggle (I.vi.47.8-9), Sansloy appears again in Book II, where he duels in order to win a woman "in [whose] loosenesse [he] tooke exceeding ioy" (II.ii.37.3). Sansloy does not actually rape Una, and so he does not suffer the magnitude of ignominy that Shechem did—he is, however, obviously still a slave to his lust. Although the links to the rape of Dinah are present, Spenser takes many liberties in shaping this old story to fit his new context. Examining connections to Genesis, however, only skims the surface: Spenser, more concerned with the Christian aspects of the Bible, focused much of Book I on New Testament themes as well.

Themes expressed in the Gospels often find their way into in *The Faerie Queene*; a particularly pervasive one is the difficulty of entering the Kingdom of Heaven. Although the Markan Jesus invites everyone to worship God and receive personal salvation, he also warns that, "It is easier for a camel to go thorowe the eye of a needle, then for the riche to enter into the

kyngdome of God" (10:25). Other Gospels then use locked-door imagery to express this. While Jesus says to Peter, "And I wyll geue vnto thee, the keyes of the kingdome of heauen" (Mt 16:19), Aldeboran the Church-robber finds "The dore… was not at his call" (I.iii.16.2, 4). In contrast, when Duessa leads Redcrosse to Lucifera's hellish Palace, "Great groups of people traueild thitherward / Both day and night, of each degree and place" (I.iv.3.1–2), and because Redcrosse is so virtuous, he must face "an yron doore" (I.viii.37.3) when he tries to enter the bowells of Orgoglio's castle. Spenser seems to be making the point that as difficult as it is for a sinner to get into heaven, it is equally difficult for someone pure of heart to find his or her way into hell. This heavily parallels with the idea of 'justification through faith', a theme propounded in the New Testament's Pauline letters.

Much of Book One has to do with ecclesiastical material, as do the epistles in the New Testament. Paul is a champion of the Church, seeing it as the propagation of Christ Himself. He teaches in his epistles "howe thou oughtest to behaue thy selfe in the house of God, whiche is the Church of the lyuyng God, the pyller and grounde of trueth" (1 Tm 3:15). Spenser takes on many of the same symbols of Jesus's life that the Church did. For instance, in Ephesians, Paul famously equates putting on armor as equipping oneself for a battle "agaynst rule, agaynst power, agaynst worldly gouernours of the darknesse of this worlde, agaynst spirituall craftynesse" (Ep 6:12). The idea of being a warrior in addition to a worshiper was extremely important in the early Church, where many martyrs suffered persecution. Crawford believes Redcrosse represents this armor, not only because of the opening passages describing the knight's own armor, but Redcrosse's general steadfastness. "In keeping with the conditions of the allegory in Book I," Crawford points out, "the Red Cross Knight is a pilgrim to eternity. He represents the Church militant" (178).

But Spenser isn't done with making connections to the early Church. One possible reading of 1 Corinthians is that Paul explains transubstantiation by saying that bread and wine become the body and blood of Christ (10:16). This interpretation was developed in the medieval Church, and one with which Spenser would likely be rather familiar. Wine in Book One of *The Faerie Queene* also has supernatural properties: it is used to treat Sansfoy's wounds (I.v.17.4), and for Redcrosse and Una's wedding celebration, "they sprinckle all the posts with wine" (I.xi.38.1), in addition to frankincense, perhaps a reference to the Nativity. Even during Redcrosse's fight with the dragon at the end of the first book, so rife with parallels to Genesis,

Spenser makes larger thematic connections to the Church. As scholar Rudolf B. Gottfried points out, "in the end Redcross[e] overcomes the dragon with the strength supplied by the water of life and the tree of life, which represent the two sacraments accepted by the reformed church" (1365).

Although it takes some ideas from the epistles, the poem draws much more heavily on Revelation. The Book of Revelation is a favorite of Spenser's, most likely due to the mythological imagery of beasts, animals, spirits, celestial bodies, and other symbolism. Much as *The Faerie Queene* itself can be an allegory for elements of the Bible, this book of the Bible is an allegory for the Kingdom of Heaven, and the end of the world. Animal imagery is strong in both pieces of literature. In Revelation, "a great multitude... stoode before the throne, and before the lambe, clothed with long whyte garmentes" (7:9)—this lamb represents Jesus. Una, the purest character in the poem, carries with her "a milke white lambe... So pure and innocent" (I.i.4.9– 5.1), suggesting that Christ is with her. In Revelation, a dragon is used to embody Satan—and in case readers did not catch that, John of Patmos drives the fact home in saying, "the dragon, that olde serpent, which is the deuyll and Satanas" (20:2). Redcrosse must face two dragons in Book One, the first a half-woman hybrid called Error, and the second a more palpable embodiment of the devil. Crawford cites Carol Kaske's claim that, "his swift and final victory over Satan both on his own behalf and that of others" (176) shows Christ as the King of Heaven, the way he is presented in Revelation, and not a deity-*cum*-human, the way the Gospels represent him. The ties between Revelation and *The Faerie Queene* are quite prevalent in the first book.

Book One, out of all the books, is probably the heaviest with direct Biblical references; that does not mean, however, that the other books do not embody some of the same components. As Book Two traces Guyon's temptation and victory, readers can see the same pattern of Old Testament allusions and New Testament themes.

Book II

Spenser expresses the quality of Temperance in the second book, although editors Roche and O'Donnell warn us "to keep in mind that the word generally means 'self-restraint' and then to let Guyon's adventures qualify and define the virtue further" (1108). Since holiness is a virtue directly addressed in the Bible, and temperance is one of the Four Cardinal Virtues (which Plato formulated and the early Church adapted), this transition between books means markedly fewer

Biblical allusions. The influence, however, remains in the nature of Guyon's adventures and the themes they embody.

Just as he did with Book One, Spenser uses his protagonist (in this case, Sir Guyon) to explore narratives from Genesis, although to a much smaller extent. For instance, Braggadocchio takes Guyon's spear just as Tamar takes Judah's staff; Braggadocchio also pursues a beautiful woman named Belphoebe, not realizing that she is armed with "deadly tooles" (II.iii.37.3), much in the same way Judah beds Tamar without imagining she has equipped herself to censure him (Gen 38). Another analogous situation is that of the squire whom Furor drags around in Canto IV. He, like Joseph, was a righteous man who fell victim to circumstance (or in this case, 'Occasion') and was enslaved. Also like Joseph, he is an ultimately good character who explains to those around him "what hard mishap him brought to such distresse" (II.iv.16.8). There is even an echo of the flood story, what with all the water in Book Two. A burning Pyrocles attempts to drown himself to escape Furor's flames—he wishes to abolish the fire of fury with water, the way God purged the world of sin with water. In Arthur's battle against Maleager, the only way the Prince can defeat the demon is to divorce him from his mother, the Earth, and throw him in a lake. In Genesis, the earth is the source of original sin (bearing the fruit from the forbidden tree, then becoming part of Adam's punishment, and producing the offering that led Cain to kill Abel), while the water of the flood is a purifying agent.

Finally, temperance, the very theme of this book, could have saved Adam and Eve from being cast out of Eden. Spenser already explored the Eden episode in Book One, but here he focuses on the Fall itself, as opposed to punishing the satanic snake. The Bower of Bliss, like Eden, is both a delightful garden and a source of great temptation. Temptation and trickery are *everywhere* in this book: Belphoebe tempts Braggadocchio and Trompart, Archimago tricks the same two men, and Phaedria tempts everyone she meets. Acrasia is evidently the most successful temptress, "many errant knights hath foule fordonne" (II.i.51.4) and is seen at the end of the book asleep with a man she has successfully seduced. Guyon sees these examples and learns from them, the same way readers are supposed to learn from Eve's lament: "the serpent begyled me, and I dyd eate" (Gen 3:13). In the first canto, Guyon is fooled and believes Archimago when he tells him that Redcrosse defiled a virgin—but by the end of the book, having mastered temperance, Guyon is able to defeat Acrasia with ease and "all those pleasant bowres and Pallace braue, / Guyon broke downe" (II.xii.83.1–2). Whitaker, citing Calvin's *Institutes*, declares that

Book II both "discuss[es] Adam's fall" and "Christ's role as a mediator", nicely tying together the two testaments and giving testimony to the holistic impact of the Bible on *The Faerie Queene*.

Temptation, the opposite of temperance, is explored thoroughly in Book Two as well as in the Gospels. What Helen Cooney calls the "central issues" of Book Two, Guyon's encounters with Phaedria and Mammon anchor this poem to the Bible by exemplifying two of the evils Christ preached against: lust (Mt 5:23) and money (Lk 12:33–34). "[B]oth have their genesis in a single Biblical text, chapter 7 of Matthew's gospel, that part of the Sermon on the Mount in which Christ warns man that 'Ye cannot serve God and riches'" (Cooney, 176), which shows Spenser's focus is not only on the New Testament, but specifically on portraying Guyon as Jesus. This allegory crops up throughout all of Book Two. Guyon's episode in the Bower of Bliss parallels strongly with Jesus's temptation in the desert (Mt 4:1–11, Mk 1:12–13, Lk 4:1–14). Maurice Evans points out that "Guyon, like Christ, chooses to undergo temptation in order that his virtue may be strengthened" and furthermore notes, "when he has come through the ordeal, like Christ he receives the ministration of angels" (88). Both Guyon and Jesus take mercy on evildoers, like when Jesus pardons the woman about to be stoned and when Guyon permits Pyrochles to live after defeating him (Joh 8:7, II.v.14.1). Additionally, both suffer in a garden before facing death: Jesus in Gethsemane and Guyon in Proserpina (Mt 26:36, II.vi.53.1). But neither character—nor the rest of the human race, according to the Bible—faces his trials alone.

The "comely Palmer" who "Him als accompanyd vpon the way" (II.i.7.1–2) bears similarities to the Holy Spirit mentioned in the New Testament epistles. Whitaker astutely notes that "Red Cross must be separated from Una before he can fall into sin, just as Guyon in Book II must be separated from the Palmer or right reason before he can be tempted" (157). In this way, Guyon's Palmer seems to act as a conscience, much like the Holy Spirit—a force which is above reason, but guides it. "And greeue not the holy spirite of God, by whom ye are sealed vnto the day of redemption" (Ep 4:30) the author of Ephesians advises, reminding his recipients that even if they 'forget' their conscience (as Guyon and the Palmer were separated), they cannot abandon the Spirit and they will be held accountable on the Last Day (as evidenced by Guyon's reunion with the Palmer [II.viii.4.8]). The Holy Spirit is also a source of hope and power (Rom 15:13), two gifts that the Palmer bestows upon Guyon time and again. As Spenser so succinctly declares in the opening to the last canto, "Guyon by Palmers gouernance, / passing through perils great, /

Doth overthrow the Bowre of blisse, / and Acrasie defeat" (II.xiii). Much in the same way that Paul draws strength from the Spirit to complete his evangelical missions overseas, Guyon too depends on the Palmer to lead him through a literal sea of troubles.

While the Palmer portrays a message from epistles, the antagonistic women in Book II provide connections to Revelation. The figure of the Great Whore is a prominent aspect of Revelation, representing Babylon, or more broadly, infidelity to God. Guyon runs into a number of women who correspond to the Great Whore, among them the mermaids, Medina's sisters Elissa and Perissa, Phaedria, and Acrasia. Like the mermaids, the Great Whore "sitteth vpon many waters" (Rev 17:1). Just as the sisters delight in the battle of Sansloy and Hudibras, the woman from Revelation is "dronken with the blood of the martirs" (17:6). Phaedria is a temptress like the Great Whore, an idea that Spenser underscores by giving them similar appearances. Phaedria, whose name means 'shining one' (Roche and O'Donnell, 1118), dons herself "With gaudie girlonds, or fresh flowrets" (II.vi.7.4), while Revelation's "woman was arayed in purple and scarlet colour, and decked with golde, precious stone, and pearles" (17:4). Both are gaudy and flashy in appearance, although Phaedria takes her gaudiness from nature while the Whore adorns herself in worldly riches. Here, Spenser removes Phaedria from the Whore a little bit, if only to make Acrasia's similarities to her more striking. Acrasia seduces a man "With cup thus charmd" (II.i.55.3), indicating that she possesses "a cup of golde" (Rev 17:4) as does the Great Whore. Acrasia also has the power to transform the men she seduces into beasts, reminiscent of the "scarlet coloured beast, ful of names of blasphemie, which had seuen heades, and ten hornes" that the Great Whore sits upon (Rev 17:3). All of these women, like the Great Whore, are either avoided or defeated by the end of the book.

The Book of Temperance stresses New Testament material more heavily than that of the Old Testament, resulting in the emergence of Guyon as a Christ-figure. Spenser carefully selects Genesis stories as well as ideas from the epistles and Revelation in order to highlight the role of personal strength in overcoming temptation.

Conclusion

The Faerie Queene is a poem so obviously influenced by the Bible that many scholars no longer feel the need to comment on this connection, turning instead to less propounded feminist or new historicist readings. Biblical ties, however, run deeper than just the "holiness" theme of Book One and the arrant holiness of the Bible. Spenser takes many narratives from Genesis,

modifies them, and creates a world for Redcrosse and Guyon suggestive of that in the Torah. Often Spenser borrows the framework and not the actual events of a narrative, instead changing details of the stories for the effect of juxtaposition, or simply to fit *The Faerie Queene's* plot. Spenser draws on Genesis narratives such as the fall of man, Jacob's fight at Peniel, Judah and Tamar, and the rape of Dinah. Equally important is the New Testament, in which the knights' quests embody concepts from the epistles and Revelation. The actions of both Redcrosse and Guyon parallel those of Jesus.

Rudolf B. Gottfried and Virgil K. Whitaker, however, caution readers not see the characters of Book One simply as one-dimensional representations of other things in spite of their essentially allegorical nature. Indeed, Redcrosse's behavior is often more evocative of Old Testament characters (he is a warrior, like Joab or Joshua), than the nonviolent, rabbinical Jesus of Nazareth. Similarly, when Guyon is tempted, he does not resist the temptations as wholly as Jesus does, but rather shows signs of weakness and consequently suffers—take, for example, the Mammon episode—even though he is ultimately considered a hero (like David and Samson). It is also worth noting that Prince Arthur, who has a small yet significant role in both books, is supposed to be the 'real' Christ figure, rescuing both heroes when they are in dire straits—which must alter our perceptions of the knights. So while Redcrosse and Guyon may personify elements that are also present in Christ's life, Spenser seems open to using both men as a representations of many different Biblical subject matters.

In the movement from Book One to Book Two—holiness to temperance—readers can see a shift from emphasis on the Old Testament to emphasis on the New Testament. This makes sense, since holiness would be more involved with references to the stories that are the core of the faith—Adam and Eve's classic mistake, Jacob's heroic battle with the angel—whereas temperance or 'self-restraint' is a quality more championed in the Gospels and epistles. We should also note that Spenser alludes most heavily from the first book of the Bible (Genesis), and the last (Revelation)—perhaps in an attempt to reconcile two very different literary works with nearly three millennia between them. I hope we can agree that he manages to do so quite effectively. Edmund Spenser's *The Faerie Queene* draws on the whole spectrum of the Holy Bible, from Genesis to Revelation.

Works Cited

Cooney, Helen. "Guyon and His Palmer: Spenser's Emblem of Temperance." *The Review of English Studies.* 51.202 (2000): 169–192. <http://www.jstor.org/stable/519339>

Crawford, John W. "The Fire from Spenser's Dragon: 'The Faerie Queene,' I. XI." *The South Central Bulletin.* 30.4 (1970): 176 –178. < http://www.jstor.org/stable/3187985>

Evans, Maurice. "The Fall of Guyon." *Spenser: A Collection of Critical Essays*. Ed. Harry Berger, Jr. Englewood Cliffs, NJ: Prentice-Hall, 1968. 88–96. Print.

Gottfried, Rudolf B. "Our New Poet: Archetypal Criticism and 'The Faerie Queene'." *PMLA*. 83.5 (1968): 1362–1377. <http://www.jstor.org/stable/1261309>

Spenser, Edmund. *The Faerie Queene.* Eds. Roche, Thomas P., and O'Donnell, C. Patrick. London: Penguin Classics, 1978. Print. [2] [3]

Whitaker, Virgil K. "The Theological Structure of the Faerie Queene, Book I." *ELH*. 19.3 (1952): 151–164. < http://www.jstor.org/stable/2871935>

The online version of the Bishop's Bible that I used can be found here:
http://www.studylight.org/desk/?query=ge+1&t=bis&st=1&new=1&l=en

---

[2] My in text citations of *The Faerie Queene* read like this: (I.i.1.1)—they signify: (Book. Canto. Stanza. Line).

[3] To be clear, I used this book not only for the actual text of *The Faerie Queene,* but also for the critical material in the footnotes. Critical notes will be cited as "Roche and O'Donnell."

Higher Education in the BRIC Countries: Promoting Democratic Development

Julie Williams

Since 2001, the combined GDP of Brazil, Russia, India, and China, collectively known as the BRIC countries, has grown to total over 20 trillion dollars (CIA). The enormity of the global consequences associated with the economic, social, and political development of the BRIC countries is apparent. Whether or not these four countries and almost three billion citizens develop strong economies, stable societies, and established democratic institutions will determine the world order for the foreseeable future. Therefore, as one of the driving forces behind politics, society, and the economy, systems and institutions of higher education in the BRIC countries are worthy of study.

Leading comparative education scholar Philip Altbach explains the ability of higher education to promote democratic development in industrializing countries through autonomy, accountability, academic freedom, the academic profession, and students. Autonomous universities promote rational debate of societal issues and faculty members often produce research and provide expertise that influences public policy. He states:

> Universities are among the few institutions in modern society that have the detachment and objectivity to pose alternatives and ask difficult questions…The university, in some ways, is an institution that is profoundly subversive of intolerance, repression, and authoritarianism. Dictatorial rulers are correctly fearful of universities, and the most repressive regimes…are in a sense right when they close the universities for extended periods (Altbach 188–194).

Altbach also credits universities for training future elites to appreciate the open discourse and critical thinking that characterize a healthy democratic society. In addition to promoting democratic values on campus, higher education institutions (HEIs) disseminate vital information to their communities and serve as interpreters of international trends and knowledge (Altbach). Altbach notes that universities are of "primary importance." to the internal development of countries as well as holding a place at the "very center of economic and cultural development." in a globalized world (Altbach XVIII). In the case of the BRIC countries, internal development and by extension, the role of HEIs in society and the economy will have a direct and immediate impact on the international system.

Systems of higher education produce informed, knowledgeable citizens who are the keystone of democratic society as well as skilled workers that are the foundation of a productive economy. Their ability to promote positive democratic development, however, is accompanied by challenges that can undermine liberalism and cause existing inequalities and deprivations to become more entrenched in society. The negative effects of higher education systems that contain aspects of inequality, elitism, corruption, and other hindrances to access and academic excellence can be understood using Amartya Sen's capabilities approach.

The capabilities approach is a paradigm that explains poverty as capability deprivation rather than simply a lack of income. The solution to capability deprivation is a social commitment to development through the expansion of freedom. Sen argues, "Development consists of the removal of various types of unfreedoms that leave people with little choice and little opportunity of exercising their reasoned agency. The removal of substantial freedoms, it is argued here, is constitutive of development" (xvii). Rich and poor countries face issues associated with persistent poverty, hunger, and restrictions on political freedom, the oppression of women, unsustainable development, and the maintenance of civil liberties. The tenet of the capabilities approach is that the freedom of individual agency is instrumental in solving these problems. Individual agency, however, is "inescapably qualified and constrained" by the social and economic situation of any given individual. Sen considers both the processes that allow individual agency and the opportunities that individuals have depending on their circumstances (Sen 17). Deprivation of the five instrumental freedoms (political freedoms, economic facilities, social opportunities, transparency guarantees, and protective security) can be a result of inadequate processes or inadequate opportunities (17).

Education, as a social opportunity, complements economic development and political freedoms (Sen 18). For example, a more educated, highly skilled workforce can increase economic productivity and in turn can result in more investment in higher education. Similarly, HEIs can be affected by public policy, but they also affect policy by producing a highly literate, critical voting population. Sen's approach can be used to analyze the challenges of access and equity faced by HEIs in the BRIC countries. The systems in all four states demonstrate evidence of systemic and often institutionalized discriminatory policies that limit the ability for individuals of disadvantaged social strata, regions, ethnicities, cultures, linguistic backgrounds, and income levels to obtain an equal opportunity to enroll at a quality university.

On a more systemic level, Altbach applies the idea of center-periphery to describe the relationship between HEIs in developing countries and HEIs in well-established, industrialized countries (Altbach 20). This concept separates the few central institutions, the knowledge producers and international research leaders, from the majority periphery institutions that distribute knowledge acquired from the central institutions. Altbach categorizes the HEI systems of Brazil, Russia, India, and China as peripheral internationally, but that does not mean that the universities within those systems are peripheral within their respective societies. In fact, the opposite seems to be true:

> Third World universities are…probably more important to their societies than the major international universities like Harvard or Oxford are to theirs. They produce the highly trained elites necessary for the operation of the modern state, are very often at the political vortex of their societies, produce cultural commentary and criticism, and, in some nations, make important contributions to defining newly established political entities (22).

Altbach argues that as a result of neocolonialism, poor access to adequate technology and educational resources, and the dominance of English as the language of academia makes the transition from the systemic periphery to the central nearly impossible (22). The BRIC countries, however, have no choice but to create world-class higher education systems to maintain their economic growth and sustain a productive and politically stable society. By using Sen's capability theory to analyze challenges of access and equity and Altbach's work on universities as tools for development to explore university autonomy, sources of funding, and state building capabilities, it is possible to assess the role of HEIs in the globally relevant BRIC countries.

<div align="center">

Expansion and Privatization

</div>

Rapid Growth

The first major and dramatic similarity between institutions of higher education in the BRICs is the rapid expansion experienced by the systems starting in the 1990s. For example, Brazil's total undergraduate population expanded from about 1.5 million in 1992 to over 3.8 million in 2003 and the total number of higher education institutions nearly doubled between 1998 and 2008 (McCowan 580). The number of college students per 10,000 individuals in Russia rose from 178 to 517 between 1993 and 2007 (Dias, et al., 23). In the same time period, the number of Russian state higher education institutions rose from 548 to 682 and non-state institutions grew from 78 to 666, with even more growth in branches and satellite campuses

(Salnikov, Burukhin 72). Finally, in 1980, three percent of university-aged Chinese citizens were enrolled in a HEI; by 2009, the number had jumped to 24 percent (Lei, et al., 44). This growth fosters a more educated populace and generates more skilled workers to contribute to the economy, two positive factors crucial to the effect of higher education on democratic development. However, the rapid rate at which higher education has expanded in these growing countries has compromised quality and equity in all four countries.

Additionally, the presence of an expanding group of educated students has, in some cases, prompted undemocratic action by political leaders who are threatened by the political activism associated with college campuses. Prime Minister Vladimir Putin and his political party, United Russia, have a reputation for undemocratic practices. Aside from a recent intervention in the journalism program at Moscow State (Putin funded a second school with the explicit purpose of training journalists to support the government), there is not much documented evidence of the PM or his party's intervention (Nemtsova 56). Despite a lack of definite proof, PM Putin's tendency to brutally stifle protests and political challengers does not bode well for student activists.

In China, the increase in enrollment has forced the People's Republic China to devise a method for governing the universities that balances the autonomy needed for academic advancement with the control desired by the Communist Party of China. The current compromise means that each public institution has an academic administration headed by a president and a separate Communist Party administration led by a party secretary (Altbach 13). Altbach calls this lack of autonomy "dysfunctional" and states that Chinese universities are subjected to "extraordinary bureaucratic controls" (16). When trying to determine the role of higher education in the democratic development of a country, it is important to understand the relevance of the institution's internal politics. China is the only BRIC country that is not an official democracy, so it is relatively unsurprising that the CCP is heavily involved in the governance of public universities. That does not make the authoritarian government's extensive involvement acceptable, however. The fact that a CCP member oversees all research, and even determines what is and is not acceptable to research, threatens the academic integrity of the entire country. Ideally, HEIs promote democracy by encouraging critical thinking, dialogue, and analysis that challenges the status quo and advocates for social and political progress. If each

Chinese campus has its own CCP censor, universities cannot carry out that critical function and therefore cannot serve as proponents of democracy.

"Budget for Non-Development"

One of the most obvious and prominent effects of rapid expansion is the inability of BRIC governments to keep up with the new institutions and students financially. In Brazil, "standards in many of the public universities have dropped in recent years due to reductions in their budgets" (McCowan 584). The public sector is often seen as "costly and inefficient, and unable to provide the diversity of provision and the responsiveness to consumers necessary in the contemporary context" (586). In 2003, public universities in Brazil were spending $13,500 per student per year, an expenditure that was "unlikely to be sustained by a government who had adopted the belief that the State's responsibility was to regulate and not provide higher education" (McCowan 456).

Similarly, between 1996 and 2006, the Russian Federation consistently allocated 11 to 13 percent of its budget to higher education every year, which was problematic once one considers the immense increase in enrollment during that same time period (Arapov 48). Between 2005 and 2007, the money spent by the government on higher education increased by 105 percent, but in 2006 that amount only equaled 3.87 percent of the GDP, which is less than the 4.8 percent allocated in 1997 and well under the industrialized country average of five to six percent (Khalin 30–44). Some scholars refer to this unfortunate economic situation as a "crisis," citing unprecedented budget cuts and a dramatic departure from the Soviet policy of generous spending in the sphere of education (14 percent of the budget in 1950) (Kolisnikov 35–48). Even O. Smolin, the vice chair of the Committee on Education and Science in the State Duma, called the 2004 budget for higher education a "budget for nondevelopment" (Plaksii 6–22).

In India there are only 350 public HEIs in a country of over a billion individuals. Therefore, it makes sense that only seven to ten percent of young adults in India attend college; there is no room for anyone other than the elite. The government of Prime Minister Singh has since tried to remedy this deficiency by building more Indian Institutes of Technology and Management, but the financial viability of his goals is yet to be seen (Bagla).

The Strengths and Weaknesses of the Private Sector

Brazil, Russia, India, and China each responded to the financial strains of a growing higher education system by allowing the private sector to pick up the slack of overburdened public systems. Each of the BRIC countries has employed unique mechanisms to allow the private sector into the sphere of higher education, and each system approaches the growth and regulation of private colleges and universities slightly differently. The privatization of tertiary education in all three countries has been strikingly similar. There was a 133 percent increase in the Brazilian private sector between 1998 and 2004; the expansion of private higher education was so intense that in the nine months between November 2001 and July 2002 an average of 2.5 private HEIs were opened every day (McCowan 457). In Russia, between the 1993–94 school year and the 1998–99 school year, tuition-paying students caused 82 percent of the growth in the higher education sector. By 2002, 50 percent of students were paying tuition, and that number has grown every year since so that at least two thirds of all spending on higher education comes directly from Russian families (Arapov 7–27). In 1980 there were no private HEIs in China; by 1999 there were 43, and in 2006 there were 278 private degree-granting institutions located predominately in the economically robust Eastern regions of the country (FienLiang, Morgan 28). One in six undergraduates currently enrolled in an institution attend a nonpublic university (Yu Ertl 43).

Unfortunately, private colleges and universities are often over-commercialized, of questionable quality, and too expensive to be truly accessible to the sectors of the population that need a higher education the most. Wrana Maria Panizzi characterizes the growth of Brazil's private higher education sector as "the kind of growth that is often accompanied by a decline in the quality of educational provision" (Panizzi 45). A newer university, Pitagoras, is affiliated with the U.S. company Apollo, which is responsible for the University of Phoenix, a for-profit college with over 400,000 students in the United States. These HEIs are more like business ventures than universities because they rely heavily on advertising and branding to promote their thousands of campuses, which are usually located in strip malls and even theme parks (McCowan 458). These institutions are often "launched quickly and indiscriminately, without necessarily complying with regulatory processes and quality assessments" (Dias 17). Not all private institutions are illegitimate, however. An example of a successful private college in India is NIIT University, born from the NIIT Corporation led by millionaires Rajenda Pawar and Vijay Thadani. The founders claim that they started the institution because it was obvious to them that

the state "does not have enough money to take on the myriad of challenges facing the country's education system" (Neelakantan 56).

The spectrum of private institutions that ranges from highly commercialized "diploma factories" to more quality learning and student-based schools colleges exists in all four BRIC countries. Complicated accreditation processes and financial strain make students of all four countries vulnerable to illegitimate unaccredited institutions. For example, one Russian university rector goes as far as to say, "a graduate with a 'nonstate' diploma has fewer rights" than a graduate of a public school. It is also claimed that graduates from non-state institutions have a lower level of knowledge, reasoning skills, maturity, and personal and professional expectations than their state school counterparts (Suspitsin, Suspitsyna 62–80). One critic of private education even said that Russian private schools that charge tuition are "turning into hotbeds of deceit, fraud, and financial skullduggery" (Kolesnikov 42).

Similarly, in China, Li FienLiang and W. John Morgan emphasize that, "There is a consensus in China that investment in private higher education should get reasonable economic profits or returns" (29). Student fees and tuition comprise over 80 percent of the school's total income and the institutions are often subsidiaries of major corporations (for example, Jili University is operated by Jili Auto Company) (29). In addition to this extreme level of educational commercialization, FienLiang and Morgan argue that because of the lower quality of private HEIs, "private institutions can only begin to recruit new students in the wake of ordinary public institutions, let alone in competition with key universities," and that private higher education should be viewed as a "necessary supplement to public provision" rather than a "complementary twin" (29–30). Finally, the authors criticize private HEIs for employing more part-time teachers who are less committed to the institution (no private schools can conduct official research in the PRC) and whose age ranges are concentrated in the "less than 30" and "older than 60" areas, implying that the instructors, "lack experience at the one end of the scale and energy and commitment on the other end of the scale" (30).

FengLiang and Morgan concluded that most lower-income families pay for education in an attempt to increase prospects for employment after graduation, but, ironically, the expansion of the private sector has resulted in college graduates having an increasingly difficult time finding adequate employment. This holds especially true for private HEI graduates in China; 43 percent of them ultimately work in non-urban (and therefore lower paying and less prestigious)

areas compared to 10 percent of their public peers, and the average starting wage for a public school graduate is significantly higher than for a graduate from a private school. Students from a disadvantaged background as well as a private institution have the lowest chances of obtaining employment in an urban area, which is exceptionally detrimental to the students and families who are financially struggling to earn a degree (FengLiang 32).

<div align="center">

Expansion Without Equity

</div>

Tristan McCowan summarizes the lack of relationship between expansion and increased equity in higher education:

> The existence of private universities, with their less competitive entrance exams, flexible hours, and, in some cases, location in areas outside the metropolitan centres, have meant that many Brazilians have obtained a university diploma who would not have been able to otherwise. Private universities are, therefore, fulfilling a role in Brazilian society, and there is clearly a strong demand for them from 'consumers'. Nevertheless, it will be argued that the growth of this sector is not in fact contributing to equity in the education system, particularly in the long run (460).

McCowan was specifically referencing Brazil, but there is evidence in all of the BRICs that the expansion of tertiary education has not necessarily resulted in increased equity and accessibility for financially disadvantaged, rural, and minority students. For example, the Organisation for Economic Co-operation and Development (OECD) created a profile that described a typical Brazilian student as male, Caucasian, a member of a nuclear family with parents who have completed at least secondary school, and a member of a household that earns over ten times minimum wage (13). Additionally, "71 percent of students are from the top quintile of family income, there is low representation of the African Brazilian populations, and universities are concentrated mainly in the wealthier South-Eastern and Southern regions" (McCowan 456). McCowan elaborates:

> Brazil's education system displays the extreme inequality that characterizes the country as a whole. While a child of an upper-middle class family is assured an education comparable to that of any developed country, the poor can expect little more than a few years at an under-resourced primary school…Less than half of those at secondary level have survived their studies without repeating a year…education opportunity on a national scale is deeply undemocratic (456).

This combination of race, income, and region is characteristic of inequality, inaccessibility, and capability deprivation in the BRIC countries.

The Cost of Education

The first limiting factor for equal access to quality higher education is the high cost of tuition. A study conducted in 2002 determined that only 11.81 percent of Brazilian 18–24 year olds could afford the majority of courses (McCowan 460). The fact that costs for courses vary depending on subject area further complicates the issue. Subjects that result in a degree with high-earning potential, like dentistry, medicine, and engineering are significantly more expensive than lower earning subjects like humanities, education and social work (460). Additionally, the cost of a degree from a private HEI depends on the quality of the institution. Therefore:

> The emergence of low-cost courses in the last 5 years is a deliberate strategy by education companies to open the HE market up to the lower-middle class, and at first sight seems a positive means of widening access. However, it will ultimately serve to reproduce inequalities by confining students of poorer families to courses and institutions providing diplomas with less value in the employment market (461).

According to M.V. Arapov, in Russia, "There are regions in the country where the total worth of goods and services produced by an individual is about half the resources that he will have to pay to obtain a higher education" (461). In 2002, 42.4 percent of young people in Russia claimed that going to college would force them to "give up everything else" and 44.8 percent of them said that obtaining an education on a tuition basis "is simply not possible" (461).

Race and Ethnicity

A second major limiting factor in achieving democratic development through higher education is the inability for BRIC countries to accommodate for racial and ethnic diversity in the admission process. Racial diversity has the potential to strengthen a democracy, but it is often contentious because there is an overlap with race and socioeconomics; for example, Afro-Brazilians constitute an economic underclass. Descendants of Africans make up almost half of the Brazilian population (the only country with a higher black population is Nigeria), but they only account for 20 percent of university students and two percent of graduates (Lloyd). Additionally, one in four black Brazilians are illiterate compared to one in 10 non-blacks and Afro-Brazilians earn on average half of the salary of whites (Lloyd). Private institutions in Brazil are not mandated by the government to help correct this problem of inequality. In the public sector, the State University of Rio de Janeiro was the first institution to reserve 40 percent of its seats for Afro-Brazilians and 10 percent for students from public secondary schools (McCowan 591). Since 2000, the federal government has attempted to enact policies that require public HEIs

to set aside half of their seats for Afro-Brazilians, indigenous peoples, and students from public schools in proportions representative of their region's demographics (591).

Quota systems are controversial and will be discussed in detail in a subsequent section, but Brazil is similar to India in that the higher education quota system has incited racial violence and intense controversy—such as dorm looting and the use of racial epithets—on university campuses. One reason for the extreme reactions stems from the quota selection process; applicants are deemed "black" or "not black" based on a picture. This system has resulted in an inconsistent selection process for students; in one case, one identical twin was admitted as "black" to the University of Brasilia and the other was denied as "not black" (Ash).

In China, 91.59 percent of the population identify as members of the Han ethnic group, the other 8.41 percent constitute 55 distinct minority groups (Zhu 14). The number of ethnic minority students in China's higher education system increased from 4,500 in 1952 to 950,000 in 2005, but that is due in large part to a proportional expansion of exclusively ethnic majority colleges, not an overall growth of ethnic equality in the system (15). These ethnic universities tend to be less prestigious than their Han majority counterparts and also generally focus on the liberal arts rather than hard sciences, technology, and engineering, which limits the minority students' employment opportunities upon graduation (17).

The fact that there are at least 128 distinct languages recognized by the Chinese government complicates the process of integrating ethnic minorities into Han-majority schools (Tsung, Clarke 58). Bilingual education at the primary and secondary level was institutionalized through policy in the 1980s and has since developed into a two-tiered system where the word for the education of the Han Chinese (*zhenggui jiaoyu)* means "regular education" and "national education," *minzu jiaoyu,* is unique enough to require a separate word (58).

Lind Tsung and Matthew Clarke conducted extensive interviews in the Xinjiang Uyghur Autonomous Region (XUAR) to learn more about the effect of linguistic identity issues on higher education. Responses varied, but they ultimately concluded that the decision to attend primary school taught in Han Chinese or in Uyghur—the dominant ethnic and linguistic minority in the region—is one of the most impactful choices in one's life (62). Uyghur-speaking students have fewer employment opportunities upon graduation and are generally considered inferior to their Han-speaking counterparts. Sen addresses this issue and suggests that the actual linguistic

decision made by groups like the Uyghurs is less important than the fact that they have the ability to make that choice in the first place:

> The pointer to any real conflict between the preservation of tradition and the advantages of modernity calls for a participatory resolution, not for a unilateral rejection of modernity in favor of tradition by political rulers, or religious authorities, or anthropological admirers of the legacy of the past. The question is not only not closed, it must be wide open for people in the society to address and join in deciding. An attempt to choke off participatory freedom on grounds of traditional values…simply misses the issue of legitimacy and the need for the people affected to participate in deciding what they want and what they have reason to accept (Sen 32).

Regional Disparities

Just as pertinent as racial and ethnic disparity is regional inequality. Democratic development requires an educated population that has access to quality schools regardless of where an individual was born. The BRICs provide urban students with the best flagship universities, leaving the majority of citizens in rural areas without access to the best tertiary education the country has to offer. In China, students from rural regions are underrepresented in all types of HEIs; for example, 4.5 percent of Beijing University's 2000 students came from rural areas in 2004 when rural residents as a whole comprise 52 percent of the general population (Yang 195).

A large part of the problem stems from quotas established by local educational authorities that mandate a percentage of students local to that city, municipality, or province be admitted to a particular HEI (often with lowered admission requirements). As a result of these quotas, prestigious public universities within major cities in eastern provinces have a mandate to enroll mostly affluent students from an urban background (Yu 49). The most significant challenge faced by rural college students, however, is related to their lack of resources in and out of primary and secondary school. In 2002, schools in Shanghai spent five times more per pupil than those in the rural Henan province (Yang 197). Fiscal decentralization and a cultural emphasis on self-reliance in China has resulted in rich urban schools constructing indoor swimming pools while poor rural schools struggle with approximately a third of the textbooks available to wealthier schools (as well as a lack of desks, chalk, and other essential teaching materials) (197). Also, urban students are more likely to have wealthy, well-educated parents, and a child whose mother has completed at least two years of college is four times more likely to enroll in an HEI than a child whose mother has never pursued a degree (Yang 196).

Regional disparities and resource deficiencies are not unique to China. A primary cause of the perceived illegitimacy of the Russian higher education system is a lack of adequate assets and infrastructure; underfunding has caused state and non-state institutions alike to suffer from a shortage of modern supplies, academic rigor, and human capital. Cramped dorm rooms, outdated computers, and aging lab equipment are among the physical problems faced by universities in Russia today (Kolesnikov 40). In 2006, only 100 Russian Foundation for Fundamental Research grants were awarded to schools outside of the two capital cities, which indicates a serious lack of emphasis on quality research in major universities in the provinces (Arapov 22).

Standardized Tests as the Solution and the Problem

One of the ways in which differences in race, location, and socioeconomic status become salient within systems of higher education is through standardized testing for admission to HEIs. Much of the stress associated with standardized testing comes from the competitive nature of the exams; in Brazil there are 8.4 candidates that take the test for every one place in a public university, and 1.5 applicants for every place in a private institution (McCowan 585). In India, the competition is even more intense with 320,000 applicants taking the All India-JEE entrance exam in April 2008 for 7,000 seats (Neelakantan 54).

The situation becomes more complicated when one considers the various mechanisms BRIC governments have constructed to accommodate linguistic, ethnic, and racial minorities as well as the ways individuals can use their socioeconomic status to manipulate the competitiveness of the exams. In Brazil, in order to have a competitive advantage on the exams for tuition-free public HEIs, wealthy parents often enroll their children in *pre-vestibulares*, or preparatory courses (McCowan 457). These courses are expensive and therefore almost completely inaccessible to students outside of Brazil's upper economic echelons. McCowan describes the situation:

> The cruel irony of Brazilian higher education, therefore, is that the majority of the free higher education places are filled by students from wealthy backgrounds who have been able to afford private primary and secondary schooling, and a *pre-vestibular* course (457).

In Russia, the United State Examination (USE), a standardized admissions test, was first tested in certain regions in 2001 and was administered for the first time nationally in 2009 (Nemtsova A28–A29). This test is an improvement for both efficiency and egalitarianism because previously, if a student in Siberia wanted to attend a university in Moscow, they had to

travel to the capital to take a specific entrance exam. The USE also addresses the needs of 153 categories of disabled or underprivileged students (handicapped students, Chernobyl victims, etc.); universities are mandated to accept lower scores from those specified individuals (A28–A29). Several of Russia's premier universities, Moscow State being the most notable example, did not initially accept USE scores and are still struggling to adjust their admissions process to accommodate this new test.

China's national entrance exam, the College Entrance Examination (CEE), is especially controversial among advocates of equitable access to higher education. Who makes the test, what is on the test, how the test is administered, and who even has to take the test are all highly politicized questions in China. The exam was instituted in 1977 and since its beginning, scores have been highly relative to a student's home province and to the universities to which the student is applying. The score cutoffs became so random and decentralized that by 2002 Shanghai and Beijing, two of China's wealthiest provinces, created their own province-specific test (Wang 22). These new tests were immediately criticized for increasing the possibility of cheating and corruption because the likelihood of the academics writing the tests being the same academics administering and teaching for the tests increased drastically (23).

Houxiong Wang conducted a subject functional difference analysis of the regional 2007 entrance exams and found them to be "prejudiced against or unfair to rural students" (Wang 32). The source of this bias was often language usage. In addition to concerns about taking a test in a non-native language, many exams contained words such as "mortgage," "chain store," "hamburger," and "public reserve funds" that caused serious comprehension difficulties for students accustomed to rural terminology (32). More obvious than subtle linguistic bias is China's "grade credit policy" that blatantly awards CEE "credits" to students with powerful parents, wealth, or some other athletic or academic talent that can usually only be developed with the help of a rich family (29). Wang, and most other scholars who have researched this subject, ultimately concluded that despite some glaring inequalities, "using objective paper tests are fairer than not using them" (33). Therefore, other means to improve fair access to higher education must be explored.

Possibilities for Improvement

The work of both Sen and Altbach is based on the possibilities for improving human capability and the potential for universities to be used as tools for democratic state building.

Therefore, evidence that the BRICs have identified weaknesses and are taking reasonable steps to combat them is just as important as evidence of weaknesses themselves. Brazil and Russia are both in the process of implementing national programs for HEI improvement. All four BRICs have attempted to develop a financial aid system for economically disadvantaged students. All four BRICs have also developed some sort of quota system to include historically excluded and underprivileged minority groups in tertiary education.

National Improvement Programs and Financial Aid

The government of Luiz Inacia Lula da Silva (Lula) responded to weaknesses in Brazilian HEIs in 2004 with the Universidade para Todos (University for All Programme, or PROUNI). This system of reforms used tax reductions to encourage private HEIs to provide vacant seats to disadvantaged and underrepresented students for free. Non-profit universities can designate 20 percent of their vacancies, whereas for-profit colleges can only set aside 10 percent (Akoojee, Nkomo 122). The 2004 reforms also modified a pre-existing student loan system, but poor families have been reluctant to engage in the system and the rates of non-payment are over 20 percent (589). Nonetheless, 277,000 students have benefited from the program (called FIES, the Programa de Financiamento Estudantil); since 2004 it has effectively provided previously unavailable tertiary education opportunities to a portion of the low-income population.

In Russia, Article 42 of the 1992 Law on Education introduced the concept of federal educational loans and grants, but it did not include any information about creating viable infrastructure to enact any sort of organized program (Bain 67). The Russian law "On Education", last amended in 2008, did suggest an experimental voucher system called GIFO (Governmental Individual Financial Obligation) that funded universities based on per-capita ratios, but it failed to truly be implemented. According to O. Bain, the Russian government continues to "nibble at the problem rather than solve it" (67).

Similarly, the Chinese government launched a small financial aid program in 1997, and in the past decade and a half it has grown to provide aid to over 29 million students (Shen 147). The system provides aid through scholarships, grants, work-study, tuition waivers, and, most prevalently, through student loans. Hong Shen's in-depth analysis of the aid framework in China revealed that the program is off to a good start, but for many of the country's poorest students there is simply not enough aid to make higher education a possibility (164).

A potential solution to Russia's educational legitimacy issues is the Bologna Process, the European Union's solution to higher education. When Russia became a participant in this process in 2003, it agreed to adopt a multilevel (baccalaureate, master's, doctorate) system of higher education, work to enable student and instructor mobility throughout participating countries, implement joint educational programs, and the use of the European transcript. Russia also began using the European Credit Transfer System (Dobren'kova 42–51). Many of the more prestigious universities throughout Europe have refused to adapt to this system, but for Russia, the Bologna Process will serve as an effective model for development in the future. The major arguments against Russia's entry into this agreement revolve around the perceived need for a uniquely Russian higher education system, but G.A. Prazdnikov is able to reply to that concern:

> The Bologna Process does not pose a threat to our own system of education in Russia (or to any other national system of education) of destroying or wiping out cultural traditions; it does not dictate lockstep standards of education. Quite the contrary: a diversity of cultures, national social systems, and education syllabi is viewed as a part of the common European intellectual treasure (5–16).

Controversial Quotas

Although all of the BRICs have implemented some sort of national quota system within their admission process, India's highly contested reservation system provides an excellent case study. In India, only the ancient caste system itself is more complicated than the politics behind the issue of reservations. Essentially, 15 percent of seats and faculty positions at public universities are reserved for members of the Scheduled Castes (SCs), 7.5 percent are reserved for Scheduled Tribes (STs), and 27 percent are reserved for non-creamy layer members of the Other Backward Castes/Classes (OBCs) (Neelakantan).

The reservation process technically begins when students sit for the All India-JEE. The general applicants take the test on white paper; OBC/SC/ST and disabled students take the test on colored paper. Then, the raw scores of the applicants are manipulated so that OBC/SC/ST score cut offs are up to 50 percent lower than the limits for general applicants (Chhapia). If enough of those seats are not filled, the students who could not pass the test even with the relaxed limits enroll in a rigorous one-year prep course that provides a summary of all the subjects covered by the tests. They are then admitted, without retaking the exam, to a public HEI. Even after that, Chief Justice Balakrishnan ruled against IIT Delhi's attempt to expel five SC

students because, "these socially and economically backward categories are to be taken care of at every stage" (Mahapatra).

The higher caste groups that oppose the current quota system claim that the situation is unfair, unnecessary, and detrimental to the quality of national institutions. They have a valid argument; wealth and educational experience vary widely among the OBC/SC/STs. The "creamy layer," the wealthiest OBC members, are a particularly contentious group because they are seen as too well off and therefore unworthy of reservations. Faculty and students make their opinions clear; strikes, self-immolation, arson, and riots demonstrate exactly how desperate Indians are for a seat in an IIT (Indian Institute of Technology) or IIM (Indian Institute of Management).

This inflammatory system is not an effective solution to India's problems of inequity and inaccessibility. A study by S. Desai and V. Kulkarni in 2008 showed that the dalit (the lowest members of the now illegal caste system) graduation rate is actually decreasing as more quotas are implemented and that a SCs chances at employment are, for the most part, still poor (Desai, Kulkarni 45). In 2007, the National Knowledge Commission suggested a deprivation index that takes into account income, gender, region, and place of residence before determining admission to an institution (Neelakantan). This shift from restrictive group policies would take the emphasis off of the social construction of caste and place it on to the individuals who are truly at the greatest disadvantage, regardless of sociopolitical label.

Possibly the most important indication that quota systems in the BRICs are ineffective is the fact that many historically disadvantaged groups cannot produce enough qualified students to fill reserved seats even with lowered admission requirements. In 2009, 1100 reserved seats were not filled because that many SC/ST/OBC students could not pass the JEE even with a preparatory course (Chhapia). In Brazil, the PROUNI system is based on the fact that there are empty seats in universities because they are inaccessible to students due to price or lack of preparation. There is hope for students who take advantage of the quota system, however. FengLiang and Morgan did find that the employment prospects and starting wages of underprivileged students who graduated from a Chinese public institution were not significantly different than those of wealthier students (35). Similarly, low-income students in Brazilian public institutions performed better on a national assessment than their wealthier counterparts in 2000 (McCowan 463). This means that a quality higher education does have the ability to improve the lives of disadvantaged students.

Conclusion

The BRIC countries are distinctly different states economically, socially, and politically. Their only obvious commonalities are the reason they were somewhat haphazardly grouped together in the first place—large population and economic potential. Therefore, the dramatic extent to which their higher education systems are similar is surprising. The systems are not identical, but the overall themes of expansion, privatization, and the challenge of increasing access and equity to disadvantaged students are obvious in each country. These themes have been covered extensively in the previous sections, but their relevance to democratic development and human capability has yet to be explicitly addressed.

The most significant obstacle standing between tertiary education in the BRIC countries and democratic development are the problems of inaccessibility and inequity. Every country with a university system faces the challenges of balancing public funds, private contributions, and educational quality in order to provide as many citizens with a higher education as possible. No system will be perfect, but McCowan lists two criteria for higher education systems to strive for to be considered democratic:

1. There must be sufficient places so that all members of society who so desire, and who have a minimum level of preparation, can participate in higher education.

2. Individuals must have a fair opportunity of obtaining a place in the institution of their choice (582).

UNESCO's World Conference on Higher Education in 1998 gives another definition of an equitable and accessible system:

> Admission to higher education should be based on the merit, capacity, efforts, perseverance and devotion, shown by those seeking access to it, and can take place in a lifelong scheme at any time, with due recognition of previously acquired skills. As a consequence, no discrimination can be accepted in granting access to higher education on grounds of race, gender, language or religion, or economic, cultural or social distinctions, or physical disabilities (UNESCO).

Finally, the OECD endorses the following definition of equity in education:

> Equity in education has two dimensions. The first is fairness, which implies ensuring that personal and social circumstances…should not be an obstacle to achieving education potential. The second is inclusion, which implies ensuring a basic minimum standard of education for all (Dias 4).

All three of these definitions argue that in order to be considered equitable and democratic, higher education has to be a somewhat realistic option for everyone. Expensive tuition, challenging tests, and cultural barriers are not inherently undemocratic as long as they are not so overwhelming that they make obtaining a higher education impossible for one segment of the population. From the perspective of the capabilities framework, individuals should not be actively or passively excluded from the ability to pursue a diploma in a field of their choosing.

Given those criteria, there is evidence of all four BRICs actively promoting equity of access. All four countries have quota or reservation systems for historically disadvantaged groups as well as a form of financial aid system to offset the rising cost of tuition. The quota systems are far from effective and inherently discriminatory, but no developed country has found a better solution to assist groups that have systematically been excluded from beneficial facets of society in the past. A. Gupta argues that fair and effective affirmative action policies are unique to democratic polities because they provide a legitimate and justified way for the elite and the underprivileged to positively interact (5). Similarly, national standardized tests are generally considered unfair because they favor students from rich cultural urban centers who have had the benefit of a wealthy upbringing and excellent primary schooling, but no country has an admission system that is considered more fair than a standardized written test.

The major unique challenges facing the BRICs are the rapid rates of higher education expansion and the extreme levels of income inequality. The private education sector grew so quickly in the 1990s that the public sector has not been able to truly catch up and the systems as a whole have not developed an adequate infrastructure to provide equal access to all. For example, most of the BRIC financial aid systems are severely underdeveloped. Additionally, the inability of HEIs in India and Brazil to fill vacant seats indicates a deep problem with not only planning on part of the HEIs, but also seriously inadequate primary and secondary schools. China has the most developed system of loans, grants, and scholarships, but the lack of autonomy from the CCP places a major limit on the university's essential function of academic independence and research. Russia's rampant corruption demonstrates underdevelopment in the area of the teaching profession and institutionalized university integrity.

UNESCO's declaration states that access for some specific disadvantaged groups must be "actively facilitated" (McCowan 582). As tertiary education systems become more established in the future, public and private parties can do more to actively facilitate access and equity. In all

four countries, national exam preparation courses could be offered at discount rates for students who cannot afford the currently exorbitant course prices. By actively promoting democratic practices in the admissions and enrollment process, BRIC governments are actively promoting democratic university development and, by extension, democratic development as a whole.

> McCowan argues:

> Reform of higher education is unlikely to be successful in isolation from primary and secondary levels, and it is difficult for any university system to correct inequalities developed through previous years of schooling. At the same time, higher education cannot absolve itself from all responsibility, placing the onus of inequities on the previous levels (581).

This point highlights how it is often easiest to dismiss higher education as simply the tip of the educational iceberg, totally dependent on the primary and secondary levels below it. That perspective ignores the abilities of the university as explained by Altbach to become a state-building entity through research and to become a democratizing agent through the facilitation of critical thinking, academic autonomy, and a knowledgeable voting base.

The criteria for equitable and democratic higher education described by McCowan, UNESCO, and the OECD can all be directly related to Sen's capability theory. Deprivations exist in the higher education systems of Brazil, Russia, India, and China—an inevitable fact. In order to promote democratic development, these deprivations must be addressed through increasing the capability of individuals to obtain a higher education. This means treating basic education as a civil right to ensure that there are enough qualified students to fill reserved seats in prestigious universities. It means guaranteeing a fair admissions process that does not give undue benefit to those of specific races, religions, or social classes. It means regulating universities to ensure quality academic programs that are worthy of time and money. Finally, it means implementing a sound financial aid system to make an investment in higher education a reasonable process for everyone involved. All of these actions will increase the capabilities of the billions of current and future students in the BRIC countries to become knowledgeable, productive, and democratically motivated citizens.

Works Cited

Akoojee, Salim, and Mokubung Nkomo. "Widening equity and retaining efficiency: Considerations from the IBSA southern coalface." *International Journal of Educational Development* 31 (2011): 118–125. Web.

Altbach, Philip G. *Comparative Higher Education.* Westport, Connecticut: Ablex Publishing, 1998. Print.

Altbach, Philip G. "One-third of the globe: The future of higher education in China and India." *Prospects* 39 (2009): 11–31. Web.

Arapov, M.V. "The Higher Education Boom in Russia: Scale, Causes, and Consequences." *Russian Education & Society* 48, no. 1 (2006): 7–27. Web.

Ash, Timothy Garton. "In Brazil I glimpsed a possible future in which there is only one race." *The Guardian*, July 11, 2007, http://www.guardian.co.uk/commentisfree/2007/jul/12/comment.brazil (accessed on March 25, 2012).

Bagla, P. "India's education bonanza instills hope—and concern." *Science* 320 (2008). www.sciencemag.org.

Bain, O. "The Cost of Higher Education to Students and Parents in Russia: Tuition Policy Issues." *Peabody Journal of Education* 76 (2001): 57–80. Print.

Central Intelligence Agency. "The World Factbook: Brazil." Central Intelligence Agency. https://www.cia.gov/library/publications/the-world-factbook/rankorder/2001rank.html?countryName=United%20States&countryCode=us&regionCode=noa&rank=2#us.

Chhapia, H. "1,100 quota seats in IITs not filled this year." *The Times of India* (25 June 2009). http://timesofindia.indiatimes.com/articleshow/4699526.cms?prtpage=.

Desai, S., and Kulkarni, V. "Changing educational inequalities in India in the context of affirmative action." *Demography 45,* (2008).

Dias, Dina, Claisy Marinho-Araujo, Leandro Almeida, and Alberto Amaral. "The democratization of access and success in higher education: the case of Portugal and Brazil." *Higher Education Management and Policy* 23, no. 1 (2011). 1–20. Print.

Dobren'kova, E.V. "Problems of Russia's Entry into the Bologna Process." *Russian Education & Society* 50, no. 4 (2008): 42–5. Web.

FengLiang, Li, and W. John Morgan. "Private higher education in China: access to quality higher education and the acquisition of labour market qualifications by low-income students." *Education, Knowledge & Economy* 2, no. 1 (2008): 27–37. Web.

Gupta, A. "Affirmative action in higher education in India and the US: A study in contrasts." *Center for Studies in Higher Education, University of California: Berkeley* (2006): 8.

Khalin, V. V. "On Financing the Training of Master's-Degree-Holders in Russian Higher Education." *Russian Education & Society* 52, no. 1 (2010): 30–44. Web.

Kolesnikov, V. N, I. V. Kucher, and V. N. Turchenko. "The Commercialization of Higher Education." *Russian Education & Society* 47, no. 8 (2010): 35–48. Print.

Lai, Desheng, Yongpo Tian, and Dahu Meng. "Differences Between the Employment of Higher Education Students from the Cities and the Countryside and the Fairness of Higher Education." *Chinese Education and Society* 44, no. 1 (2011): 3–26. Web.

Lloyd, Marion. "Affirmative Action, Brazilian-Style." *The Chronicle of Higher Education* 56 (2008): http://proxy.geneseo.edu:2048/login?url=http://search.ebscohost .com/login.aspx?direct=true&db=ehh&AN=45036880&site=ehost-live&scope=site.

Mahapatra, D. "IITs can't expel SC, ST students on merit excuse." *The Times of India* (13 August 2009).

McCowan, Tristan. "Expansion without equity: An analysis of current policy on access to higher education in Brazil." *Higher Education* 53, (2007): 579–598. Web.

McCowan, Tristan. "The growth of private higher education in Brazil: implications for equity and quality." *Journal of Education Policy* 19, no. 4 (2004): 454–470. Web.

Neelakantan, S. "Industrialists create new private universities in India." *Chronicle of Higher Education* 56 (2010).

Nemtsova, A. "At Moscow State U., 2 Journalism Schools, Divided by Politics." *Chronicle of Higher Education* 56, no. 2 (2009): A50–A51. Print.

Panizzi, Wrana Maria. "Public Universities: A Benchmark for Higher Education in Brazil." *Higher Education Management and Policy* 15, no. 1 (2003): 43–55. Print.

Plaksii, S. S. "Exploitation of College Instructors Is the Main Threat to Higher Education and the Development of Russia in the Twenty-First Century." *Russian Education & Society* 47, no. 5 (2005) 6–22. Print.

Pradnizkov, G. A. "The Bologna Process and Its Relevance to Present-Day Education." *Russian Education & Society* 48, no. 7 (2006): 5–16. Print.

Salnikov, N. N., and Burukhin, S. S. "Current State of the Structure of Higher Education in Russia." *Russian Education & Society* 51, no. 11 (2009): 72. Web.

Sen, Amartya. *Development as Freedom.* New York: Anchor Books, 1999. Print.

Shen, Hong. "Access to higher education by student aid in China: results from the national survey of 100,000 students." *Evaluation & Research in Education* 22, nos. 2–4 (2009): 145–166. Web.

Suspitsin, D., and T. Suspitsyna. "Private Higher Education in Russia." *European Education* 39, no. 3 (2007): 62–80. Web.

Tsung, Linda, and Matthew Clarke. "Dilemmas of identity, language and culture in higher education in China." *Asia Pacific Journal of Education* 30, no. 1 (2010): 57–69. Web.

Wang, Houxiong. "Research on the Influence of College Entrance Examination Policies on the Fairness of Higher Education Admissions Opportunities in China." *Chinese Education and Society* 43, no. 6 (2010): 15–35. Web.

Yang, Xu. "Access to higher education for rural-poor students in China." *Educational Research for Policy & Practice* 9 (2010):193–209. Web.

Yu, Kai, and Hubert Ertl. "Equity in Access to Higher Education in China." *Chinese Education and Society* 43, no. 6 (2010): 36–58. Web.

Zhu, Zhiyong. "Higher Education Access and Equality Among Ethnic Minorities in China." *Chinese Education and Society* 43, no.1 (2010): 12–23. Web.

Indonesian Tobacco: A Consumer Culture of Exploitation

Francis Joseph Mandracchia

Globalization is a pervasive, man-made force that is on par with a force of nature in its ability to shape the way the world functions. It is the opening of markets, peoples, cultures, and governments into a global village. It is space and time becoming increasingly irrelevant as technology becomes increasingly relevant in daily life. It is the consequences of these increasingly common interactions as seen in Barber's (1992) "Jihad vs. McWorld" (Lechner & Boli 32). Barber's essay describes the ways in which globalization and local culture make war on each other and the nation state. His binary underscores the paradox of a global village: a world that is defined by evaporating political borders and free-flowing economies, cultures, and peoples. Under globalization, nations have seen great economic prosperity through trade liberalization policies, opening markets to the world economy (Barber 95). Institutions like the International Monetary Fund (IMF) and the World Bank argue that trade liberalization is essential to development of economies. Trade liberalization is a large force behind globalization and has been known to cause unintended catastrophes in which people and nations are exploited (McMichael 151).

Sociologist Immanuel Wallerstein's World Systems theory seeks to better understand these processes and the dynamics of and stemming from globalization. The theory draws on the neo-Marxist literature of exploitation to explain the relative position of nation states in the world capitalist economy in three categories (core, semi-periphery and periphery). The categories are essentially economic zones with core countries such as the United States, representing nations with large, developed economies and high living standards. Nations like the Congo are in the periphery, defined as countries that have seen little economic growth, relatively low living standards, and exploitation from the core. Comparatively, the semi-periphery nations are stuck somewhere in between the former and the latter (Lechner 2001). Indonesia lies claustrophobically confined in between the

two categories in the semi-periphery. Though it looks to the core in hopes of greater wealth and prosperity, it nervously looks back at its past in fear of reverting to its peripheral status.

Since the 1970s, when the country first made efforts to attract transnational corporations and capital, it has seen trends pointing to development. World Bank data (2011) reveals growing literacy rates, gross domestic products, household expenditures, and life expectancy, all indicators of development. As such, the Indonesian people have greater participation in the world economy and in consumer culture. The tobacco sector is one of the industries in which Indonesia has experienced the greatest growth. Globalization has played a major part in increasing tobacco consumption both globally and within Indonesia. Ways in which tobacco has been imposed on Indonesia, from the colonial era to present day, and the consequences stemming from it are exploitative by nature. The consumer market and culture of tobacco threatens to reverse economic development by undermining the health of Indonesian smokers and non-smokers.

The inception of tobacco in Indonesian society was caused by global interaction. The Spanish imported the plant in 1601 to the island of Java, an island in the Indonesian archipelago (Arnez 50). Prior to the advent of tobacco, the betel nut served as both a stimulant and an important fixture of Indonesian culture for thousands of years (Staples & Bevacqua 3). Indonesians would chew the nut much like chewing tobacco. Reid claims, "It was a social necessity for every adult in society. To refuse to offer betel, or to refuse to take it when proffered, was esteemed a deadly insult" (Reid 531). The nut was given high value because of its many uses. While it was used as a means of social integration, it was also used in marriage ceremonies, adulthood rituals, and tributes to the deceased, among other uses. The cultural importance of the betel nut, coupled with a large supply and low relative cost, made the nut a ubiquitous object in Indonesian society. Tobacco would come to be used as an additive in the betel's recipe, along with spices and other organic material (532).

As early on as the eighteenth century, it was understood that tobacco and the betel nut had similar stimulating effects on the human body (Penzer). With its similarities to tobacco in a social and pharmacological sense, the betel nut was usurped by tobacco in a relatively short amount of time (Arnez 50). The nut and tobacco served certain social

functions in Indonesian society (Porth, Neutzling & Edwards). In the case of the betel nut, it served as a mechanism for social congregation and amicability stemming from its stimulating properties. Tobacco, though lacking the significance the nut had in rituals and other important rites of passage, served a similar function as the nut in terms of daily social interaction. The congruence of the two commodities coupled with the coercive assimilation of Indonesians by their Dutch colonizers, may explain the relatively quick change from betel nut to tobacco.

While transnational linkages were a major actor in the diffusion of Dutch culture during this era, there was no mass media that rapidly broadcast information and culture. Instead, education was used as an essential vehicle of cultural appropriation to implement Dutch ideals into Indonesian society. The Dutch saw betel nut chewing as an extremely primitive act. Many young Indonesians opened their perspectives to the Dutch way of life and in doing so hoped to attain a higher status within society. Along with a higher status came access to education opportunities. Many rejected the betel nut for the more modern and Western white cigarette (Arnez 51). During the colonial era, educated Indonesians participated in their changing, Dutch-influenced culture by following many of the 'scripts' necessary to assimilate. The patriarchal society excluded women from education and in turn, the immediate conversion from the betel nut to tobacco.

> In areas such as Tana Toraja, where betel-chewing among both sexes is still common today, the first generation to have been exposed to education was also the first to abandon betel. The whole image of "modernity" that education conveys to a young person is contradictory to the chewing of betel. Their association with education, and subsequent employment in the modern sector of the economy, is no doubt part of the reason why men generally abandoned betel-chewing a generation or so ahead of women (Reid 539).

These scripts usually envisioned a schism in the roles of each gender into separate spheres. Males would often enter into the public labor force for cash crops, or in some cases the modern work force, after receiving an education (McMichael 34). The Dutch sought to educate and employ Indonesian elites as civil servants, even though Indonesian labor was traditionally reserved for agriculture. The motives behind this initiative were to cut administrative costs by placing Indonesians into the primarily Dutch-run civil service sector, embedding Dutch favor into educated Indonesians and setting an example for those in the lower echelons of society (Ricklefs 156).

Coupled with a long Islamist influence, the Dutch ideology constructed a domestic sphere for women that defined their status as wives and "childrearers." As patriarchy was the defining gender discourse of the West during the colonial period, men were primary actors in almost all facets of Indonesian society, excluding the home (Blackwood 864). This left little room for women to receive an education, especially since "the cause of women's education…was never taken up as a government priority" (Ricklef 157). Women, and the lower classes, for a brief period, saw little in switching to tobacco. This could be attributed to their exclusion from participating in a predominately male, elitist culture. Exposure to education was important to the insemination of Western practices and ideals in Indonesia at that time (Arnez 51).

Indonesian education was at its height when education became accessible to the lower classes at the end of the nineteenth century. By 1930, there were over 1.7 million Indonesians being educated in Western educational institutions, as compared to about 270,000 Indonesians only 30 years prior (Ricklef 159). Betel nut use continued to decline through 1950. By this time it appears as though the nut was almost entirely gone from daily Indonesian life (Arnez 50). Tobacco became the drug of choice in social settings. Indonesians chewed and smoked tobacco, enjoying their new-found "modernity."

Just as education was a gentrified structure, the ways in which the different classes within Indonesian society smoked also varied. The social elite of Indonesia usually smoked white imported cigarettes, while the lower classes would smoke hand-rolled kreteks (Hanusz 25). Kreteks are a hand rolled cigarette composed of cloves and tobacco, wrapped in cornhusks or cigarette paper (13). The Western white cigarette was easily identifiable in contrast to the hand-rolled kretek. While the imported Western cigarette was neat and white in appearance, the kretek varied in its color and its shape, and it left stains on its user. These apparent differences in appearance made smoking a conscious act that was closely tied to class image. Hanusz claims, "[e]ven if they personally liked the taste of kretek, well to-do Indonesians…would light up a white cigarette in public and only enjoy their kretek in the privacy of their home" (25).

The cigarette industry again felt the effects of globalization in the 1970s that expanded the scale of the industry. The first event that spurred overall economic growth was an oil boom in 1974. The boom was mostly due to increasing oil prices mandated by

OPEC (Organization of the Petroleum Exporting Countries), of which Indonesia was a member (Booth 77). Investment capital, both foreign and domestic, flooded the Indonesian economy during this time (264). The influx of investment capital was abetted by an upsurge in government revenue from the growing economy, specifically in oil. Taking into account historical exchange rates, government revenue from oil alone went from 25,504 rupiah million ($86 million US dollars) to 1,540,000 rupiah million ($3.7 billion) in a matter of eight years, from 1968 to 1976 (Antweiler 3; Palmer 48). Growing tax revenues led to a rise in the number and size of low-interest loans given out by the government. Many industries grew as a result of the ascending Indonesian government and economy. The kretek industry was included in this trend as the loans attracted entry into the market by many (Hanusz 29).

Another phenomenon associated with the growth of the kretek industry was the ability to mass-produce their product. Kreteks were solely hand rolled, until 1968 when the Indonesian government granted a few large kretek–producing firms licenses to mechanize their production. By the 1980s, most firms were mechanized and this allowed several companies to dominate the market in future years. Firms seen on today's market such as Djarum, Gudang Garam, and Sampoerna, were mechanized in 1976, 1978, and 1983, respectively (Tarmidi 90). Today these firms are the top three market share holders in the tobacco economy, suggesting the mechanization of kretek production was essential to firms' development and success ("Tobacco Industry Profile—Indonesia" 3).

As such, many small businesses exited the industry because of their inability to compete with larger firms, thus consolidating the industry. They were unable to mechanize in time to compete with the larger firms that initially received mechanization licenses and they were more susceptible to fluctuations in commodity prices because of corruption within the Indonesian government. This can be seen in the aggregate data on the number of firms within the cigarette industry. In 1961, there were a total of 941 companies producing cigarettes. By 1985 there were only143 companies making the commodity (Arnez 57).

Prior to mechanization, many firms were entering the kretek market because of the considerable gains the industry proffered. This led to a rise in employment and a subsequent labor shortage. Mechanization came at an ideal time for an industry that was

facing labor shortages (Saptari 14). The large firms that received the licenses to mass-produce experienced great benefit from this since they were able to cut labor costs as well as mass produce the commodity (Tarmidi 97). Data from the Barber et al. report (2008) reveals that after only after 1970 did cigarette production significantly increase. The industry manufactured 48 billion cigarettes in 1970 and expanded this to just over 200 billion sticks by 2000 (40). Though this data may provide a picture of the tobacco consumption that was developing, it does not take into account the large portion of the population that hand-rolled kreteks (Reid 540).

Once the industry had the capacity to turn out kreteks, they needed to expand their market. The companies turned to the middle and upper classes, which had been smoking white cigarettes, a symbol of modernity. The mechanization of kretek production gave the kretek a uniform appearance, similar to white cigarettes. Kreteks were no longer a poor man's smoke, one that he had to hand roll. To convey this large firms spent a great deal of money marketing the kretek as a "sophisticated smoke for a sophisticated generation" (Hanusz 29). Packaging emulated that of foreign white cigarette brands in an attempt to appeal to the young, professional class in Indonesian society. The kretek was becoming "a precise marker of the achieved status of the host who dispenses this new narcotic" (Reid 542).

The predominance of tobacco use in Indonesia and the rapid growth of the industry primed the market for transnational corporations' increased attention—and capital. The two major transnational corporations, Philip Morris International and British American Tobacco, were both part of the Indonesian economy by 1984. However, it was not until the aughts that the companies began acquiring local kretek firms (Lawrence & Collins 97). This was largely a response to the unsuccessful attempts to extend white cigarette consumption passed 14.2 percent of the market (Barber et al. 74).

In 2005, Philip Morris International became a majority shareholder in the kretek brand Sampoerna, attaining 97.5 percent of shares. Philip Morris Indonesia is another company owned by the conglomerate that operates a factory that produces much of the firm's white cigarettes, including Marlboro, a top seller among white cigarettes in Indonesia ("Tobacco Industry Profile—Indonesia," 3). British American Tobacco (BAT) existed as early on as 1908 and has been a significant force in the industry since

(Lawrence & Collins 97). By 2009, BAT acquired Benotel, a local brand of kretek, holding 85 percent of shares. Benotel is also a conglomerate which owns a number of other holdings like development, trade, and industrial enterprises as well as an amusement park. The company also operates through BAT, producing predominately white cigarettes with popular international brands such as Lucky Strike and Pall Mall. It has been reported that Lucky Strike has been particularly successful in gaining a niche among Indonesian smokers. Recently, Philip Morris International has become the dominant firm in the tobacco industry as a whole, commanding a 37.8 percent market share after acquiring BAT and its subsidiary Benotel ("Tobacco Industry Profile—Indonesia," 3–4).

The make-up of local firms in the tobacco market composes a majority at 62.2 percent. While this proportion of the market is significant, there are a large number of firms in this sector. It is estimated that there are approximately 5,000 local firms sharing the local market. What is unique about the Indonesian tobacco market is its resistance to domination by transnational corporations. As one transnational tobacco executive explained, "tiny factories can compete in the same market as huge multinationals." Performance of local companies is considerable, considering the presence of TTCs. Gudang Garam, a local corporation, has a 22.1 percent market share and is the top producer of kretek in the world (5).

The consumer tobacco market has been the impetus in the growth of the industry. There is a sizeable and growing demographic consuming tobacco products. With a total population of 245,613,043 people and 63.1 percent of males and 4.5 percent of females smoking by 2004, Indonesia has a sizeable tobacco population (Barber et al. 62; World Bank). Consumer preference indicates that kretek is the ideal tobacco product in the tobacco market, with 88.1 percent of smokers preferring kreteks to a mere 11.9 percent preferring white cigarettes. Tobacco composes 11.5 percent of total household income, which is a significant portion considering health and education account for 2.3 percent and 3.2 percent, respectively (64–5).

Advertising in Indonesia has been an innovation essential in appealing to consumers. The use of marketing has managed to socially reconstruct the perceived image of tobacco consumer culture into one that is as attractive as it is young and

masculine. As reflected by the large number of males who are smokers, advertising is particularly made relevant to images of masculinity. Advertisements also take advantage of inherent social cues in Indonesian society. Firms place advertisements near communities' culturally significant locations like mosques. Outside the places of worship large ads that simply read, "wish you well," are clearly sponsored by companies like Bentoel, a kretek firm (Reynolds 86).

Strategic promotions and sponsorships are aimed at young audiences. Though regulation prohibits advertising specifically aimed at children, these events are indirectly targeted towards young audiences. Sponsored shows similar to American Idol as well as concerts with international artists like Alicia Keys and the Cranberries are run by tobacco companies and draw much attention through attendance and media coverage. Sports events such as soccer tournaments and drag racing are also planned by the tobacco industry. Many times the intent of youth culture marketing is to attract smokers early, so the industry may gain lifelong consumers ("Tobacco Industry Profile—Indonesia," 8).

Factors involving price play a large role in the consumption of cigarettes, including kreteks. With the average pack of cigarettes (16 sticks per pack) costing around 0.58 US dollars, Indonesia has the most affordable cigarettes in South Eastern Asia. Beyond looking at price, Guindon et al. (2003) calculated the affordability of the commodity by dividing cigarette prices by South East Asian countries' per capita gross domestic product (GDP) and found these results still held, with Indonesia at the very bottom in terms of expense (Barber et al. 23). Contributing to the low cost of cigarettes is the Indonesian government's low tax rates that have remained relatively stagnant for a number of years (19). Another contributing factor to low relative cost is the ability of transnational corporations to invest readily available capital into expanding and mechanizing efficient production ("Tobacco Industry Profile—Indonesia," 3).

Cost is an important factor in consumption as it determines demand levels. Analysis of the effects of price on demand can be better understood by analyzing demand curves. In observing the Indonesian market, analysis has consistently found demand to be inelastic. Essentially, as the price of a good increases, consumers will still be willing to purchase the good, at any price (Barber et al. 27). Tobacco's addictive properties make elasticity more inelastic as consumers who are addicted will be more willing to pay

higher prices to meet their tobacco needs (33). Income elasticity, which studies the relationship between income and consumption, has also consistently found cigarettes to be a normal good. As such, when incomes increase, the consumption of the good increases (27). World Bank data reveals that incomes have been on the rise since the country opened its borders, which can be directly linked to higher tobacco consumption rates because the product is a normal good (World Bank). Moreover, the diversity in tobacco products at many different price ranges allows for consumers in most economic positions to afford the commodity ("Tobacco Industry Profile—Indonesia," 7).

Many of the conditions of the tobacco market have made it ideal to attract new entrants. It is estimated that over 70 percent of these entrants begin smoking between the ages of 5 and 19 (Barber et al. 66). Some estimates indicate that 20 percent to 25 percent of ten-year-old boys smoke (van der Meij and Hüsken 50). Even more troubling is the fact that kreteks are more appealing to children as the cloves in the cigarette make the smoke sweeter. Many times the harsh flavor of a white cigarette will dissuade most children from smoking; however, the kretek's taste is attractive to many first-time smokers. While there is little evidence on the effects of kreteks relative to white cigarettes, data has shown higher carbon monoxide, nicotine and tar emissions in kreteks (52). While no data is available, social workers have reported infant-smoking cases have even been on the rise (Putzel et al). The conditions of the market and the prevalence of cigarette advertising have created prime conditions for smokers to begin at an increasingly younger ages. The average age of smoking initiation has gone from 18.8 in 1995 to 17.4 by 2004 (Barber et al. 66).

Increasing childhood smoking rates are one of many effects of the market failures in the Indonesian tobacco economy. Lack of regulation has created an environment where consumers, especially children, are bombarded with cigarette advertisements. The Indonesian government's lack of regulation has allowed ads to be extremely prevalent and invasive. Some of the few restrictions the industry faces are limiting advertising on television to only be broadcasted between the hours 9:30 p.m. and 5 a.m. as well as regulation requiring health warnings must be posted on all advertisements, with no minimum size or amount time the warnings must be shown for (Barber et al. 81; Danardono 99). Advertisements have been known to be placed right outside playgrounds

and schools (Rorandelli; Putzel et al.). The Centers for Disease Control found that, on average, 90 percent of children had been exposed to at least one cigarette advertisement over the span of one month (cited in Barber et al. 12). Lack of information on the health risks associated with smoking is also a prevalent issue. One study found Indonesian boys from 12 to 17 years old believed smoking one to two packs a day was not harmful to health, even though they were fully aware of the warning labels on the packaging (Ngi 797). Even less is understood about the health impacts of secondhand smoke.

Approximately 97 million Indonesian non-smokers are frequently exposed to secondhand tobacco smoke ("Tobacco Burden Facts"). Of the 200,000 deaths that are reported annually one-eighth, or 25,000, are attributed to secondhand smoke (Barber et al. 6). Large numbers of non-smokers are affected by secondhand smoke largely because there is little regulation where smokers can and cannot smoke. The government has no laws banning smoking in designated areas; however, some health facilities and schools have been known to ban use ("Indonesia: Tobacco Policy Status"; van der Meij & Hüsken 58).

With tobacco deeply engrained into Indonesian culture and exposure to the drug at unprecedented levels, the effects of the tobacco smoke on health and the economy are sure to be widespread. As the number of smokers is expected to continue to rise, illness will increase significantly as well. Little data is available on the effects of tobacco on public health in Indonesia because statistics are just beginning to emerge on the consequences of 25 years of smoking on health. Estimates available are extremely limited because of the lack of information as well as the low rates of health services use. Some predict the health care costs of tobacco related illness to be between $484 million and $1.2 billion. These estimates, however, are based on the healthcare costs of other countries, even further affecting the estimate's accuracy (Barber et al. 13).

Future projections seem to point towards disconcerting trends. As nations continue to develop and tobacco consumption rises—as well as healthcare costs related tobacco, developing nations like Indonesia will face an "epidemiological trap, saddling their new health care systems [due to development] with considerable new health costs while having to continue treat infectious disease [related to tobacco]" (Kawachi &

Wamala 53). These aforementioned projections do not account for the possibility of expansion of the tobacco market to women in Indonesia.

Though the market has remained untouched by tobacco, trends in development have shown that as development increases, women's tobacco consumption increases as well. As global influence (namely Western influence) increases, entrenched cultural norms on gender begin to change. Further, transnational tobacco companies exploit this narrowing gender gap in order to expand their market to the female population (Hartwig et al. 204). Marketing towards women in developing nations takes cues from this and exploits the image of freedom and vanity, as women respond to different cues in marketing. Since women have both reproductive and productive duties, and consequent high stress levels, as the intersection of these two roles often cause, they are more likely to turn to smoking for stress relief (207).

Increases in smoking rates among women will lead to higher rates of secondhand smoke exposure among the family, especially among children. Illness and death associated with secondhand smoke, such as cancer and heart disease, will rise as well. Among children there will be severe effects, such as sudden infant death syndrome and asthma. Beyond this children will be raised in homes and in a society where smoking is seen socially acceptable (Hartwig et al. 230; Barber et al. 12).

Not only do the social spheres of Indonesian society face regression, but the Indonesian economy also has the potential to be greatly hindered by tobacco consumption. There is no doubt that constant, prolonged use of tobacco is harmful to health, as it is expensive. Health and the economy go hand in hand, as human capital is only worth as much as the ability of its labor force to contribute to the economy. A healthy work force means more productivity and less sick days. Healthy consumers can take full advantage of their working capacity. Greater and widespread accumulation of wealth in Indonesia would lead to increased investment in the economy (Barber et al. 8).

Contrasted with this is potential for unbounded tobacco consumption. Smoking can cause illness and decreased well-being. Statistically smoking is known to increase the rate of general infections (9). With the rise in illness comes more sick days (Eriksen & Mackay 40). Manual labor is greatly affected by this, as there usually is no compensation for missed work time. Smoking also causes increased death rates and many of these

mortalities occur when a person is in the productive span of their life. Early death of parental figures can also cause decreases in quality of life, including higher dropout rates, of their children (Barber et al. 9). Further, expenditures on tobacco in low-income households has been shown to increase poverty as it takes away from human capital investments such as money spent on healthcare and education (Esson & Leeder 35). Food is also affected by expenditures on tobacco consumption. Malnutrition is directly affected by smoking in the poorest of households. The father, who is usually the primary provider of income, will spend much of his wages to pay for his cigarette addiction and leave little left to feed his family (Semba et al. 9).

Tobacco is largely a consequence of globalization's ability to disperse commodity flows cross-nationally. Trade liberalization, a catalyst of globalization, which has ushered in gains in economies and qualities of life, has also embedded the potential to see these developments degenerate. As time goes on, emerging trends will point to the damaging effects of this commodity on people and society. Many of these models have played out in countries like the United States, where the tobacco industry was able to attain much success through the exploitation of the legal system, modern science and the trust of the people (Brand). Fortunately, the relative stability of the United States afforded its citizens, for the most part, retributive justice.

Countries like Indonesia face even more pervasive exploitation by tobacco companies, especially transnationals that are well aware of and have historically used the damaging effects of their product as a defense in litigation in the United States. Indonesian consumers are largely unaware of the effects of the product and the Indonesian government fears action in the defense of its people in fear of the fleeting capital of these corporations. Indonesia stands at the pivotal point of a dilemma—the balancing of economic interests with the interests and health of the people who make an economy possible.

Works Cited

Antweiler, W. "Foreign currency units per 1 U.S. dollar, 1948–2009." *Pacific Exchange Rate Service*. University of British Colombia, Sauder School of Business, 2009. Web. http://fx.sauder.ubc.ca/etc/USDpages.pdf

Arnez, M. (2009). Tobacco and kretek: Indonesian drugs in historical change. *Austrian Journal of South-East Asian Studies, 2*(1), 49–69.

Barber, S., Ahsan, A., Adioetomo, S. M., & Setyonaluri, D. *Tobacco economics in Indonesia.* Paris: International Union Against Tuberculosis and Lung Cancer, 2008. Web. 17 Nov 2011. http://www.tobaccofreecenter.org/files/pdfs/en/Indonesia_tobacco_taxes_report_en.pdf

Blackwood, E. "Gender transgression in colonial and postcolonial Indonesia." *The Journal of Asian Studies* 64.4 (2005): 849–879. Print.

Booth, A. *The Indonesian Economy in the Nineteenth Century and Twentieth Centuries: A History of Missed Opportunities*. New York: St. Martin's Press Inc., 1998. Print.

Brandt, A. M. *The Cigarette Century: The Rise, Fall and Deadly Persistence of the Product That Defined America*. New York: Basic Books, 2007. Print.

Danardono, M., et al. "Reading Culture from Tobacco Advertisements in Indonesia." *Tobacco Control*, 18.2 (2009): 98+. Web http://go.galegroup.com/ps/i.do?id=GALE%7CA198818815&v=2.1&u=genese&it=r&p=HRCA&sw=w

Hanusz, M. *Kretek: The Culture and Heritage of Indonesia's Clove Cigarettes*. Jakarta: Equinox Publishing, 2000. Print.

Hartwig, K. A., Kickbusch, I., & List, J. M. (Eds.). *Globalization, Women and Health in the 21st Century*. New York: Palgrave Macmillan, 2005. Print.

van der Meij, D., & Hüsken, F. *Reading Asia: New Research in Asian Studies*. Richmond: Curzon Press, 2001. Print.

Eriksen, M., & Mackay, J. "The tobacco atlas." *Who.com.* World Health Organization, Tobacco Free Initiative, 2002. http://www.who.int/tobacco/statistics/tobacco_atlas/en/

Esson, K. M., & Leeder, S.R. "The millennium development goals and tobacco control."
	Who.int. World Health Organization, 2004. Web.
	http://www.who.int/tobacco/publications/mdg_final_for_web.pdf

"Indonesia: Tobacco burden facts." *Tobacco Free Center, International Facts Website.*
	Campaign for Tobacco-Free Kids, 2010. Web.
	http://tobaccofreecenter.org/files/pdfs/en/Indonesia_tob_burden_en.pdf

"Indonesia: Tobacco policy status." *Tobacco Free Center International Facts Website.*
	Campaign for Tobacco-Free Kids, 2011.
	Web.http://tobaccofreecenter.org/files/pdfs/en/Indonesia_tob_policy_en.pdf

Lawrence, S., & Collins, J. "Competing with kreteks: transnational tobacco companies,
	globalization and Indonesia." *Tobacco Control,* 13.2 (2004): ii96–ii103. Web.

Lechner, F. "Globalization Theories." *The Globalization Website.* Emory University,
	2001. Web. http://www.sociology.emory.edu/globalization/theories01.html

Lechner, F., & Boli, J. *The Globalization Reader*. Malden: Blackwell Publishing, 2008.
	Print.

McMichael, P. *Development and Social Change: A Global Perspective*. London: Pine
	Forge Press, 2008. Print.

Ngi, N., Weinhall, L., & Ohman, A. "If I don't smoke, I'm not a real man—Indonesian
	boys' views about smoking." *Health Education Research* 22.6 (2007): 794–804.
	Print.

Palmer, I. *The Indonesian Economy Since 1965*. London: Frank Cass and Company, 1978.
	Print.

Penzer, N. M. *Poison-Damsels and Other Essays in Folklore and Anthropology*. London:
	Chas. Sawyer, 1952. Print.

Porth, E., Neutzling, K., & Edwards, J. "Functionalism." *Department of Anthropology.*
	University of Alabama, 2009. Web.
	http://anthropology.ua.edu/cultures/cultures.php?culture=Functionalism

Reid, A. "From betel-chewing to tobacco smoking in Indonesia." *The Journal of Asian
	Studies* 44.3 (1985): 529–547. Print.

Reynolds, C. "Tobacco advertising in Indonesia: 'The defining characteristics for
	Success'." *Tobacco Control,* 8.1 (1999): 85–88. Web.

Ricklefs, M. C. *A history of modern Indonesia since* c.*1930.* Stanford: Stanford University Press, 1933. Print.

Rorandelli, R. (2011). "Tobacco in Indonesia." *Italian Collective of Documentary Photographers.* Terra Project Photographers, 2011. Web. http://www.terraproject.net/en/photographers/rocco-rorandelli/tobacco-in-indonesia

Staples, G. W., & Bevacqua, R. F. (2006). "Areca catechu (betel nut palm), ver. 1.3." In: Elevitch, C. R. (ed.). *Species Profiles for Pacific Island Agroforestry. Permanent Agriculture Resources (PAR), Hōlualoa, Hawaii.* Traditional Tree Initiative, 2006. Web. <http://www.traditionaltree.org>.

Saptari, R. *The political economy of smoking; The case of the kretek cigarette industry in Indonesia.* P. Boomgard, R. Sciortino, & I. Smyth (Eds.). Leiden: KITLV 1996. Print.

Tarmidi, L. "Changing structure and competition in the kretek cigarette industry." *Bulletin of Indonesian Economic Studies* 32.3 (1996): 85–107. Print.

"Tobacco industry profile—Indonesia." *Tobacco Free Center International Facts Website.* Campaign for Tobacco-Free Kids, August 2009. Web. http://www.tobaccofreecenter.org/files/pdfs/en/IW_facts_countries_Indonesia.pdf

"Vanguard." Prod. C. Putzel, A. Simmons, J. Plunkett. Current TV, New York. 2011 June 28. Television.

Wamala, S., & Kawachi, I. *Globalization and Health.* New York: Oxford University Press, 2007. Print.

World Bank Data. (2011). *Indonesia.* Retrieved from http://data.worldbank.org/country/indonesia

Investigation of Compounds in *Anaphalis Margaritacea*

Lauren Healy

ABSTRACT

The roots and tops from *Anaphalis margaritacea*, commonly referred to as Pearly Everlasting, were extracted using a mixture of ether and petroleum ether and analyzed through the use of various spectroscopic techniques. Dr. Ferdinand Bohlmann previously reported a thirteen carbon chlorinated polyacetylene found in the roots of *Anaphalis* species that was of much interest due to its structural similarity to known antibacterial and antifungal compounds. The goal was to successfully isolate and characterize this compound, known as (E)-5-chloro-2-(octa-2,4,6-triynylidene)-5,6-dihydro-2H-pyran, as well as identify other compounds previously unidentified in *A. margaritacea*.

Although this particular compound has not been identified yet, various other compounds have, including terpenes, unsaturated compounds, and ring systems. These compounds are of interest in their own right, as they too may be useful as pharmaceuticals or antimicrobial agents.

LIST OF TABLES

LIST OF FIGURES

Figure 4: Proposed Structure of an Isolated Flavonoid

Figure 5: Phytosterols

Figure 6: Silver Ion SPE Column

LIST OF ABBREVIATIONS

AL#1………………………………………………………………Alumina Column #1

AL#2………………………………………………………………Alumina Column #2

AL#3………………………………………………………………Alumina Column #3

AL#4………………………………………………………………Alumina Column #4

AL#5………………………………………………………………Alumina Column #5

AL#6………………………………………………………………Alumina Column #6

AL#7………………………………………………………………Alumina Column #7

AL#8………………………………………………………………Alumina Column #8

AL#9………………………………………………………………Alumina Column #9

AL#10………………………………...………………………………..Alumina Column #10

AL#11……………………………………………………………….Alumina Column #11

AL#12……………………………………………………………….Alumina Column #12

SC#1……………………..…………………………………………….Silica Column #1

SC#2………………………………………………………..….Silica Column #2

SC#3……………………………………………………………….Silica Column #3

SC#4………………………………………………………………..Silica Column #4

SC#5………………………………………………………………..Silica Column #5

SC#6……………………………………………………………Silica Column #6

MeOH……………………………………………….………………...Methanol

EtOAc………………………………………………………...Ethyl Acetate

INTRODUCTION

Natural products have an extremely large role in the pharmaceutical industry. Between 60% and 75% of all cancer and infectious disease drugs currently on the market approved for human treatment are estimated to either be natural product derived or mimic the actions of a natural product. As many as 79% of all anti-infectives are of natural origin along with up to 70% of anti-migraine drugs on the market, and there are over 100 natural product based drugs in trial at this time (4).

The current method used by the pharmaceutical industry is called high throughput screening (HTS), which uses a mass targeting approach to identifying potential drugs. Target molecules are placed into wells and treated with anything ranging from a potential ligand to a new drug via an automated process. If any of the substances in the wells react with the treatment, this is considered a hit and these substances are then further processed. For every 100,000-300,000 samples, there are about 100-300 hits, which lead to 1-2 potential "lead compounds" that have potential to actually work. This is a hit rate of only about 0.001% and only a 1% chance the hit will actually turn out to be useful (10). In comparison, there are roughly 7,000 known metabolites from plants, bacteria, and fungi known as polyketides. Twenty of these have led to successful pharmaceuticals, which is about a 0.28% hit rate. The importance of natural product study for the advancement of the drug industry and benefit of human health is clear from this data.

Long before the modern drug industry however, various Native American tribes were already reaping the benefits natural products had to offer. For example, *Oplopanax horrius* (more commonly known as Devil's Club) was used by the First Nation's people to treat a wide variety of ailments, including diabetes, tuberculosis, headaches, and lung hemorrhages. Investigation into the inner bark of this plant revealed several compounds that exhibited antimycobacterial and antifungal activity, including several polyynes[7]. Mycobacteria are rod shaped bacterial that are responsible for infections such as tuberculosis and leprosy, so any compounds that show an ability to kill mycobacteria may have use as a new drug against these infections (8). Also, polyacetylenes ranging in length from 15 to 17 carbons were isolated from the stony coral *Montipora* and were proven successful as cytotoxic agents against several types of human tumor cell lines, with a particularly high cytotoxicity towards a line of human skin cancer⁻ (3).

Anaphalis margaritacea, a rhizomatous dicot belonging to the Asteracea family that blooms perennially in the spring and summer, was also used to treat a broad spectrum of ailments (1). The Cherokee, Cheyenne, Mohegan, and Iroquois all used the aerial parts of A. margaritacea as a treatment for respiratory problems such as colds, coughs, and asthma, whereas the roots were used by the Iroquois and Okanagan-Colville to treat gastrointestinal disorders including "poison stomach", dysentery, and diarrhea (2).

Previous investigations into the chemical composition of the roots by Dr. Ferdinand Bohlmann revealed the compound (E)-5-chloro-2-(octa-2,4,6-triynylidene)-5,6-dihydro-2H-pyran (Figure 1), which was later also found in species of *Gnaphalium* and *Dicoma* (5,3). This particular compound is of interest since it is both a polyacetylene and chlorinated; two characteristics of many currently successful pharmaceuticals.
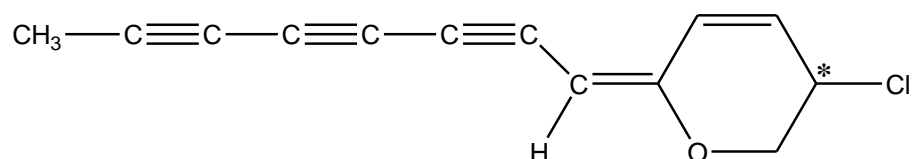
Figure 1: The structure of (E)-5-chloro-2-(octa-2,4,6-triynylidene)-5,6-dihydro-2H-pyran. The asterisk marks the one chiral carbon.

The nature by which the plants produce these compounds is still not entirely known. It is speculated that the polyacetylenes begin via the crepenynate pathway, which begins with the combination of acetyl-coA and malonyl-coA and the initiation of fatty acid synthesis to produce an eighteen-carbon fatty acid. Acetylenases and desaturases add the double and triple bonds, and carbons are cleaved off via alpha and beta cleavage to yield the lesser carbon products. The ring most likely forms after the desaturation process occurs (9).

How and when the chlorine is added is also not known, although it is thought to occur after the ring closes via an S_N2 attack on a hydroxyl group present in that position (11). The carbon to which the chlorine attaches is chiral, but the chirality has yet to be determined. If this compound is to be synthesized for use as a pharmaceutical it is imperative that the chirality is known. This is not only to see if there is only one enantiomer

present in the plant that is acting as the antimicrobial agent, but one enantiomer could have negative consequences if ingested.

The reason plants produce these compounds is also uncertain. However, polyacetylenes and related compounds tend to be tissue specific and transient in nature, leading to the thought they may be produced in response to a particular stressor or premade in anticipation of a stressful situation then released upon exposure to the stress (12). This could pose a problem when attempting to isolate these types of compounds because if the plants used for extraction were not exposed to the stressor, they may not produce the compounds.

There have been very few compounds successfully isolated, characterized, and reported from *Anaphalis margaritacea*, so any compounds found and identified are of great interest. Other types of compounds expected include terpenes, phenylpropanoids, sterols, carotenoids and other pigmentation molecules, fatty acids, triacylglycerols, flavonoids, and other secondary metabolic products. Careful spectroscopic analysis of fractions collected after column chromatography can help identify the compounds collected, which, once isolated, can then be further investigated for properties such as antimicrobial activity.

METHODS

Two separate root extractions and one top extraction of *Anaphalis margaritacea* were analyzed over a period of one year. The root extractions were done using the same technique in both cases. The samples of *A. margaritacea* were collected and the roots were separated from the tops. The tops were allowed to dry and the roots were washed with warm water. A total of 588g of roots were used for the second extraction, which were then cut into small pieces and further chopped in a food processor. These roots were then placed into 1000mL Erlenmeyer flasks and extracted with a mixture of 2:1 petroleum ether to ether (hereon referred to as the nonpolar extract). The roots were then extracted with pure methanol (hereon referred to as the polar extract). Data from the first and second root extractions can be found in the appendix (Table 1-Table 12).

During the second extraction, the nonpolar solution was poured off of the roots and into separate flasks, then evaporated down and transferred into a smaller flask. There was a residue remaining on the inside of the flask that would not dissolve in the petroleum

ether:ether mixture. This residue was dissolved in methanol and placed in a separate flask, in hopes of obtaining a crystal.

The tops were treated in a manner similar to the roots, also extracted with a 2:1 mixture of petroleum ether to ether, followed by pure ethyl acetate, and finally pure methanol. All extracts were stored in a freezer until further analysis.

The root extracts from extract 1 were the first to be examined, and were evaporated onto aluminum oxide for analysis by column chromatography using a rotary evaporator. The nonpolar extract was a pale yellow powder and underwent column chromatography in accordance with the details present in appendix table 1. The polar extract was a slightly darker yellow-orange color and was also exposed to column chromatography in accordance with the details in appendix table 4.

The nonpolar and polar root extracts from extraction 2 were treated in the same manner as the root extracts from extraction 1, and the specific data can also be found in the appendix (Table 9-Table 12). The only difference in the top extraction was the extraction using 2000mL of ethyl acetate following the petroleum ether:ether extraction. This fraction was a deep brown color, and the methanol extraction that followed was a deep green color. Any fractions showing distinct spots when observed using TLC were then analyzed by GCMS. Any samples that were confirmed as pure by GCMS were then analyzed by NMR (Appendix Tables 13-14, Appendix Figures 1-3).

All alumina columns were prepared by placing a small amount of glass wool in the tip of the column, followed by about 0.5cm of sand. The column was then filled about halfway with hexane and the desired amount of $Al_2O_{3(s)}$ was added, either dry or as a slurry mixed with hexane. A second 0.5cm layer of sand was added then finally the sample was added on top of that. Once started, the columns were not allowed to dry until the most polar of the solvents (typically methanol) had been used. The silica columns were prepared in a similar manner, except the silica was always loaded as a slurry mixed with solvent.

RESULTS

IR data was taken from both the polar and nonpolar extracts from extraction 1 (Table 1). When attempting to dissolve the polar extract in ethyl acetate, some fine white particles were observed in the sample. The IR was taken both of the solution containing these white particles and after they had been filtered off. The results are as expected with the polar fraction exhibiting a significantly stronger hydroxyl and carbonyl stretching and the nonpolar fraction exhibiting stronger C-H stretching. A smaller hydroxyl peak can still be seen in the nonpolar extract, which is most likely due to any long chain fatty acid residues or only very slightly oxidized hydrocarbons that would have dissolved in the ether used for that extraction. A UV-Vis absorption spectrum was also obtained for the nonpolar extract 1, with a maximum absorbance at 360.1nm.

Nonpolar IR (cm^{-1})	3390.03, 2917.75, 2849.78, (2400-1800)*, 1603.48, 1496.68
Polar IR (unfiltered) (cm^{-1})	3202.1, 2161.64, 2049.84, 1979.77, (1700)*, 1631.48
Polar IR (filtered) (cm^{-1})	3389.66, 2927.59, 1694.34, 1606.4

*Table 1: IR bands from the nonpolar and polar extracts from extraction 1. The white particles observed were filtered off of the polar sample, than the sample was reanalyzed. *These values are not exact due to inability to get a computer-generated value off of the IR, so they were determined via inspection.*

Fraction 50 from AL#1 showed a clean peak at 29.808min with a mass spectrum as indicated below:

M/Z: 244.0, 216.0, 201.0, 173.0, 145.0, 127.0, 115.0, 91.0 (g/mol)

This was then run on the NMR to obtain the data seen in Table 2.

Fraction 50	
^{1}H NMR	δ:2.027, 2.033, 2.039, 2.044, 2.050, 2.801, 2.834, 3.873, 5.473, 5.479, 6.075, 6.081, 7.701, 6.740, 6.861, 6.883, 7.296, 7.336, 7.505, 7.526, 8.815 ppm
^{13}C NMR	δ:55.679, 87.860, 99.975, 115.791, 116.395, 129.191, 134.571 ppm

Table 2: NMR data collected from AL#1 fraction 50. The sample eluted out in 70% EtOAc in hxane, showed an absorbance when spotted on a silica TLC plate, was dissolved in acetone-d and analyzed on a Varian 400MHz NMR (For HSQC and TOCSY Data see Appendix Figure 1).

AL#2 yielded numerous clean fractions, including fraction 38 and 184 (Table 3). These fractions have not yet been analyzed by NMR to confirm the proposed structure.
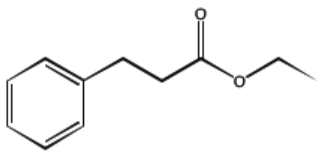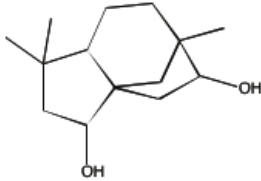
Fraction:	38	184
Eluted in:	Hexane	50% Ether/50% Hexane
Retention Time (GC):	17.768 min	21.619 min
M/Z:	178, 160, 133, 104, 91, 77, 65, 51, 39, 29	238, 220, 205, 187, 182, 164, 150, 135, 121, 107, 93, 81, 69, 55, 41, 29
Library Structure:	Benzenepropanoic acid ethyl ester	4, 4, 8–trimethyltricyclo[6.3.1.0(1,5)]-dodecane-2,9-diol
Chemical Structure:		

Table 3: GCMS data from pure fractions 38 and 184 from AL#2. The spectra were consistent with the structures listed from the GCMS library search.

AL#2 fractions 89-97 showed similar GCMS spectra, so the fractions were recombined, then a ^{1}H NMR and HSQCAD were obtained (Table 5, Appendix Figure 2). The carbon NMR had too much noise to be accurately interpreted.

^{1}H NMR	δ:0.715, 0.735, 0.786, 0.803, 0.822, 0.855, 0.877, 0.899, 0.943, 0.969, 1.003, 1.020, 1.061, 1.082, 1.125, 1.157, 1.180, 1.206, 1.225, 1.244, 1.266, 1.287, 1.316, 1.335, 1.359, 1.397, 1.430, 1.457, 1.514, 1.540, 1.573, 1.592, 1.645, 1.681, 1.712, 1.732, 1.754, 1.792, 1.963, 1.987, 2.014, 2.045, 2.078, 2.109, 2.130, 2.190, 2.258, 2.277, 2.296, 2.339, 2.356, 2.375, 2.394, 2.413, 4.076, 4.093, 5.820, 5.971, 7.307 ppm

Table 4: NMR results from recombined AL#2 fractions 89-97. This data is consistent with that of a sterol of some sort (for HSQCAD see Appendix Figure 2).

Fractions 1 and 2 from AL#8 eluted out in pure hexane into 50mL Erlenmeyer flasks as a very pale yellow liquid. The solvent was allowed to evaporate and a white/pale yellow residue was left behind. An attempt to dissolve this in methanol to transfer to a smaller vial resulted in the formation of a gel like substance that was insoluble in methanol. Ethyl acetate was then used with a similar reaction. Finally, hexane was used and the sample was scraped off of the bottom of the flask into a test tube. A distinct gel layer formed on top of a more liquid looking layer in the test tube (Figure 2). When agitated or heated, the gel would go back into solution and allowing the test tube to sit undisturbed caused the gel to form again. The gel was analyzed by IR and NMR with the details present in Table 4.
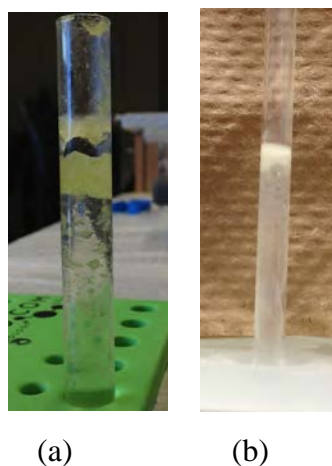
Figure 2: The gel that appeared after methanol was added to fraction one from AL#8. In both the test tube (a) and the NMR vial (b) there is a clear distinction between the gel like layer on top and the more fluid layer on the bottom.

(a)   (b)

Fraction 1 from AL#8	
IR	Only C-H stretching significant
^{1}H NMR	δ: 0.762, 0.779, 0.792, 0.810, 0.826, 1.002, 1.017, 1.185, 1.530, 1.557, 1.606, 1.944, 1.969 ppm
^{13}C NMR	δ: 14.108, 22.685, 23.413, 26.385, 29.358, 29.652, 29.691, 31.921, 32.192, 125.011, 135.191 ppm

Table 5: Spectroscopic analysis of the gel from fraction 1 from AL#8. The IR was run on just the gel collected from the top of the test tube and the NMR data was collected with the sample dissolved in acetonitrile-d on a Varian 400MHz NMR (Appendix Figure 3).

DISCUSSION

While the (E)-5-chloro-2-(octa-2,4,6-triynylidene)-5,6-dihydro-2H-pyran was not successfully isolated at this point in time, there were several other compounds that were, including phenylpropanoids, a sterol or sterol derivative, various terpenes, and a wax or fatty acid residue.

Fraction 38 from AL#2 yielded benzenepropanoic acid ethyl ester (Table 3), which is a phenylpropanoid. After the citrate cycle, phosphoenol pyruvate is converted into 3-dehydroshikimate, which then is converted into shikimate, then chorismate, via the shikimate pathway. Chorismate is converted to prephenate, then phenylalanine, which is used to synthesize the phenylpropanoids (Figure 3). These are well known and studied compounds that help plants respond to stressors such as UV exposure, wounds, infections, and pollutant exposure by scavenging free radicals (13).

PEP   3-DHSH   SH   CH   PP   Phe
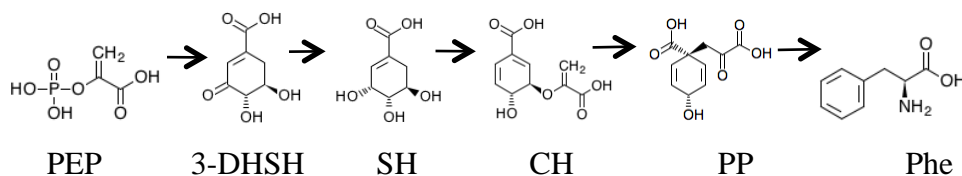
Figure 3: The pathway used to obtain the phenylpropanoids. Phosphoenol pyruvate (PEP) is converted stepwise into 3-dehydroshikimate (3-DHSH), shikimate (SH), chorismate (CH), prephenate (PP), and then finally phenylalanine (Phe), which is used to create the phenylpropanoids (12).

Interestingly enough, these phenylpropanoids have also been shown to exhibit antitumor activity similar to the polyacetylenes present in *Montipora*. Phenylpropanoids collected from the plant *Smilax china* acted as effective cytotoxins on prostate tumor spheroids. In addition to acting as a cytotoxin, these compounds significantly reduced the amount of angiogenesis and proliferation in a range of cancers including murine melanoma, human adenocarcinoma, and uterine carcinoma (13).

These phenylpropanoid compounds can also give rise to flavonoids, which is what is thought to have been isolated in AL#1 fraction 50. The TOCSY showed three distinct spin systems corresponding to the lone double bond and the two benzene rings. There is a proton peak at 3.873ppm that corresponds to the methoxy group, and another peak at 8.815ppm that corresponds to the hydroxyl group (Appendix Figure 1). The exact positioning of the substituents has not been fully determined, but the structure in Figure 4 is what is currently proposed. The hydroxyl group positioned at this location on the ring would allow for hydrogen bonding with the doubly bonded oxygen and would give the most stability. Also, the proton NMR showed two doublet of doublets, with one having twice the integration of the other, and placing the hydroxyl group in this location would give that pattern. The methoxy group must therefore also be located as a para-substituent on the other benzene to give the classic doublet of doublets associated with a p-substituted benzene ring.
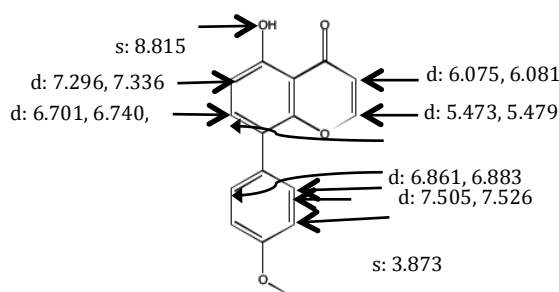
Figure 4: The proposed structure of the flavonoid isolated from AL#1 fraction 50. This compound was eluted in 70% ethyl acetate in hexane and analyzed by GCMS and NMR to determine its structure. The proton shifts are indicated (Appendix Figure 1).

Flavonoids are also known for their ability to scavenge free radicals and thus protect the plant from damage. They have known roles in signaling processes such as activating of seeds during germination, pollinator attractants, and mediating plant-host symbiotic relationships, such as in legumes (14). Knowing the important roles these compounds play

in plants, it is not surprising that they are found in a wide range of species, including *Anaphalis*. However, they are typically present in such a low amount that it becomes difficult to accurately elicit structures.

A careful, thorough analysis of the NMR data is needed for accurate identification, but the general trend observed from AL#2 fractions 89-97 is consistent with a sterol. Exact identification of the sterol has not been determined due to the difficulty of analyzing the NMR data. There are numerous possibilities for substituents and, along with the large size of these types of compounds, which leads to a high number of signals in the 1-4ppm range (Appendix Figure 1).

The isolation of a sterol is, again, unsurprising. Phytosterols play important roles in membrane identity and help regulate membrane fluidity and permeability. This can help the plant cells adapt to changing temperatures, as well as changes in pH or membrane potential. Since the cell membrane acts as a barrier between the internal and external environments of the cell, it is unsurprising that phytosterols can also interact with membrane bound enzymes, regulating things such as Na^+/K^+ pumps, cellular proliferation, and cellular differentiation (15). Since the phytosterol was collected with pure hexane as the eluent, it must not be heavily oxidized, although the proton NMR does show a higher chemical shift at 7.307ppm, indicative of at least one hydroxyl group (Figure 5).
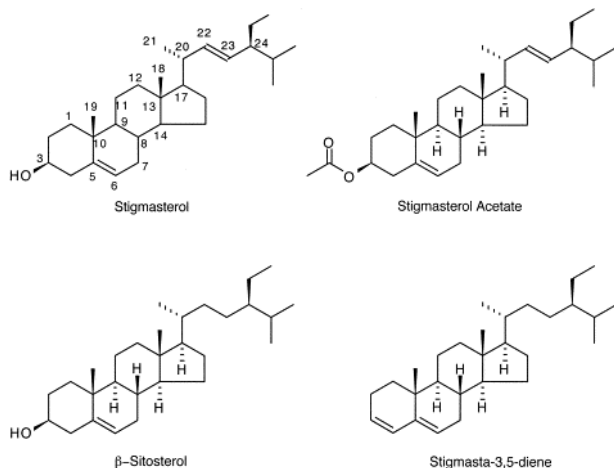
Figure 5: Examples of common phytosterols. They are mainly hydrophobic compounds present in cellular membranes and serve various functions from regulating the membrane function to interacting with enzymes, impacting signaling. The phytosterol isolated from AL#2 fractions 89-97 is most consistent with a structure similar to these, although the exact identity is still unknown (15).

The gel fraction IR showed only bands associated with C-H stretching, consistent with a wax or methyl ester of some sort. A study on the epicuticular and intracellular layer of rose leaves showed a high number of alkanes, alkenes, and methyl esters in the epicuticular layer, which is consistent with the IR. The study also found that the intracellular layer contained slightly more polar compounds, such as fatty acids and primary alcohols (16). They had a difficult time separating the intracellular layer from the epicuticular layer, so the gel that was recovered from AL#8 could possibly contain a mixture of both epicuticular and intracellular wax layers.

The gelling nature upon exposure to methanol is highly suggestive of an aliphatic hydrocarbon with maybe one or two degrees of unsaturation in the form of double bonds. This would account for the NMR data, which shows vinylic protons and carbons present in the gel (Appendix Figure 3). The reason for the gelling is due to aggregation of the long chain hydrocarbons when placed in a polar environment, such as methanol. An attempt was made to spin and heat the sample while taking an NMR with the sample in methanol-d_4, but the sample was not homogeneous enough to obtain an accurate NMR. The gel was then divided up onto several plates and different solvents were dropped over it to see what it was most soluble in. Acetonitrile-d_3 proved to be the best candidate and so the sample was re-dissolved in it and the NMR was obtained that way.

The isolation of an epicuticular wax is, again, unsurprising since the extract it was obtained from was the top extract. The leaves, stems, and flowers were all included in the top extraction, and plants use these waxes on their leaves as a barrier between the internal and external environments. Although water loss is an important issue plants must face, it is believed that the intracellular layer is more involved with this than the epicuticular wax, which is thought to serve a role in UV protection and gas exchange through the stomata. They also serve an ecological role by signaling to insects that feed off of or pollinate the plant (16).

In his paper, Bohlmann stated he was able to isolate yellow crystals (5). As of now, no yellow crystals large enough for analysis by x-ray diffraction have been isolated. Small yellow crystalline like substances have been found and attempts have been made to generate larger crystals by layering methanol and methylene choride over the sample, as well as the use of heated petroleum ether. Fraction 34 from AL#6 showed long, white crystals that are being recrystallized with the hopes of obtaining an x-ray diffraction and structurally

identifying them. This fraction eluted out in acetone, so either the crystals are an impurity in the eluent or they are extremely polar.

The columns used are either made of Al_2O_3 or silica gel, with the Al_2O_3 showing slightly better separation than the silica. Although separation of compounds with very different polarities is easily accomplished, alumina and silica are not good enough to separate compounds that are structural isomers of one another or compounds that are extremely similar. Some of the GCMS data suggests that differences causing the peaks in an impure sample are as small as the addition of a methyl group to a compound. In order to separate these compounds, either a different type of column must be employed or even smaller fractions need to be collected. The fractions collected ranged from 5mL to 50mL depending on which column was run, but perhaps fractions as small as 1mL need to be collected to ensure complete separation. The issue with collecting such small fractions is that the analyte then becomes spread over multiple fractions rather than being concentrated all in one fraction. Capillary zone electrophoresis has also been successfully used to isolate nine flavonoids with only minor structural differences from *Anaphalis margaritacea,* and may be a useful technique to try with some of the fractions collected showing various structural isomers present (16).

FUTURE DIRECTIONS

The ethyl acetate and methanol extracts still need to undergo column chromatography. Although the rule of thumb is to use 10g of solid phase for every gram of sample, a minimum of 20g of $Al_2O_{3(s)}$ should be used for the extraction, as that is what gave the best separation for the previous extracts. If using silica gel, the maximum weight the columns used previously can accommodate is 10g, so either a larger column would have to be purchased or less of the sample would have to be introduced. Silica gel would work better for separation of a fraction collected after the initial column is run on the extract.

The fractions previously collected still have to be analyzed. Fractions from AL#1-AL#4 and SC#1-SC#4 have already been analyzed and those deemed important enough for further investigation either through purity or containing compounds that may be of interest to try and isolate through GCMS analysis have been saved and labeled. These fractions should be reanalyzed for any changes since they have been left to sit for a while, and some compounds may have oxidized or decomposed. These samples are from extraction 1, so

they could also be compared to similar fractions collected from extraction 2 to see if any similar compounds were isolated.

Fractions from AL#5-AL#8 are from extraction 2 and need to be extensively analyzed. There were numerous samples that contained interesting compounds, but were not taken further. In AL#6, the column was capped off after fraction 13 was collected and left to stand overnight with ¼'' of 50% EtOAc/50% hexane mixture to ensure the column did not dry. In the morning, a yellow solid substance was seen on the tip of the column and around the stopcock. These were isolated in fraction 14, and an attempt should be made to try and obtain crystals from this fraction. Fraction 35 was especially dark yellow in color and fraction 37 had a cloudy white precipitate form as it was eluting from the column, and was also never analyzed.

AL#8 contained the gel substance, which still needs to be properly identified. Further NMR investigations are needed into this compound, as well as the possible sterol isolated from AL#2 fractions 89-97 and the flavonoid from AL#1 fraction 50. Recrystallization is underway of the crystals found, and hopefully an x-ray diffraction can be obtained from these to give an exact structure. Since Bohlmann reported yellow crystals, further samples should be investigated for crystals and layered with methanol and methylene chloride in order to try and get crystals to crash out.

If any of the initially reported compound is found through GCMS and then confirmed by NMR, crystals will need to be grown in order to obtain an x-ray diffraction and determine the stereochemistry of the chlorinated carbon. If there is any way to stress the plants pre-extraction, that is also something to consider trying since these compounds could potentially be increased in number if the plant is under distress. Any increase in amount would make them easier to isolate, since only 11mg of (E)-5-chloro-2-(octa-2,4,6-triynylidene)-5,6-dihydro-2H-pyran was extracted from 2kg of *A. margaritacea* roots. This means that the approximate 0.5kg of roots would yield, at best, roughly 2.75mg of this compound.

Another thing to consider is changing the type of column used to try and optimize the separation. Cis/trans FAMEs were successfully isolated using a silver ion SPE tube, in which the silver ion interacts more strongly with cis pi electrons than with the less accessible trans pi electrons (Figure 6) (17). This leads to the trans compounds eluting before the cis. Since the silver is acting as a pi electron acceptor, it is a reasonable

assumption that this might work for triple bonds as well and give even better separation than what is currently observed.
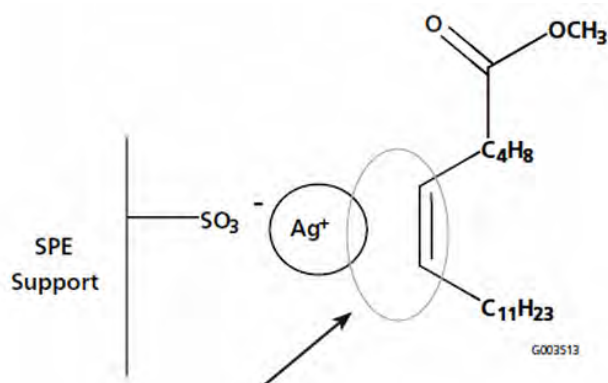
Figure 6: Stationary phase of the silver ion SPE column interacting with an analyte. The silver ion acts as a pi acceptor and interacts more strongly with trans double bonds, which have greater accessibility to their pi electrons. This type of column may help increase separation from A. margaritacea extractions.

ACKNOWLEDGEMENTS

Appendix

Fraction	Solvent	TLC
1	100% Hexane	N
2	100% Hexane	N
3	100% Hexane	N
4	100% Hexane	N
5	100% Hexane	N
6	100% Hexane	N
7	100% Hexane	N
8	5% EtOAc	N
9	5% EtOAc	N
10	5% EtOAc	N
11	5% EtOAc	N
12	5% EtOAc	N
13	5% EtOAc	N
14	10% EtOAC	N
15	10% EtOAC	N
16	10% EtOAC	N
17	10% EtOAC	N
18	10% EtOAC	N
19	10% EtOAC	N
20	25% EtOAc	N
21	25% EtOAc	N
22	25% EtOAc	N
23	25% EtOAc	N
24	25% EtOAc	N
25	25% EtOAc	N
26	35% EtOAc	N
27	35% EtOAc	N
28	35% EtOAc	N

29	35% EtOAc	N
30	35% EtOAc	N
31	35% EtOAc	slight
32	45% EtOAc	slight
33	45% EtOAc	slight
34	45% EtOAc	slight
35	45% EtOAc	slight
36	45% EtOAc	N
37	45% EtOAc	N
38	50% EtOAc	N
39	50% EtOAc	slight
40	50% EtOAc	slight
41	50% EtOAc	Y
42	50% EtOAc	Y
43	50% EtOAc	Y
44	60% EtOAc	Y
45	60% EtOAc	Y
46	60% EtOAc	Y
47	60% EtOAc	Y
48	60% EtOAc	Y
49	60% EtOAc	Y
50	70% EtOAc	Y
51	70% EtOAc	Y
52	70% EtOAc	Y
53	70% EtOAc	Y
54	70% EtOAc	Y
55	70% EtOAc	slight
56	80% EtOAc	slight
57	80% EtOAc	slight

58	80% EtOAc	slight
59	80% EtOAc	slight
60	80% EtOAc	N
61	80% EtOAc	N
62	90% EtOAc	N
63	90% EtOAc	N
64	90% EtOAc	N
65	90% EtOAc	N
66	90% EtOAc	N

67	90% EtOAc	N
68	100% EtOAc	N
69	100% EtOAc	N
70	100% EtOAc	N
71	100% EtOAc	N
72	100% EtOAc	N
73	100% EtOAc	N
M	Methanol	N

Table 1: Fractions collected from alumina column 1, which used 10g of Al_2O_3 as the stationary phase and the polar root extract from extract 1 dried onto 1.0120g of Al_2O_3.

Fraction	Solvent	TLC
1	60% EtOAc	N
2	60% EtOAc	Y
3	60% EtOAc	Y
4	60% EtOAc	Y
5	60% EtOAc	Y
6	60% EtOAc	Y
7	60% EtOAc	N
8	60% EtOAc	N
9	60% EtOAc	N
10	60% EtOAc	N
11	60% EtOAc	N
12	60% EtOAc	N
13	60% EtOAc	N
14	60% EtOAc	N
15	60% EtOAc	N
16	60% EtOAc	N
17	60% EtOAc	N
18	60% EtOAc	N
19	60% EtOAc	slight
20	60% EtOAc	slight
21	60% EtOAc	N
22	60% EtOAc	N
23	100% EtOAc	N
24	100% EtOAc	slight
25	100% EtOAc	slight
26	100% EtOAc	slight
27	100% EtOAc	slight
28	100% EtOAc	slight
29	100% EtOAc	slight

*Table 2: Fractions collected from silica column 1, which used 5g of silica gel as the
stationary phase and recombined fractions 46-71 from alumina column 1.*

Fraction	Solvent	TLC
1	5% MeOH/95% dichloromethane	Y
2	5% MeOH/95% dichloromethane	Y
3	5% MeOH/95% dichloromethane	Y
4	5% MeOH/95% dichloromethane	Y
5	5% MeOH/95% dichloromethane	Y
6	5% MeOH/95% dichloromethane	Y
7	5% MeOH/95% dichloromethane	Y
8	5% MeOH/95% dichloromethane	Y
9	5% MeOH/95% dichloromethane	Y
10	5% MeOH/95% dichloromethane	Y
11	5% MeOH/95% dichloromethane	Y
12	5% MeOH/95% dichloromethane	Y
13	5% MeOH/95% dichloromethane	Y
14	5% MeOH/95% dichloromethane	Y
15	5% MeOH/95% dichloromethane	Y
16	5% MeOH/95% dichloromethane	Y
17	5% MeOH/95% dichloromethane	Y
18	5% MeOH/95% dichloromethane	Y
19	5% MeOH/95% dichloromethane	Y
20	5% MeOH/95% dichloromethane	Y
21	5% MeOH/95% dichloromethane	Y
22	5% MeOH/95% dichloromethane	Y
23	5% MeOH/95% dichloromethane	Y
24	5% MeOH/95% dichloromethane	Y
25	5% MeOH/95% dichloromethane	Y
26	5% MeOH/95% dichloromethane	Y
27	5% MeOH/95% dichloromethane	Y

Table 3: Fractions collected from silica column 2, which used 10g of silica gel as the stationary phase and recombined fractions from silica column 1.

Fraction	Solvent	TLC		Fraction	Solvent	TLC
1	Hexane	Y		28	Hexane	N
2	Hexane	Y		29	Hexane	N
3	Hexane	Y		30	Hexane	N
4	Hexane	Y		31	Hexane	N
5	Hexane	Y		32	Hexane	N
6	Hexane	Y		33	Hexane	N
7	Hexane	Y		34	Hexane	N
8	Hexane	Y		35	Hexane	N
9	Hexane	Y		36	Hexane	N
10	Hexane	Y		37	Hexane	N
11	Hexane	Y		38	Hexane	N
12	Hexane	Y		39	Hexane	N
13	Hexane	Y		40	Hexane	N
14	Hexane	Y		41	Hexane	N
15	Hexane	Y		42	Hexane	N
16	Hexane	Y		43	Hexane	N
17	Hexane	Y		44	Hexane	N
18	Hexane	Y		45	Hexane	N
19	Hexane	Y		46	Hexane	N
20	Hexane	Y		47	Hexane	N
21	Hexane	Y		48	Hexane	N
22	Hexane	Y		49	Hexane	N
23	Hexane	Y		50	Hexane	N
24	Hexane	Y		51	1% Ether	N
25	Hexane	Y		52	1% Ether	N
26	Hexane	N		53	1% Ether	N
27	Hexane	N		54	1% Ether	N
				55	1% Ether	N

56	1% Ether	Y		86	1% Ether	Y
57	1% Ether	Y		87	1% Ether	Y
58	1% Ether	Y		88	1% Ether	Y
59	1% Ether	Y		89	1% Ether	Y
60	1% Ether	Y		90	2% Ether	Y
61	1% Ether	Y		91	2% Ether	Y
62	1% Ether	Y		92	2% Ether	Y
63	1% Ether	Y		93	2% Ether	Y
64	1% Ether	Y		94	2% Ether	Y
65	1% Ether	Y		95	2% Ether	Y
66	1% Ether	Y		96	2% Ether	Y
67	1% Ether	Y		97	2% Ether	Y
68	1% Ether	Y		98	2% Ether	Y
69	1% Ether	Y		99	2% Ether	Y
70	1% Ether	Y		100	2% Ether	Y
71	1% Ether	Y		101	2% Ether	Y
72	1% Ether	Y		102	2% Ether	Y
73	1% Ether	Y		103	2% Ether	Y
74	1% Ether	Y		104	2% Ether	Y
75	1% Ether	Y		105	2% Ether	Y
76	1% Ether	Y		106	2% Ether	Y
77	1% Ether	Y		107	2% Ether	Y
78	1% Ether	Y		108	2% Ether	N
79	1% Ether	Y		109	2% Ether	N
80	1% Ether	Y		110	2% Ether	N
81	1% Ether	Y		111	2% Ether	N
82	1% Ether	Y		112	2% Ether	N
83	1% Ether	Y		113	2% Ether	N
84	1% Ether	Y		114	2% Ether	N
85	1% Ether	Y		115	2% Ether	N

116	2% Ether	N		146	10% Ether	Y
117	2% Ether	N		147	10% Ether	Y
118	3% Ether	N		148	10% Ether	Y
119	3% Ether	N		149	10% Ether	Y
120	3% Ether	N		150	10% Ether	Y
121	3% Ether	N		151	10% Ether	Y
122	3% Ether	N		152	10% Ether	Y
123	3% Ether	N		153	10% Ether	Y
124	3% Ether	N		154	10% Ether	Y
125	3% Ether	N		155	25% Ether	Y
126	3% Ether	N		156	25% Ether	Y
127	3% Ether	N		157	25% Ether	Y
128	3% Ether	N		158	25% Ether	Y
129	3% Ether	N		159	25% Ether	Y
130	3% Ether	N		160	25% Ether	Y
131	3% Ether	N		161	25% Ether	Y
132	3% Ether	N		162	25% Ether	Y
133	3% Ether	N		163	35% Ether	Y
134	3% Ether	N		164	35% Ether	Y
135	3% Ether	N		165	35% Ether	Y
136	3% Ether	Y		166	35% Ether	Y
137	3% Ether	Y		167	35% Ether	Y
138	3% Ether	Y		168	35% Ether	Y
139	3% Ether	Y		169	35% Ether	Y
140	3% Ether	Y		170	35% Ether	Y
141	3% Ether	Y		171	35% Ether	Y
142	10% Ether	Y		172	35% Ether	Y
143	10% Ether	Y		173	50% Ether	Y
144	10% Ether	Y		174	50% Ether	Y
145	10% Ether	Y		175	50% Ether	Y

176	50% Ether	Y		206	50% Ether	Y
177	50% Ether	Y		207	50% Ether	Y
178	50% Ether	Y		208	75% Ether	Y
179	50% Ether	Y		209	75% Ether	Y
180	50% Ether	Y		210	75% Ether	Y
181	50% Ether	Y		211	75% Ether	Y
182	50% Ether	Y		212	75% Ether	Y
183	50% Ether	Y		213	75% Ether	Y
184	50% Ether	Y		214	75% Ether	Y
185	50% Ether	Y		215	75% Ether	Y
186	50% Ether	Y		216	75% Ether	Y
187	50% Ether	Y		217	75% Ether	Y
188	50% Ether	Y		218	75% Ether	Y
189	50% Ether	Y		219	75% Ether	Y
190	50% Ether	Y		220	75% Ether	Y
191	50% Ether	Y		221	75% Ether	Y
192	50% Ether	Y		222	75% Ether	Y
193	50% Ether	Y		223	75% Ether	Y
194	50% Ether	Y		224	75% Ether	Y
195	50% Ether	Y		225	75% Ether	Y
196	50% Ether	Y		226	75% Ether	Y
197	50% Ether	Y		227	100% Ether	Y
198	50% Ether	Y		228	100% Ether	Y
199	50% Ether	Y		229	100% Ether	Y
200	50% Ether	Y		230	100% Ether	Y
201	50% Ether	Y		231	100% Ether	Y
202	50% Ether	Y		232	100% Ether	Y
203	50% Ether	Y		233	100% Ether	Y
204	50% Ether	Y		234	100% Ether	Y
205	50% Ether	Y		235	100% Ether	Y

236	100% Ether	Y		256	Methanol	Y
237	100% Ether	Y		257	Methanol	Y
238	100% Ether	Y		258	Methanol	Y
239	100% Ether	Y		259	Methanol	Y
240	100% Ether	Y		260	Methanol	Y
241	100% Ether	Y		261	Methanol	Y
242	100% Ether	Y		262	Methanol	Y
243	100% Ether	Y		263	Methanol	Y
244	100% Ether	Y		264	Methanol	Y
245	100% Ether	Y		265	Methanol	Y
246	100% Ether	Y		266	Methanol	Y
247	100% Ether	Y		267	Methanol	Y
248	100% Ether	Y		268	Methanol	Y
249	100% Ether	Y		269	Methanol	Y
250	100% Ether	Y		270	Methanol	Y
251	100% Ether	Y		271	Methanol	Y
252	Methanol	Y		272	Methanol	Y
253	Methanol	Y		273	Methanol	Y
254	Methanol	Y		274	Methanol	Y
255	Methanol	Y		275	Methanol	Y

Table 4: Fractions collected from alumina column 2, which used 25g of Al_2O_3 as the stationary phase and the nonpolar root extract from extract 1.

Fraction	Solvent	TLC
1	Hexane	N
2	Hexane	N
3	Hexane	N
4	Hexane	N
5	Hexane	N
6	Hexane	N
7	Hexane	N
8	Hexane	N
9	Hexane	N
10	Hexane	N
11	Hexane	N
12	Hexane	N
13	Hexane	N
14	Hexane	N
15	Hexane	N
16	Hexane	N
17	Hexane	N
18	Hexane	N
19	Hexane	N
20	Hexane	N
21	Hexane	N
22	Hexane	N
23	Hexane	N
24	Hexane	N
25	Hexane	N
26	Hexane	N
27	Hexane	N
28	Hexane	N
29	Hexane	N

30	Hexane	Y
31	Hexane	N
32	Hexane	N
33	Hexane	N
34	Hexane	N
35	Hexane	N
36	Hexane	N
37	Hexane	N
38	Hexane	N
39	Hexane	N
40	Hexane	N
41	Hexane	N
42	Hexane	N
43	Hexane	N
44	Hexane	N
45	Hexane	N
46	Hexane	Y
47	Hexane	N
48	Hexane	N
49	Hexane	N
50	Hexane	N
51	Hexane	N
52	Hexane	N
53	Hexane	N
54	Hexane	slight
55	Hexane	slight
56	Hexane	N
57	Hexane	N
58	Hexane	Y
59	Hexane	N

60	Hexane	N
61	Hexane	N
62	Hexane	N

63	Hexane	N
64	Hexane	N

Table 5: Fractions collected from alumina column 3, which used 10g of Al$_2$O$_3$ as the stationary phase and recombined fractions 30-36 from alumina column 2.

Fraction	Solvent	TLC
1	2% Ether	Y
2	2% Ether	Y
3	2% Ether	Y
4	2% Ether	Y
5	2% Ether	Y
6	2% Ether	Y
7	2% Ether	Y
8	2% Ether	Y
9	2% Ether	Y
10	2% Ether	Y
11	2% Ether	Y
12	2% Ether	Y
13	2% Ether	Y
14	2% Ether	Y
15	2% Ether	Y
16	100% Ether	Y
17	100% Ether	Y

Table 6: Fractions collected from alumina column 4, which used 10g of Al$_2$O$_3$ as the stationary phase and recombined fractions 1-12 from alumina column 2.

Fraction	Solvent	TLC
1	Hexane	Y
2	Hexane	Y
3	Hexane	Y
4	Hexane	Y
5	Hexane	Y
6	Hexane	slight
7	Hexane	slight
8	Hexane	N
9	Hexane	N
10	Hexane	N
11	Hexane	N
12	Hexane	N
13	Hexane	N
14	Hexane	N
15	Hexane	N
16	Hexane	N
17	Hexane	N
18	Hexane	N
19	Hexane	N
20	Hexane	N
21	Hexane	N
22	Hexane	N
23	Hexane	N
24	Hexane	N
25	2% Ether	N
26	2% Ether	N
27	2% Ether	N
28	2% Ether	Y
29	2% Ether	Y
30	2% Ether	Y
31	2% Ether	Y

32	2% Ether	slight
33	2% Ether	slight
34	2% Ether	N
35	2% Ether	N
36	2% Ether	N
37	2% Ether	Y
38	Ether	Y
39	Ether	Y
40	Ether	Y
41	Ether	slight
42	Ether	slight

Table 7: Fractions collected from silica column 3, which used 5g of silica as the stationary phase and recombined fractions 1-12 from alumina column 2 (all fractions from alumina column 4).

Fraction	Solvent
1	20% EtOAc
2	20% EtOAc
3	20% EtOAc
4	20% EtOAc
5	20% EtOAc
6	20% EtOAc
7	20% EtOAc
8	20% EtOAc
9	20% EtOAc
10	20% EtOAc
11	20% EtOAc
12	20% EtOAc
13	20% EtOAc
14	20% EtOAc
15	20% EtOAc
16	20% EtOAc

17	20% EtOAc
18	20% EtOAc
19	20% EtOAc
20	20% EtOAc
21	20% EtOAc
22	20% EtOAc
23	30% EtOAc
24	30% EtOAc
25	30% EtOAc
26	30% EtOAc
27	30% EtOAc
28	30% EtOAc
29	30% EtOAc
30	30% EtOAc
31	30% EtOAc
32	30% EtOAc
33	30% EtOAc

34	40% EtOAc		66	60% EtOAc
35	40% EtOAc		67	60% EtOAc
36	40% EtOAc		68	60% EtOAc
37	40% EtOAc		69	60% EtOAc
38	40% EtOAc		70	60% EtOAc
39	40% EtOAc		71	60% EtOAc
40	40% EtOAc		72	60% EtOAc
41	40% EtOAc		73	70% EtOAc
42	40% EtOAc		74	70% EtOAc
43	40% EtOAc		75	70% EtOAc
44	40% EtOAc		76	70% EtOAc
45	50% EtOAc		77	70% EtOAc
46	50% EtOAc		78	70% EtOAc
47	50% EtOAc		79	70% EtOAc
48	50% EtOAc		80	70% EtOAc
49	50% EtOAc		81	70% EtOAc
50	50% EtOAc		82	70% EtOAc
51	50% EtOAc		83	70% EtOAc
52	50% EtOAc		84	70% EtOAc
53	60% EtOAc		85	70% EtOAc
54	60% EtOAc		86	80% EtOAc
55	60% EtOAc		87	80% EtOAc
56	60% EtOAc		88	80% EtOAc
57	60% EtOAc		89	80% EtOAc
58	60% EtOAc		90	80% EtOAc
59	60% EtOAc		91	80% EtOAc
60	60% EtOAc		92	80% EtOAc
61	60% EtOAc		93	80% EtOAc
62	60% EtOAc		94	80% EtOAc
63	60% EtOAc		95	80% EtOAc
64	60% EtOAc		96	80% EtOAc
65	60% EtOAc		97	80% EtOAc

98	90% EtOAc		121	EtOAc
99	90% EtOAc		122	20% MeOH/80% EtOAc
100	90% EtOAc		123	20% MeOH/80% EtOAc
101	90% EtOAc		124	20% MeOH/80% EtOAc
102	90% EtOAc		125	20% MeOH/80% EtOAc
103	90% EtOAc		126	20% MeOH/80% EtOAc
104	90% EtOAc		127	20% MeOH/80% EtOAc
105	90% EtOAc		128	20% MeOH/80% EtOAc
106	90% EtOAc		129	20% MeOH/80% EtOAc
107	90% EtOAc		130	20% MeOH/80% EtOAc
108	90% EtOAc		131	50% MeOH/80% EtOAc
109	90% EtOAc		132	50% MeOH/80% EtOAc
110	90% EtOAc		133	50% MeOH/80% EtOAc
111	EtOAc		134	50% MeOH/80% EtOAc
112	EtOAc		135	50% MeOH/80% EtOAc
113	EtOAc		136	50% MeOH/80% EtOAc
114	EtOAc		137	50% MeOH/80% EtOAc
115	EtOAc		138	Methanol
116	EtOAc		139	Methanol
117	EtOAc		140	Methanol
118	EtOAc		141	Methanol
119	EtOAc		142	Methanol
120	EtOAc		143	Methanol

Table 8: Fractions collected from silica column 4, which used 8g of silica as the stationary phase and recombined fractions 253-271 from alumina column 2.

Fraction	Solvent	TLC
1	Hexane	Y
2	Hexane	Y
3	Hexane	Y
4	Hexane	Y
5	1% ether	Y
6	1% ether	Y
7	1% ether	Y
8	1% ether	Y
9	2% ether	Y
10	2% ether	Y
11	2% ether	Y
12	2% ether	Y
13	10% ether	Y
14	10% ether	Y
15	10% ether	Y
16	10% ether	Y
17	25% ether	N
18	25% ether	N
19	25% ether	N
20	50% ether	N
21	50% ether	N
22	50% ether	Y
23	60% ether	N
24	60% ether	N
25	80% ether	N
26	80% ether	N
27	5% methanol	N
28	5% methanol	N
29	5% methanol	Y
30	10% methanol	Y

31	10% methanol	Y
32	Methanol	Y

Table 9: Fractions collected from alumina column 5, which used 20g of Al_2O_3 as the stationary phase and the nonpolar root extract from extract 2.

Fraction	Solvent
1	hexane
1	methanol
2	hexane
2	methanol
3	hexane
4	hexane
5	hexane
6	hexane
7	25% EtOAc
8	25% EtOAc
9	25% EtOAc
10	25% EtOAc
11	25% EtOAc
12	50% EtOAc
13	50% EtOAc
14	50% EtOAc
15	50% EtOAc
16	50% EtOAc
17	75% EtOAc
18	75% EtOAc
19	75% EtOAc
20	75% EtOAc
21	75% EtOAc
22	EtOAc
23	EtOAc

24	EtOAc
25	EtOAc
26	50% EtOAC/50% Acetone
27	50% EtOAC/50% Acetone
28	50% EtOAC/50% Acetone
29	50% EtOAC/50% Acetone
30	50% EtOAC/50% Acetone
31	50% EtOAC/50% Acetone
32	50% EtOAC/50% Acetone
33	acetone
34	acetone
35	acetone
36	acetone
37	methanol
38	methanol
39	methanol
40	methanol

Table 10: Fractions collected from alumina column 6, which used 20g of Al_2O_3 as the stationary phase and the polar root extract from extract 2.

Fraction	Solvent	TLC			
1	hexane	N	13	hexane	N
2	hexane	N	14	hexane	N
3	hexane	N	15	hexane	N
4	hexane	N	16	hexane	N
5	hexane	N	17	hexane	N
6	hexane	N	18	hexane	N
7	hexane	N	19	hexane	N
8	hexane	N	20	hexane	N
9	hexane	N	21	hexane	N
10	hexane	N	22	hexane	N
11	hexane	N	23	hexane	N
12	hexane	N	24	hexane	N
			25	hexane	N

26	hexane	N		46	10% Ether	N
27	hexane	N		47	10% Ether	N
28	hexane	N		48	10% Ether	N
29	hexane	N		49	10% Ether	N
30	hexane	N		50	10% Ether	N
31	hexane	N		51	10% Ether	N
32	hexane	N		52	50% Ether	N
33	hexane	N		53	50% Ether	N
34	hexane	N		54	50% Ether	N
35	1% Ether	N		55	50% Ether	N
36	1% Ether	N		56	50% Ether	N
37	1% Ether	N		57	50% Ether	slight
38	1% Ether	N		58	Ether	Y
39	1% Ether	N		59	Ether	Y
40	5% Ether	N		60	Ether	Y
41	5% Ether	N		61	Ether	Y
42	5% Ether	N		62	Ether	Y
43	5% Ether	N		63	Ether	Y
44	5% Ether	N		64	Ether	Y
45	5% Ether	N		65	Ether	slight

Table 11: Fractions collected from alumina column 7, which used 20g of Al_2O_3 as the stationary phase and recombined fractions 1-5 from alumina column 6.

Fraction	Solvent
1	hexane
2	hexane
3	hexane
4	hexane
5	hexane
6	hexane
7	hexane
8	5% EtOAc
9	5% EtOAc
10	5% EtOAc
11	5% EtOAc
12	5% EtOAc
13	5% EtOAc
14	5% EtOAc
15	5% EtOAc
16	10% EtOAc
17	10% EtOAc
18	10% EtOAc
19	10% EtOAc
20	10% EtOAc
21	10% EtOAc
22	50% EtOAc
23	50% EtOAc

Fraction	Solvent
24	50% EtOAc
25	50% EtOAc
26	50% EtOAc
27	50% EtOAc
28	50% EtOAc
29	50% EtOAc
30	80% EtOAc
31	80% EtOAc
32	80% EtOAc
33	80% EtOAc
34	80% EtOAc
35	80% EtOAc
36	80% EtOAc
37	EtOAc
38	EtOAc
39	EtOAc
40	EtOAc
41	EtOAc
42	EtOAc

43	EtOAc
44	EtOAc
45	EtOAc
46	EtOAc
47	EtOAc
48	EtOAc
49	EtOAc
50	EtOAc
51	EtOAc
52	EtOAc
53	EtOAc
54	EtOAc
55	EtOAc
56	EtOAc
57	EtOAc
58	EtOAc
59	EtOAc
60	EtOAc
61	EtOAc
62	EtOAc
63	methanol

Table 12: Fractions collected from alumina column 8, which used 20g of Al_2O_3 as the stationary phase and the nonpolar top extract.

Date	Contents	Solvent	GCMS Settings	Results
9/27/11	alumina column 1 vial 50	acetone-d	50°C for 15 min ramp at 10°C/min to 150°C ramp at 10°C/min to 250°C hold for 15 min	very impure - possible GCMS contamination
9/28/11	p-methoxychloroethyl benzene	MeOH	same as above	impurities seen, re run
9/28/11			ran column at 300°C for approx 3 hours	
9/28/11	alumina column 1 vial 50	acetone-d	50°C for 15 min ramp at 10°C/min to 150°C ramp at 10°C/min to 250°C hold for 15 min	very impure. Didn't see peak with RT at 30min as in previous run
9/29/11			ran column at 300°C and injector inlet at 400°C for approx 1/2 hour	
9/30/11	methanol	n/a	300°C for 15 min	clean
30-Sep	methanol	n/a	50°C for 15 min	clean
9/30/11	p-	MeOH	50°C for 15 min,	very impure

1	methoxychl oroethyl benzene		ramp at 5°C/min to 150C, hold for 15 min	
10/4/1 1	alumina column 3 vial 54	MeOH	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	possibly pure, just isomers of the same compound
10/4/1 1	alumina column 2 vial 162	MeOH	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	approx 5-10 compounds at low concentrations
10/4/1 1	alumina column 2 vial 182	MeOH	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	approx 5-10 compounds at low concentrations
10/20/ 11	ether	n/a	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	
10/20/ 11	petroleum ether	n/a	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	
10/20/ 11	first root extract	2:1 pet ether: ether	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	
10/27/ 11	ether	n/a	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	still impure

10/27/11	crude MeOH extract sample	MeOH	50°C for 15 min, ramp at 5°C/min to 150C, hold for 15 min	small amounts of several polar compounds, maybe rerun at higher temp
1/17/12	EtOAc	EtOAc	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	pure
1/17/12	petroleum ether extract from roots	2:1 pet eth:eth	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	tons of peaks
1/17/12	EtOAc extract from roots	EtOAc	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	tons of peaks
1/17/12	MeOH extract from roots	MeOH	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	tons of peaks
2/1/12	hexane	hexane	50°C for 15 min ramp at 10°C/min to 150°C ramp at	pure

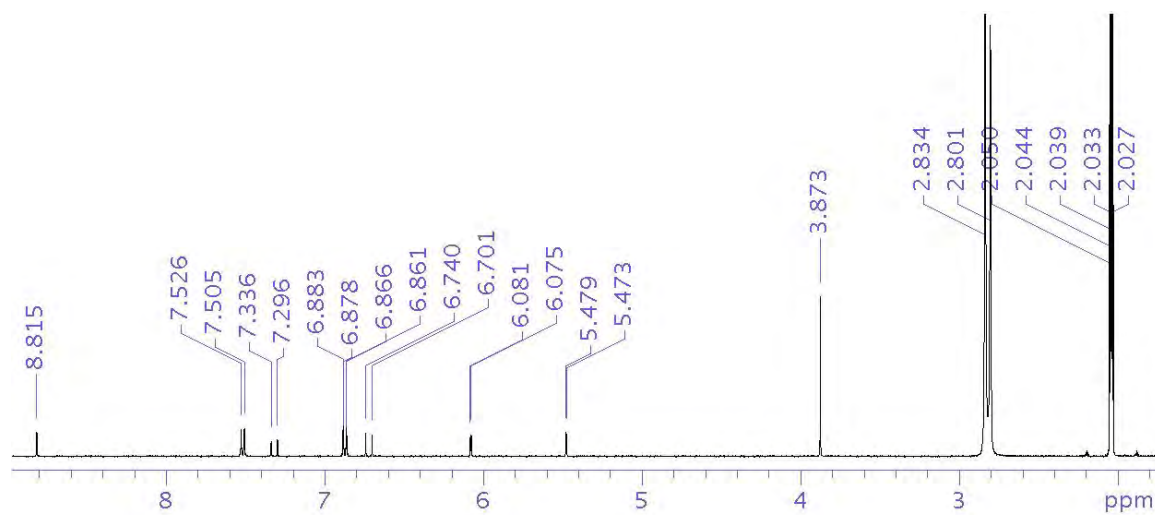|			15°C/min to 250°C hold for 15 min	
2/1/12	alumina column 5 fraction 5	1% ether in hexane	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	mainly long hydocarbon chains, various isomers
2/1/12	alumina column 5 fraction 22	50% ether in hexane	50°C for 15 min ramp at 10°C/min to 150°C ramp at 15°C/min to 250°C hold for 15 min	impure. Lots of ring systems - sterol derivatives?
2/9/12	alumina column 5 fraction 1	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/9/12	alumina column 5 fraction 2	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/9/12	alumina column 5 fraction 3	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to	

			250°C hold for 10 min	
2/9/12	alumina column 5 fraction 4	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/14/1 2	alumina column 5 fraction 6	1% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/14/1 2	alumina column 5 fraction 7	1% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/14/1 2	alumina column 5 fraction 8	1% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/14/1 2	alumina column 5 fraction 9	2% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10	

			min	
2/14/12	alumina column 5 fraction 10	2% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/22/12	alumina column 5 fraction 12	2% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/22/12	alumina column 5 fraction 13	10% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/22/12	alumina column 5 fraction 14	10% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/22/12	alumina column 5 fraction 15	10% ether in hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	

2/23/12	alumina column 6 fraction H2	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/23/12	alumina column 6 fraction 3	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/23/12	alumina column 6 fraction 4	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/23/12	alumina column 6 fraction 5	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/23/12	alumina column 6 fraction 14	50% EtoAC/50% Hex	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/28/1	alumina	hexane	100°C for 15 min	

	2	column 7 fraction 1		ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/28/1	2	alumina column 7 fraction 2	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/28/1	2	alumina column 7 fraction 3	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/28/1	2	alumina column 7 fraction 4	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	
2/28/1	2	alumina column 7 fraction 5	hexane	100°C for 15 min ramp at 15°C/min to 150°C ramp at 15°C/min to 250°C hold for 10 min	

Table 13: GCMS Data from various runs on different samples. The method used is listed, as are any relevant results.

Date	Contents	Solvent	NMR Tests Run	Results
7/26/11	Silica Column 1 Vial 6	CDCl3	Proton, Carbon, gCOSY	very dilute. No interpretation possible
7/26/11	Silica Column 2 Vial 12	CDCl3	Proton, Carbon, gCOSY	very dilute. No interpretation possible
7/26/11	Recombined Alumina Column 2 Vials 39-40	CDCl3	Proton, Carbon, HSQCAD, gCOSY	has multiple peaks below 3ppm. Possibly consistant with sterol? May still be contaminated
9/8/11	alumina column 1 vial 50	acetone-D	Proton, carbon, gCOSY, HSQC, TOCSY	nothing on carbon, proton looks good. Investigated in scifinder
9/20/11	alumina column 1 vial 49	CDCl3	Proton, Carbon, HSQCAD	multiple peaks, impure
22-Sep	alumina column 1 vial 51	CDCl3	Proton, Carbon, HSQCAD, gCOSY	multiple peaks, impure - seemed to match up with 49
10/4/11	alumina column 3 vial 54	CDCl3	Proton	very dilute. Unable to tell If pure
11/8/11	extract #3 from 2nd extraction	CD3OD	Proton, Carbon, HSQC	proton and carbon yielded results with 3-4 compounds
11/29/11	50% EtOAc/50% Hex	CD3OD	proton, carbon	proton showed numerous compounds, carbon was mostly flatline

	washing of MeOH extract			
11/29/11	MeOH extract	DMSO	Proton, Carbon, HSQC	proton and carbon were very messy - tons of compounds present

Table 14: NMR data from various samples. The tests run and relevant results are shown.
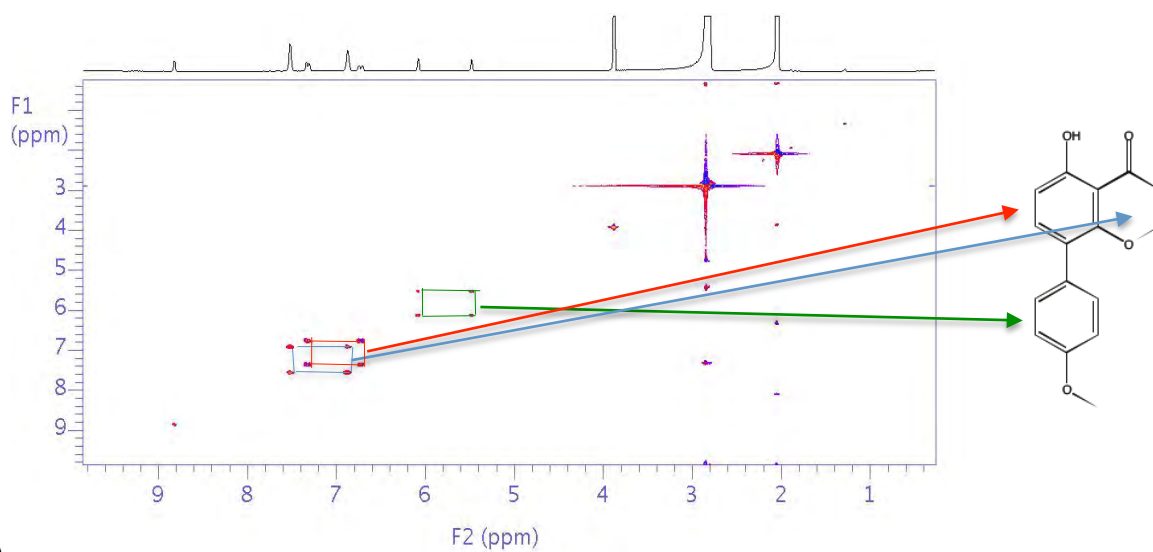

(a)

(b)

(c)

These protons should be coupled to one another, but the sample was very dilute so it most likely did not show up

These protons are coupled to one another, so they must be located next to one another
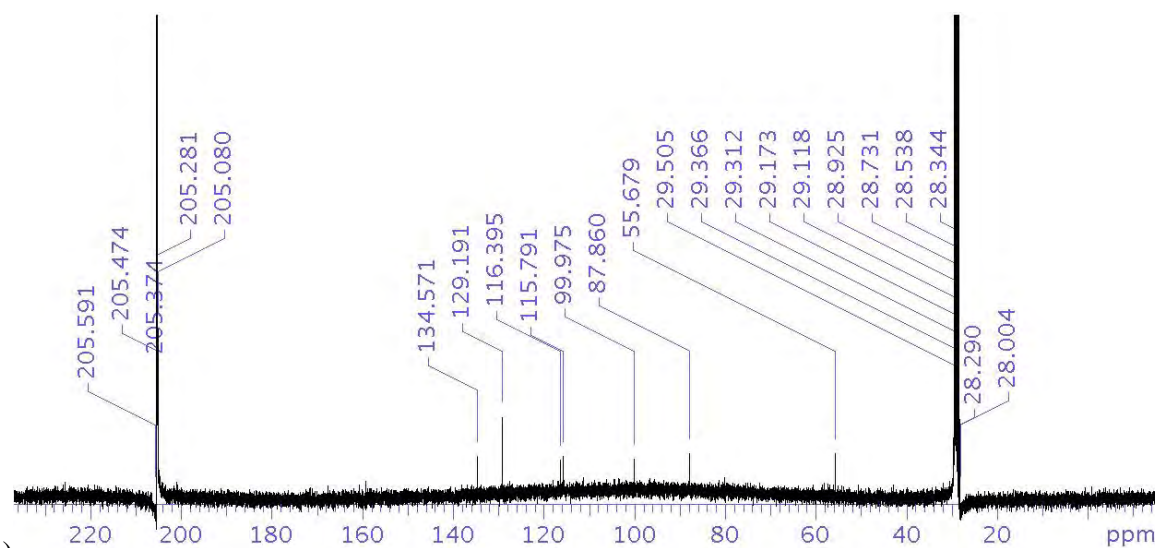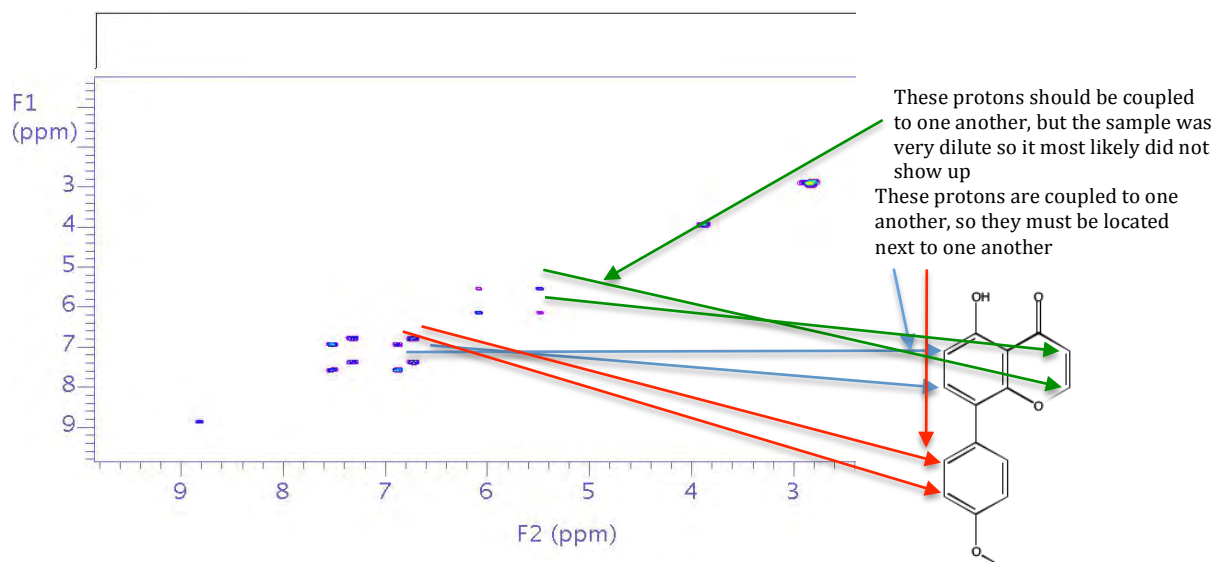
(d)

Figure 1: (a) ^1H NMR (b) ^{13}C NMR (c) COSY (d) TOCSY spectra of AL#1 fraction 50, whose proposed structure is shown next to the COSY and TOCSY with appropriate spin system

designations.

(a)
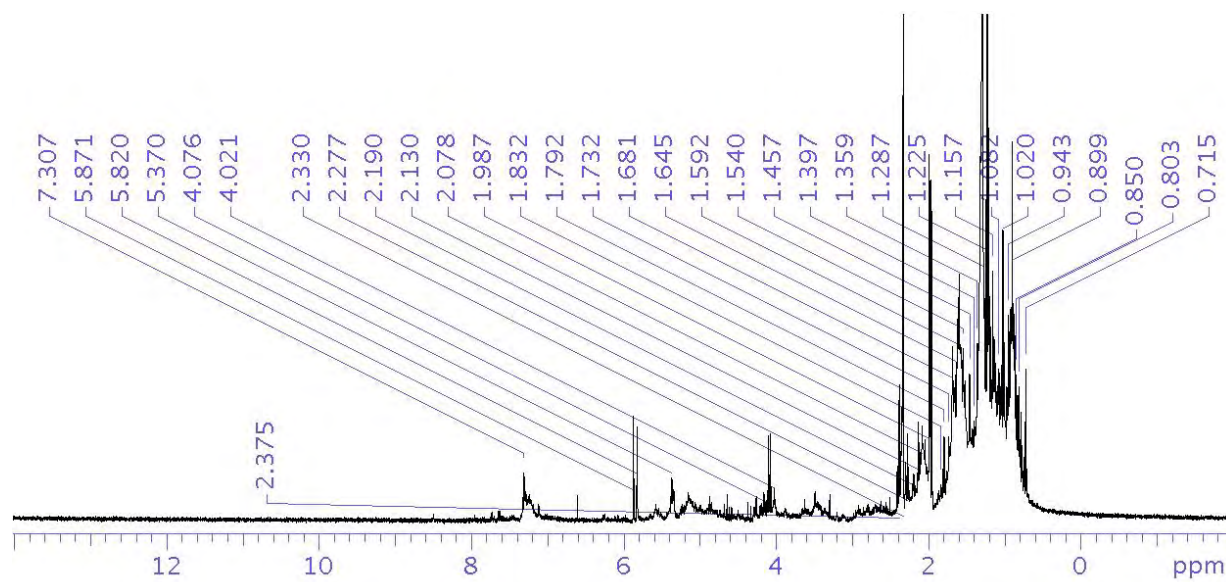
(b)
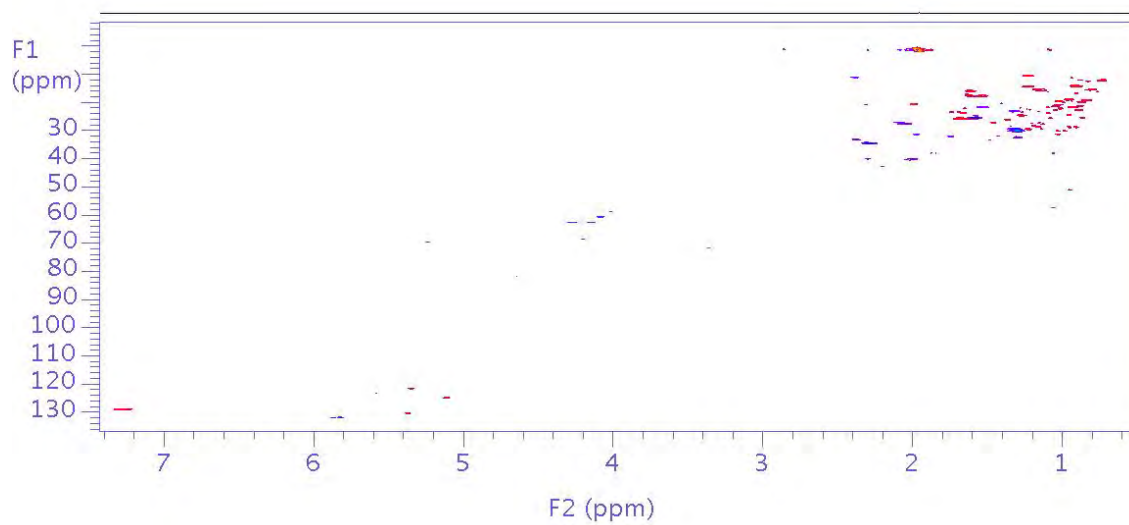
Figure 2: (a)[1]*H NMR and (b) HSQCAD data from AL#2 fractions 89-97. This is indicative of some sort of phytosterol, but further investigation is necessary in order to elucidate an exact structure.*

(a)

(b)

(c)
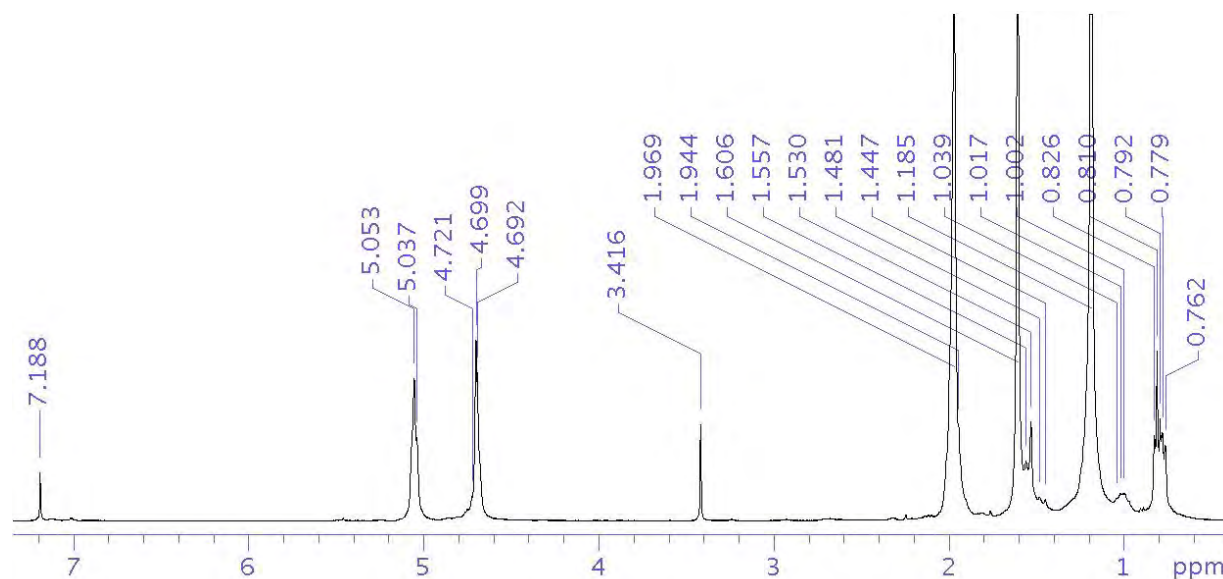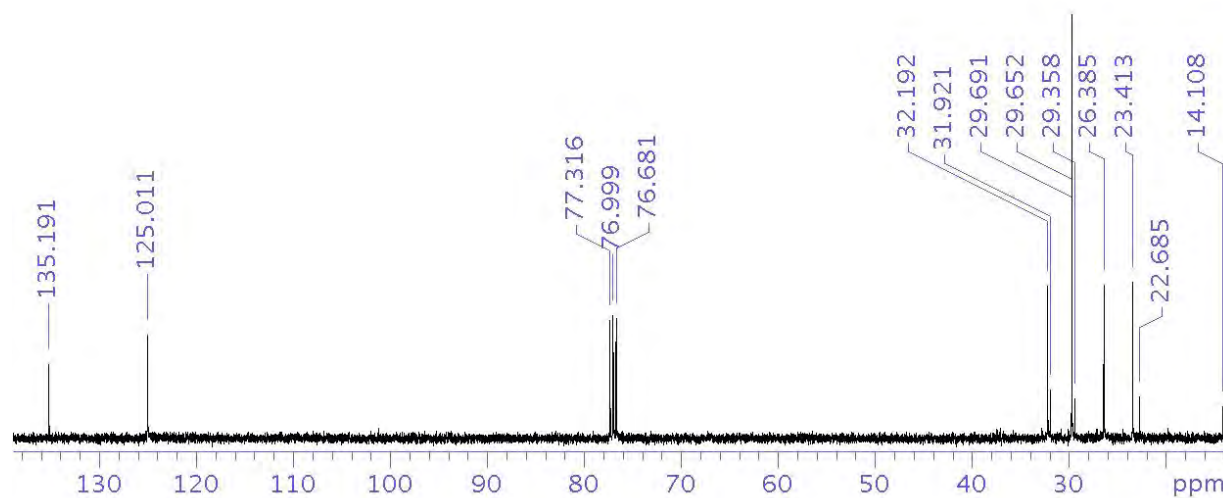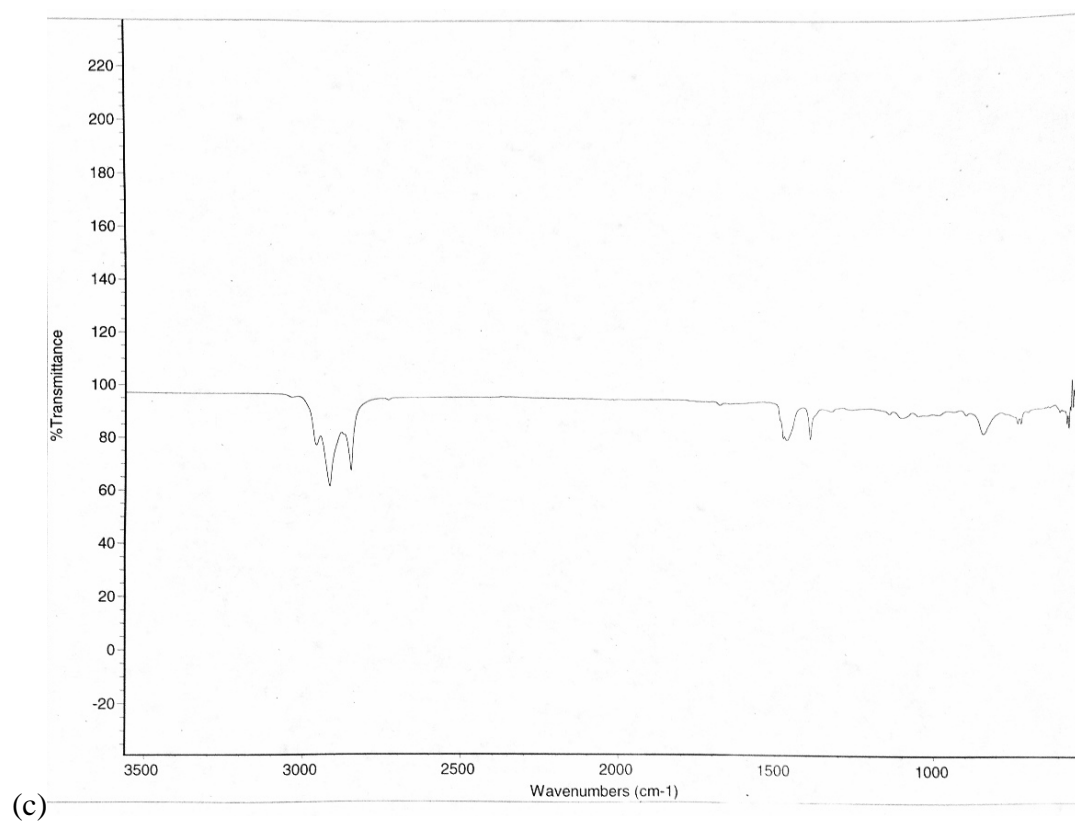
Figure 3: (a) ¹H NMR, (b) ¹³C NMR, and (c) IR data from AL#8 fraction 1, which gelled upon exposure to methanol. The data is suggestive of some sort of long chain hydrocarbon, which is typically seen in epicuticular waxes.

Works Cited

"Biosynthesis of Phenylpropanoids—Reference Pathway." *Kegg Encyclopedia of Genes and Genomes*, 1 Oct 2012. Web. 25 Apr 2012.

http://www.genome.jp/kegg/pathway/map/map01061.html.

Bohlmann, F., Arndt, C., & Zdero, C. *Chem. Ber.* 99. (1966):1648–1651. Print.

Bohlmann, F., & Weber, R. "Synthese von Enolatherpolyinen aus *Anaphalis* aus *Gnaphalium Arten*." *Chem. Ber.* 105. (1972): 3036–3040. Print.

Bok, H. B, et al. *J. Nat. Prod.* 63. (2000): 1511–1544. Print.

Britt, P. F, et al. "Mechanistic investigation into the formation of polycyclic aromatic hydrocarbons from the pyrolysis of plant steroids." *Fuel*. 80. (2001):1727–1746. Print.

Engvild, Kjeld C. "Chlorine Containing Compounds in Higher Plants." *Phytochemistry.* 25. (1985): 781–791. Print.

"High Throughput Screening in Review." *Innovaro Pharmlicensing*. Web. 14 April 2012. http://pharmalicensing.com/public/articles/view/1005568086_3befc0562a952

Kobaisy, M., et al. "Antimycobacterial Polyynes of Devil's Club *Oplopanax horrdus),* a North American Native Medicinal Plant." *J. Nat. Prod.* 60. (1997): 1210–1213. Print.

Korkina, L. G. "Phenylpropanoids as naturally occurring antioxidants: From plant defense to human health." *Cell and Mol. Bio.* 53. (2007): 15–25. Print.

Minto, Robert E., & Blacklock, Brenda J. "Biosynthesis and Function or Polyacetylenes and Allied Natural Products." *Progress in Lipid Research*. 47. (2008): 233–306. Print.

"Mycobacterium." *Encyclopedia Britannica Online*. Web. 28 Feb 2012. http://www.britannica.com/EBchecked/topic/400274/Mycobacterium.

"Native American Ethnobotany". *University of Michigan-Dearborn*, 2003.Web. 22 Feb 2012. http://herb.umd.umich.edu/herb/search.pl.

Newman, David J, Gordon M. Cragg, & Kenneth M. Snader. "Natural Products as Sources of New Drugs Over the Period 1981–2002." *J. Nat. Prod.* 66. (2003): 1022–1037. Print.

"Plants Profile." *USDA Natural Resources Conservation Service*. Web. 22 Feb 2012. http://plants.usda.gov/java/profile?symbol=ANMA.

Ren, Yan-Zhao, Yu Zhang, & Yan-Ping Shi. "Simultaneous determination of nine flavonoids in *Anaphalis* margaritacea by capillary zone electrophoresis." *Talanta*. 78. (2009): 959–963. Print.

Samanta, A., Das, G., & Das, S. K. "Roles of Flavonoids in Plants." *Int. J. Pharm Sci Tech.* 6. (2011): 12–35. Print.

Shimelis Olga, Craig Aurand, & An Trinh. "Fraction of cis/trans FAMEs using Discovery Ag-Ion SPE." *Sigma-Aldrich*. Web. 28 Apr 2012. http://www.sigmaaldrich.com/technical-documents/articles/reporter-eu/fraction-of-cis-trans.html

"It's Not As Easy As It Looks":
College Students' Experiences With Romantic Relationship Development and Decline

Katelyn Palka

ABSTRACT

As the number of young adults enrolled in colleges and universities increases, so too does the need to study and understand the communicative processes through which young adults engage in romantic relationship development and decline. Grounded in the model of interaction stages in relationships (Knapp & Vangelisti, 2009), the researcher conducted eight interviews with college women to investigate how young adult women use interpersonal communication in order to develop, maintain, and ultimately terminate their romantic relationships with those of the opposite sex. The researcher describes the ten stages of coming together and coming apart using the language of the model. Implications of these findings are discussed to provide insight to researchers and individuals who are looking to understand how we can effectively learn more about the ways in which young adult women experience relational development and decline.

Keywords: interpersonal relationships, dating, coming together, coming apart

"It's Not As Easy As It Looks":

College Students' Experiences With Romantic Relationship Development and Decline

The study of interpersonal communication, in any context or focus, can serve as quite useful in explaining the ways in which humans develop relationships with one another, especially with those of the opposite sex. Since interpersonal communication is so unique and special, it is important that we understand how to use it, in order to have healthy interpersonal relationships. The dynamics of relationships can be rather complex and they continuously change as relationships develop. Parks (2012) found that the building and maintaining of these interpersonal relationships are crucial to the health of individuals as one's interpersonal relationships serve as grounds on which to increase or decrease one's social skills, overall health and well-being, and that poor interpersonal relationships can have a negative impact on future relationships. His findings suggest that inadequate interpersonal relationships can lead to health

and wellness complications such as mental illnesses, violence and suicide, stress-induced illness in the cardiovascular system, a poorer immune system, or risky health practices (Parks, 2012). In turn, it can be argued that understanding the dynamics of interpersonal communication can help one build and maintain stronger relationships and, in turn, impact their health physically, emotionally, and psychologically (Parks 2012).

For the purposes of this study, interpersonal communication will be defined as "the kind of communication that happens when the people involved talk and listen in ways that maximize the presence of the personal" (Stewart 33). In the present study, I will use Knapp and Vangelisti's (2009) stage model as a lens to understand the ways in which young adult women use interpersonal communication in order to develop, maintain, and potentially terminate their romantic relationships with those of the opposite sex.

A Model of Interaction Stages in Relationships

Knapp and Vangelisti (2009) described common stages experienced during the development and decline of romantic relationships. This model was designed to explain the stages and communication experienced by committed romantic partners. However, it should not be forgotten that romantic relationships are complex and take various forms. The model has two sections, five stages that describe the process of coming together and five stages that describe the process of coming apart. Each process will be discussed in turn.

Coming Together

The model opens with the stage labeled *initiating*, in which we "scan the other person" (Knapp &Vangelisti 37). The stage is characterized by communication that makes the person appear to be likable or pleasant. It is essentially a time to make yourself known the person of your desire and to express your interest in them (Knapp & Vangelisti 2009).

Following the *initiating* phase is that of *experimenting*. Moving past simply expressing interest in another person, in this stage two people work to find a common ground, an integrating topic that they both share. In this stage people engage in phatic communication, commonly referred to as small talk, in which they discover more about one another. Many relationships never move past this stage as commitment is limited, and if there is a lack of integrating information, two people may not find a use in furthering a relationship past this stage (Knapp &Vangelisti).

If two people have engaged in the *experimenting* stage and are able to achieve an increased level of closeness, Knapp and Vangelisti (2009) argue that these two people have left the *experimenting* stage and enter into the *intensifying* stage. It is in this stage that two people are infatuated with one another and begin expressing direct commitment to one another. The two people are maintaining independence while concurrently fusing their personalities. They feel more comfortable with one another and are able to increase personal disclosure, sharing previously withheld information that would make them vulnerable (Knapp & Vangelisti).

If the partners move to the next stage of closeness, then they proceed to the *integrating* stage, in which two people develop a greater sense of interdependence. In this stage, two people's social circles merge and they are treated as a package deal. They are publically committed to one another, have heightened synchrony, and use phrases such as "we" and "our" in place of "I" or "my." Couples also develop personal idioms such as inside jokes that are private between the two of them. As a result, it should be clarified that achievement of this stage does not mean each person has lost their individuality or that they engage in complete togetherness; it simply means that their interdependence is more heightened than in previous stages (Knapp & Vangelisti).

Institutionalization of the relationship happens in the stage following Integrating. In the *bonding* stage, couples perform a public ritual such as marriage, in which they use a formal agreement to announce their commitment to the world. This stage can act as a "powerful force in changing the nature of the relationship 'for better or for worse'" (Knapp & Vangelisti 42). This is because when a couple engages in Bonding, they create a more concrete relationship that is harder to break free from.

Coming Apart

Knapp and Vangelisti label the first stage of coming apart as *differentiation*. In this stage couples no longer use "we" or "our" and instead use "I" or "my" as they regain a greater sense of independence from the other person. Typically communication during *differentiating* is focused around what differences people have in regards to areas such as interests or friends. It is possible for couples to move in and out of this stage of their relationship, and after a while they begin to develop expectations for how to deal with the conflict experienced in this stage. In this stage, couples have a tendency to experience a great deal of conflict or fighting, though it is entirely possible to have peaceful differentiating (Knapp & Vangelisti).

If a relationship continues to come apart, couples then enter the *circumscribing* phase of the relationship, in which communication qualitatively and quantitatively decreases. Here, couples communicate less than they used to and when they communicate it tends to be very restricted, only covering superficial topics. Couples discover more and more topics that cause conflict, and find that they bicker about even the littlest of things. This then causes more topics to become restricted, further hindering the breadth of topics covered between the two (Knapp & Vangelisti).

Following *circumscribing*, couple then find themselves in the *stagnating* stage, in which couples barely communicate, because they know how interactions will go and know they won't be particularly pleasant. There is a loss of an emotional connection between the two people and even the subject of their relationship becomes taboo. When couples find themselves stuck in this stage, it is generally because they are avoiding the pain of a break-up or still hold hope to revive the relationship to its former, happier state (Knapp & Vangelisti).

While they are *stagnant*, couples may not talk as much or as effectively. Should their relationship continue to deteriorate, couples will then begin partaking in activities characteristic of the *avoiding* stage. In this stage, couples literally make a sense of physical distance; their communication is actually designed to avoid any face-to-face or voice-to-voice interactions. One person or both will feel a lack of desire to spend any time or energy with the other and in cases where only one feels this way they may drop subtle hints or be quite blunt in communicating that they do not wish to be around the other person. Three forms of distancing may occur in this stage: avoidance, disengagement, and cognitive dissociation (Knapp & Vangelisti). In avoidance, couples may try to prevent interaction or reduce interaction should they have to be around one another. In disengagement they may actually hide information about themselves or what they are doing and communicate less personally. Cognitive dissociation can be found when one person or both disregard the other's messages or show that they do not value the other as highly. Typically cognitive dissociation happens when physical separation is unattainable, and in that case a couple could be in the same location but act as though the other doesn't even exist (Knapp & Vangelisti).

In the final stage of the model, *terminating*, couples literally end their romantic relationship. This can happen for multiple reasons; however regardless of the reason, it is the communication to the other partner that they no longer wish to continue the relationship. This

stage generally has a summary statement in which one partner provides rationale for the other as to why they wish for the relationship to end, a "decreased-access message" clarifying that the relationship is terminating, and addressing future friendships or enemies to avoid awkward interactions later on (Knapp & Vangelisti).

Dynamic Nature of Romantic Relationships

It can be argued that communication and relationships between romantic partners is complex, taking various forms. Afifi and Faulkner (2000) studied individuals who engage in friends with benefits type relationships, or cross-sex relationships that involve sexual activity. They discovered that about one-third of their participants experienced a friends-with-benefits type relationship and "they have important implications for our understanding of the way these relationships are experienced (Afifi & Faulkner 218)." The Institute for American Values (2001) found that "college women say it is rare for college men to ask them on dates" (5) and "'hooking up,' a distinctive sex-without-commitment interaction between college men and women, is widespread on-campuses and profoundly influences campus culture" (4).

In addition to changes in relationship types, technology has played a role in changing the way that men and women engage in romantic relationships. Another study revealed, "texting and sexting are relatively common in young adult romantic relationships" (Drouin &Landgraff 1). Furthermore, texting is becoming the dominant medium for interpersonal communication amongst teenagers and young adults (Drouin & Landgraff). It is important to apply Knapp and Vangelisti's Model of Interaction Stages in Relationships to new research, as communication is dynamic and always changing.

As a result of this changing nature of relationships, researching romantic relationships among college students could prove to be both eye opening and informative, as college acts as a unique time in one's life. The U.S. Department of Education (2010) reported that undergraduate enrollment has increased by 34% from the year 2000 to 2009, and they project that it will continue to increase through the years. With such an increase in enrollment, it can be argued that college is and will continue to be a shared experience by many. Learning from a popular demographic could provide as an informative reference for others who are or plan on attending college and engaging in romantic relationships while there.

The following research questions will guide the present study:

RQ1: How do young adult women communicatively engage in the development of relationships with romantic partners?

RQ2: How do young adult women communicatively engage in the decline of those relationships?

Method

Data Collection

Recruitment of participants was done through purposive sampling in which participants were chosen because they had experiences relevant to the phenomenon of interest (Baxter & Babbie). In order to participate, individuals had to (a) be at least 18 years of age and (b) have had a significant adult (i.e., after turning 18), heterosexual romantic relationship that lasted at least six months and had since ended. Participants were asked to engage in face-to-face, semi-structured interviews, each lasted approximately one hour. Semi-structured interviews allowed for flexibility of follow up questions while still providing structure in that each interview was directed by the order of questions. Each of the eight interviews was transcribed, yielding 260 single-spaced pages of data.

Questions in the interview focused on the communicative processes characteristic of each stage of the model. The model provided framework for the structure of the questions as they were intentionally asked in an order that coincided with the order of the coming together stages followed by the coming apart stages. For example, one question asked during the Integrating stage of the interview was, "Did you and your romantic partner us any nicknames or have any inside jokes with one another? What are some examples of these?" This question was framed using the notion that during the integrating stage couples develop personal idioms with one another. The interview protocol that guided this research appears in the Appendix.

Participants

Participants' ages at the time of their romantic relationship ranged from 18 years to 20 years. Their romantic partner's age at the time of the romantic relationship ranged from 18 years to 20 years. At the time of the interview, participants' ages ranged from 20 years to 21 years. Of the 8 participants, 7 described themselves as currently single and 1 described themselves as being in an open relationship with an opposite sex partner for the past 18 months. Since turning 18, participants' had had between 1 and 3 romantic relationships. Their most significant romantic relationship, the relationship in question, lasted between 7 months and 26 months and had ended

between 2 weeks ago and 19 months ago in relation to the time of the interview. All participants were female and, of the 8 participants, 7 self-identified as Caucasian and 1 self-identified as Asian. In regards to personality, 2 of the 8 participants identified themselves as introverted and 6 identified themselves as extroverted. As far as sexual orientation is concerned, 7 participants self-identified as heterosexual and one self-identified as bisexual. All participants were college students and were either in their junior or senior year.

Data Analysis

After collecting the data, a qualitative thematic analysis was performed using Knapp and Vangelisti's model as a framework. Using Smith's (1995) guidelines for qualitative analysis, I first thoroughly reviewed a single transcript and inductively analyzed the transcript for evidence of all stages in the model, creating themes that emerged in the data regarding each step. Next, I analyzed the remaining seven transcripts. After completing this process, I organized the themes relevant to each stage. Finally, I implemented exemplars to support each of them. Exemplars, or ideal representations found in the data, were used to establish credibility as well as increase understanding of the themes discovered in each stage.

Findings

In the following text, the themes that emerged in each stage of model are presented. Themes are then supported and explained through exemplars from interviews with participants.

Coming Together

Initiating. In the initiating stage, participants engaged in flirtatious interactions in which they expressed initial interest in their romantic partner or vice versa. The flirting took place through two methods: physical engagement and non-physical engagement.

Regarding physical engagement, participants nonverbally expressed interest in the other person. One participant explained how she and her romantic partner met at a fraternity party. "We just kind of bumped into each other, like physically bumped into each other and started talking that way (1: 27–28, Note: parenthetical citations include the interview number followed by line numbers from the transcript). This couple's physical encounter is what initiated their communication with one another, eventually blossoming into a relationship.

On a deeper level of physical engagement, a different participant actually engaged in sexually physical activity as a way of initiating their relationship. When asked who initiated the

relationship and how they initiated the relationship, the participant responded with, "I don't remember because we drunk hooked up…but I guess afterwards would probably be him to like continue it" (5: 16–18). In the physical cases, participants used physical contact, be it sexual or non-sexual, in order to express their attraction to the other person.

Regarding non-physical initiation, some participants were often asked for their numbers by the other person or in other cases, their romantic partner disclosed to a third party that they liked the participant (and that third party informed the participant). In one scenario of non-physical initiation, the participant and her romantic partner were attending a hat making activity in their college union, she explained:

"We were making the hats and he was like, I was like 'oh, I don't know what to put on my hat,' and he was like, "you should make a hat that says I heart…' and then like his name and I was like 'aw, I don't know, maybe,' and then I ended up making it". (2: 42–49)

Her partner was indirectly expressing interest by suggesting she put "I heart" with his name on her hat. She then reciprocated that interest by writing it on her hat and wearing it.

Experimenting. After initiation, participants encountered the experimenting phase. Experimenting is characterized by phatic communication, or small talk among participants with the goal of establishing common ground. Participants voiced that they were "just trying to get to know one another just like basic information about like ourselves" (7:46–47). That phatic communication used in getting to know one another focused on three themes: personal relationship, activities, and interests. These themes were all ways participants got to know their romantic partners and assessed how much they had in common with them.

In terms of personal relationships, participants revealed that they talked about family and friends with their romantic partner as a way to get to know more about one another. When asked what she talked about on her first "date" with her romantic partner, one participant explained, "we just talked about…just like stupid stories about my friends and his friends" (6: 56–58). She disclosed funny stories about her friends, he did the same, and in doing so, they each learned a little more about the personal relationships that they had with other people.

In addition to the personal relationships, participants also explained that they discussed activities, or things they were doing or involved in with their romantic partner as a way to get to know one another and establish common ground. One participant found common ground with

her partner regarding school work and explained, "We were both in the same major…we could have conversations about that and about the classes we were in" (4: 76–78). Another participant, when asked what topics they talked about during this time in their relationship described that she and her partner discussed "just kind of what happened during the day like what class he had come from, plans for later that night" (8: 93–94). These were simple topics in order to understand what the other did during the day, "nothing really heavy" as one participant called it.

Regarding interests, participants and their romantic partner discussed things that they enjoyed doing or subjects that they found particularly appealing. One participant said she and her romantic partner would talk about "music, what we like to do…" (3: 47–48). Another participant found a lack of common ground in the area of interests, and actually questioned how much she wanted the relationship to progress because of it:

> "He said he did not like sports and that kind of bothered me a little bit just because I know I was a big sports person and I was kind of second guessing if some of our interests were too different" (1: 78–80).

This participant showed that during the experimenting phase, it is possible that couples may not continue to progress in their relationship for fear of too many differences regarding interests or other areas. However, she decided that their lack of common ground regarding sports was not enough to justify a decision to end the progression of the relationship.

Intensifying. Following experiment, participants experienced the intensifying stage, or in other words, the "lovey-dovey" time in a relationship. During the intensifying stage it was evident that participants and their romantic partners were infatuated with one another and began to increase spending their time together. A summary of what all participants experienced was perfectly expressed when one participant said, "Every spare moment I had I would spend with him" (8: 157). In this stage it was found that participants and their partners engaged in establishing relational identity and testing.

In order to establish relational identity, it was found that participants asked one another exactly "what they were doing" in the relationship and created rules about not seeing other people. One participant explained:

> "We had been seeing each other for a few weeks and when we were hanging out one day he just kind of asked 'like what exactly are we doing?'…we both kind of agreed it was more; like, he initiated the talk about becoming, like dating, but it was a mutual thing and

then we kind of discussed like what he meant and like not seeing other people and things
like that" (1: 16–18, 35–37).

At this point, couples began to define the relationship. For example, they agreed to commit to
one another and to not see other people.

While engaging in testing, participants purposefully did things to see if their partner felt
the same way by expressing their feelings or emotions to the other person or through disclosing
previously withheld information that would make them vulnerable. One participant decided to
disclose information regarding her relationship with her father and her experience with her ex-
boyfriend:

"After a couple of months I brought up like…I told him about my most recent ex-
boyfriend like before I met him was like, um, how it ended really badly and stuff and
like, and then I, like I don't have a good relationship with my father and I told him about
that and how I haven't like seen him or anything in like, I think it is been like 13 years or
so" (2: 96–97, 104–107).

Another participant displayed testing when she expressed her feelings for her romantic partner to
see how he would react:

"Just like saying like 'I love you' when I left the room and like stuff or like left and he
was with a group of his friends and like just seeing his reaction to that and seeing if he
responded and but like reacted the same way as he would if we were alone, which he did"
(5: 170–173).

Her expression of "I love you" could have gone unreciprocated, hence she was vulnerable, and in
her mind she was testing his true feelings for her and felt he might not respond as honestly or as
openly in front of his friends as he did when the two of them were alone.

Integrating. Following their "lovey-dovey" time, couple proceeded to the integrating
stage. In the integrating stage it was found that participants and their romantic partners displayed
commitment to one another both within their relationship and to people outside of their
relationship. They shared these displays privately between themselves but they also publicly
showed their commitment to one another as a couple.

The display of commitment between participants and their partners occurred internally
through the establishment of rituals, personal idioms, and the display of emotion. One participant
disclosed how she and her romantic partner would talk before bed on a nightly basis:

"We would say goodnight to each other in a special way, we would always say goodnight

to each other, and um, so that was always nice for both of us to just get to talk to the

person one more time before we went to sleep" (7: 252–254).

This display was private between the two of them and not shared with other people, making it internal. It was something that was special between the two of them and made them feel reassured and connected to one another.

External displays of commitment included showing affection in public, displaying commitment using technology, and partners giving jewelry to other participants. One participant discussed how she and her boyfriend used technology to outwardly display their commitment to one another. She explained, "We were in a like relationship on Facebook…we like really like Facebook I guess, um we like wrote things to each other like publically just like made references to like stupid things on our walls and stuff" (5: 313–316). Technology, specifically social media, was a major way in which participants and their romantic partners made their relationships public to others.

Bonding. The act of bonding, in which couples make their relationships more permanent through institutionalizing the relationship, was not seen in any of the participants' relationships. Since all participants were college students, they and their romantic partner may not have yet reached that level of seriousness in their relationship.

Thus far, findings have detailed the ways in which participants developed their romantic relationships. Next, I will discuss findings detailing how participants experienced the decline of their romantic relationships.

Coming Apart

Differentiating. In the differentiating stage, participants and their romantic partners engaged in establishing independence from one another. They did so through either positive individualism or negative individualism. As a result, participants perceived differentiating as either helpful or hurtful to their relationships.

Those participants that experienced positive individualism explained that this separate time away from their partner helped them to not feel smothered or suffocated by the relationship. It also helped them not get sick of one another. One participant explained that a way she would assert her independence from her partner was by spending time with her own friends. She described:

"Well we like you, spend our separate time from each other like personally with just my friends and just his friends. Having space from each other was good because I feel like if we're together like 24/7 I wouldn't have my independence and like my alone time that's like needed" (6: 359–362).

Through having her separate time, she was able to maintain a sense of independence, which was important to her. Another participant who also experienced positive individualism said that her independent time was good for her relationship because "being with someone all of the time is just a lot to handle" (3: 228–229).

Those participants who felt they experienced negative individualism expressed that doing so made them and their partners realize different viewpoints or interests they had regarding what was important to them. This created small conflict, very characteristic of the differentiating phase. In one interview, the participant described how her joining an activity that her partner was not a part of created issues:

"I decided to join another activity on campus knowing full well how he felt about it and that he did not approve…it was something that was important to me and it was something that I needed and wanted as part of my college career…me joining the organization became a point of contention between the two of us…he claimed that I wasn't as committed to him as I was to the activity" (4: 348–351, 359–361).

Circumscribing. If the relationship continued to head south, the couple then proceeded to the circumscribing stage. In the circumscribing stage, couples experienced a decrease in the quality and quantity of their communication. Findings suggest this occurred through restricted communication and an argumentative inclination.

Participants and their romantic partners engaged in restricted communication with one another by avoiding certain topics and simply talking less to one another. One participant explained the kinds of discussions, or lack thereof, that she and her romantic partner would have:

"Literally the communication would just be about like us arguing and the problems we had like between us it was never about like what we like anymore or what our days were going like and topic wise I don't think we talked about anything…we literally avoided talking about anything else besides the two of us" (1: 421–426).

She and her partner no longer covered all of the topics they used to and communication was limited to talking about the two of them rather than what they did that day or how their lives outside of the relationship were.

Participants also expressed experiencing an argumentative inclination, in which they and their partners fought or bickered over insignificant matters, decreasing the quality of their communication. For example, in one interview the participant explained that she would get mad at her partner for almost everything he did. "I would get mad if he was late to work and was just being lazy about it or something like that. And then, but actually in reality I was angry about something else" (7: 444–446). She fought him over everything, even if what he was doing, like not going to work, had nothing to do with the relationship. This argumentative inclination also dealt with the tone in which participants and their partners used when they spoke to one another. One participant expressed that her partner would send her "catty" texts such as "Can you actually let me in and pick up my phone calls?" (5: 589), whereas before in the relationship he would have never sent her a message of that nature.

Stagnating. Following the circumscribing stage is the stagnating stage, in which couples lose any kind of emotional investment. In the stagnating stage, data supports that participants were stuck in a relationship that wasn't getting any better. They were not emotionally involved, yet they were not quite ready to let go. Findings suggest participants engaged in relational detachment and relational preservation.

While engaging in relational detachment, participants were essentially going through the motions of the relationship, and one participant described her communication with her partner as "task based instead of like emotion based" (5: 717–718). Another participant explained that she and her partner weren't "in it" like they used to be. "I think unintentionally we avoided like regular conversations. I'm not sure why, it could just be because we just weren't in the relationship anymore…our heart wasn't in it anymore" (7: 516–518). The emotional detachment they were experiencing made it very difficult for her and her partner to mend the relationship as they were still helplessly trying to do.

Furthermore, participants also expressed that they and their partners engaged in relational preservation, in that they were holding onto the last bit of life that their relationship had, or were simply prolonging the inevitableness of a break-up by convincing themselves that it would get better. One hopeful participant explained, "I knew it was in a bad place, I kind of looked at it as

like I know it is in a bad place but we can fix it…I think I was in a lot of denial" (8: 627–629). She recognized that the relationship was going south quite steeply but still kept herself "convinced" that it would get better and she and her partner could fix their relational issues.

Avoiding. In each case, the stagnating stage was followed by the avoiding stage. In the avoiding stage, it was discovered that participants created a sense of physical distance between them and their partners through restricted togetherness and technology to make it so that they were away from their partner more than they were together.

While practicing restricted togetherness, participants made conscious efforts to limit the amount of time that they spent with their partner. They did this by only hanging out for a short period of time, making other plans, or lying about what they were doing. In one case, the participant admitted to making up lies about what they were doing so that they did not have to be around their romantic partner. "I would try and like say like I'm going to go take a shower, but really like I would just go to my room and just be alone and just be by myself," (2:612–614) she explained. She would purposefully tell her partner she was doing something for the sole intent of being away from him because she wanted to be. Another participant explained, "I would go over there and like tell myself to leave in an hour," (1:380–381) making sure she did not spend any more time than necessary with her partner.

In order to physically distance themselves from their partner using technology, participants admitted to either themselves or their partner ignoring the other's text messages or phone calls, or lying to the other and responding that they were too busy to talk at that moment. One participant described how her romantic partner would make excuses as to why he couldn't talk to her:

> "I just remember like getting a lot of messages like 'can't talk right now, in lab' or like 'can't talk now, I'll call you later'…I think he almost used it as an excuse sometimes…it was easier to just blame it on the lab than it was to just say that he was avoiding it" (4: 663–669).

Other participants recall purposefully ignoring messages received from their romantic partner and were able to physically distance themselves by not answering any technological form of communication.

Terminating. In the final stage of the relationship, terminating, participants and their romantic partners ended the relationship. The manner in which this communicative stage occurred was either destructive or constructive.

Those participants who terminated their relationships using destructive communication described a lot of yelling, screaming, and negative put downs towards the other person. One participant even disclosed physical abuse during termination. The following is a description of one participant's experience after she disclosed to her boyfriend that she had cheated on him. "He said he did not wanna be with me anymore but it was not like a calm manner it was yelling and screaming and 'I hate you, don't ever talk to me again'" (1: 531–532).

Other participants expressed that the termination of their relationship with their romantic partner was done through rational explanations and mutual understanding of the issues of the relationship. One participant described, "I just remember him saying that we needed to figure out who we were and um, he thought that we could only do that by breaking up…we just kind of logically talked about it" (8: 777–778, 802). Here, participants used more positive communication during termination, rather than simply tearing the other down by saying or doing something harmful.

Discussion

The present study sought to apply Knapp and Vangelisti's (2009) model of relationship development to current college-aged relational experiences. Knapp and Vangelisti (2009) argued "students of human communication find concepts and principles easier to learn when they can analyze and test them in the context of common experiences" (p. ix). This study helps research regarding how college students engage in relational development and decline. As a result, it provides a deeper understanding of how young adults communicatively navigate their way through the development of their romantic relationships. With the existing research regarding various forms of relationship (i.e., Bogle, 2008; Afifi & Faulkner, 2000), looking at the model now that we know such relationships exist can help those studying communication understand how the model functions with the emergence of changing technologies and relationship types as mentioned earlier.

This study offers important implications for the study of interpersonal relationships in that it brought up a point that not all couples engage in relational development as traditionally expected. One participant experienced a "hook up" during the initiating stage of her relationship,

and according to the model, physical expressions of interest is characteristic of the intensifying stage (Knapp & Vangelisti). This finding supports existing research regarding the reality and frequency of hooking up among college students. Bogle (2008) found that "the dominant culture/sexual script for most of the twentieth century (i.e., asking someone out for a date as the first stage toward finding an intimate partner) is no longer being used by most college students" (p. 44). It was also found that "hooking up is an outgrowth of how college students socialize today" (Bogle 29), meaning that more college students engage in hooking up as a way of having relations with those of the opposite sex. Such findings could support an expansion of the initiating stage to include these sorts of displays of attraction.

This study also showcases the complex nature of the coming apart stages of relationships (i.e., differentiating, circumscribing, stagnating, avoiding, and terminating). When one looks at the coming apart stages individually, they seem quite clear and unique; however, when applying those stages to data its distinguishing stages can be tricky. The lines between the coming apart stages are not always clear. Activity in each stage appeared to be very similar, in that you could find participants discussing how they or their romantic partner would choose to hang out with their own friends in both the differentiating and avoiding stage. The only distinction was that in differentiating, participants and their partners were hanging out with their own friends for positive reasons, such as needed individual time and in avoiding, participants and their partners were hanging out with their own friends for negative reasons, such as they were sick of being around the other person. It was enlightening to see this complexity of coming apart in that it is important to understand the reasoning behind why couples communicate certain things and not just what they are communicating.

Furthermore, post-relational dissolution appeared to play a significant role in experiences had by participants. There exists research regarding how couples continue to communicate post-termination (i.e., Koenig Kellas, Bean, Cunningham, & Cheng). Such findings suggest that Knapp and Vangelisti's Model and other models regarding relationship stages could improve by adding a stage to bring attention to communication that may occur post-termination. Such examples of this would include still being friends, having to share custody of children, or having to communicate at work (Koenig Kellas et al.). Many participants of this study explained that they experienced at least some kind of interaction with their romantic partner after breaking up, simply because they had class together or decided to still be friends. However, much of the

literature in existence deals with the ways that partners communicate post-divorce, failing to address that communication post-termination could arguably exist in relationships that never made it to marriage, such as college relationships. From some of the findings of this study, it could be interesting for future researchers to look at ways in which college students engage in post-relational dissolution.

While this study provides important new insight about interpersonal relationships, it was not without limitations. For example, only one member of the relationship was interviewed. As a result, only one perspective of the communication was taken into account. It could be helpful to engage in dyadic interviews, in which both the male and female of the same relationship share their perspective on the coming together and coming apart of the relationship through communicative processes.

Second, participants were all women. Because of this, only a female perspective was taken into account. In the future, it could be useful to engage in interviews with both men and women in order to compare and contrast how men and women communicatively engage in relationship development and decline. Having a male perspective could increase the breadth and depth of a future study, as it has the potential to add other viewpoints or understanding.

Lastly, the present study only interviewed students of one university. It could be beneficial in future research of this nature to look at students from various colleges or universities in various parts of the country as different universities may have different relational cultures. This could broaden the demographic spectrum of a study and potentially open eyes to any contrastable data from different universities or areas of the country.

In addition to other propositions of future research, it could be beneficial for researchers to look at the ways in which technology, specifically Facebook and other social media, affect how college students engage in relational development and decline. Many participants of this study voiced that they were "Facebook official," or publically in a relationship on Facebook. It could prove interesting to research how this specifically affects coming together and coming apart, especially since Facebook did not exist in 1996 when this model was first introduced.

Conclusion

The present study helps to illustrate the importance of studying the dynamics of interpersonal communication. The maintenance of romantic relationships is not simple, and can be important to the health of those partaking in the relationship to engage in healthy

communication patterns. As Parks (2012) found, maintaining healthy relationships may be associated with one's overall health and well-being. It can be strongly argued that it is still important for future researchers to continuing looking at the dynamics of relationship as they are ever changing. Such research can prove to benefit everyone, as interpersonal relationships are unique and special due to their complex nature.

Works Cited

Afifi, W. A., & Faulkner, S. L. (2000). On being "just friends": The frequency and impact of sexual activity in crosssex friendships. *Journal of Social and Personal Relationships, 17*(2), 205–222. doi: 10.1177/0265407500172003

Baxter, L. A., & Babbie, E. (2004). *The basics of communication research*. Belmont, CA: Wadsworth.

Bogle, K. A. (2008). *Hooking up: Sex, dating, and relationships on campuses*. New York, NY: New York University.

Drouin, M., & Landgraff, C. (2011). Texting, sexting, and attachment in college students' romantic relationships. *Computers in Human Behavior* (2011). doi: 10.1016/j.chb.2011

Institute for American Values. (2001). *Hooking up, hanging out, and hoping for Mr. Right: College women on dating and mating today*. Retrieved December 10, 2011, from http://www.americanvalues.org/html/r-hooking_up.html

Knapp, M. L., & Vangelisti, A. L. (2009). *Interpersonal communication and human relationships*. New York, NY: Pearson Education, Inc.

Koenig Kellas, J., Bean, D., Cunningham, C., & Cheng, K. Y. (2008). The ex-files: Trajectories, turning points, and adjustments in the development of post-dissolutional relationships. *Journal of Social and Personal Relationships, 25*(1), 23–50. doi: 10.1177/0265407507086804

Parks, M. (2012). Personal relationships and health. In J. Stewart (Ed.), *Personal relationships and personal networks* (11th ed., pp. 42–51). Mahwah, NJ: Lawrence Erlbaum Associates.

Smith, J. A. (1995). Semi-structured interviewing and qualitative analysis. In J. A. Smith, R. Harré, & L. V. Langenhove (Eds.), *Rethinking methods in psychology* (pp. 9–26). Thousand Oaks, CA: Sage.

Stewart, J. (2006). Communicating and interpersonal communicating. In J. Stewart (Ed.), *Bridges not walls* (9th ed., pp. 16–39). New York, NY: McGraw-Hill.

U.S. Department of Education, National Center for Education Statistics. (2011). *Digest of Education Statistics*, 2010 (NCES 2011–015), Chapter 3.

Appendix

Interview Protocol

Please make note that during this interview you will be asked questions based only on what you consider to be the most significant heterosexual, adult, romantic relationship that has lasted at least 6 months and has since ended. Please do not discuss any other romantic relationship besides this one during the interview.

INITATION

1. Who initiated the relationship? How did they/you initiate the relationship?
2. If your ex initiated the relationship, did you also express or interest in that person? How so? Or vice versa, if you initiated the relationship, did your ex express interest in you as well? How so?
3. In what ways did you communicate with this person as you or he/she was initiating the relationship?

EXPERIMENTING

1. Describe your first "date," or when you first started hanging out with this person on a romantic level.
2. What topics did you discuss?
   a. What topics did you avoid?
   b. Why did you avoid them?
3. Did your partner ever bring up anything at this beginning stage of your relationship that made you second-guess your feelings about them?
   a. If so, what was it that they brought up and why do you think it made you feel that way?
4. Were there things your partner brought up that made you decide to want the relationship to progress?
   a. If so, what were these topics?
5. Were there instances in this early stage of your relationship that you brought up something you wish you hadn't?
   a. If so, what was it that you brought up and why do you wish you hadn't?

INTENSIFYING

1. Now, think back to the time when you started to have "serious" feelings for this person – what we sometimes call the "lovey dovey" time in a relationship, what was your communication like?
   a. What new topics did you discuss?
   b. What new topics did you avoid?
2. During this time, did you do anything to see if that person felt the same way about you as you did for him or her?
   a. Describe what you did.

3. Do you remember your partner ever doing anything to try to determine if you felt the same way about him or her and as he or she did about you?
   a. Describe what they did.
   b. How did you respond?

INTEGRATING

1. Now, think back to the time when your relationship became evident to others. In other words, when people saw the two of you as a couple. Characterize your communication at this time.
   a. What new topics did you discuss?
   b. What other changes did you make to the way you communicated?
   c. Were there topics you avoided at this time? If so, what were they? Why did you avoid them?
   d. How did you communicate your commitment to one another?
2. What sort of things did you do to let other people know the two of you were a couple?
3. Did you and your romantic partner use any nicknames or have any inside jokes with one another? What are some examples of them?
   a. Were these nicknames or jokes private between the two of you, or did other people know about them as well?
      i. If there were cases of both, what kinds of jokes/nicknames were public and which were private?
4. Were there certain routines you and your partner practiced on a regular basis that were special between the two of you?
   a. What sort of things were they, and why were they special?
5. What made you feel closest to your romantic partner during this time?

BONDING

1. Have you or your partner made any sort of public display of commitment?
   a. If so, what did you do and why was this important to you?
   b. How did your relationship change after that time?
   c. What new challenges did your relationship face following that display of commitment?

DIFFERENTIATING
1. In what ways did you separate yourself or assert your independence from your romantic partner? Has this been helpful or hurtful to your relationship?
   a. In what ways did your partner separate himself or herself or assert their independence from you? Did you find this helpful or hurtful to your relationship?
   b. Did the separation ever result in conflict?
      i. If so, how did you address that conflict?
      ii. How did you partner address that conflict

CIRCUMSCRIBING

1. Think back to the time in your relationship when it first started to go "downhill" or fall apart. Did you experience a lot of conflict during this time?
   a. What were the conflicts about?
   b. What was your communication like during this time?
   c. Were there any topics you began to avoid that you did not use to? If so what were those topics and why did you begin to avoid them?
   d. What did you do to express your dissatisfaction or unhappiness to your partner?
   e. What did your partner do to express their dissatisfaction or unhappiness to you?
2. Did you take any steps to repair the relationship at this point?
   a. Did your partner take any steps to repair the relationship at this point?

STAGNATING

1. Assuming the last part of your relationship did not go very well and it continued to fall apart. How did you feel about your partner at this point? How did you feel about your relationship at this point?
   a. Think of the top three words you could use to describe the relationship at this point.
   b. Why did you list each term you listed?
2. What was your communication like at this point?
   a. What topics did you discuss in your relationship at this point?
      i. Were there any topics you began to avoid at this point? Why did you avoid these topics?
3. Did you take any steps to repair the relationship at this point?
   a. Did your partner take any steps to repair the relationship at this point?

AVOIDING

1. During this turbulent time in your relationship, did you ever try to create distance between the two of you?
   a. What did you do to create the sense of distance?
   b. How did your partner react to this distance you were creating?
   c. How did your partner create distance between the two of you?
      i. How did you react to your partner creating distance?
2. What made you feel farthest from your romantic partner?
   a. Did you take any steps to repair that distant feeling?

TERMINATING

1. Did you or your partner end the relationship?
2. How did you or your partner end the relationship?
   a. What did you/they say in ending the relationship?
   b. Did you/they end the relationship in person? If not, what channel did they use?
   c. How did the other communicatively respond when they ended the relationship?
3. If you could go back, would you make any changes to the way the relationship ended?
   a. Why, and what changes would you make?

POST-RELATIONAL DISSOLUTION

1. Since you and your ex broke up, have you talked to each other?
   a. If so, why?
   b. Who initiated the communication?
   c. How did you communicate (face-to-face, phone, etc.)?
   d. What were the topics you discussed?
   e. Do you still talk?
   f. What were your feelings on the continued communication after ending the relationship?
   g. How is your communication with that person now different from when you were together?
2. Did any communication after breaking up make you change your mind about the break up?

OTHER QUESTIONS

1. What would you do differently in your next romantic relationship based on the events of this one?
2. Was there anything in this romantic relationship that you regretted and wish you could take back or do over?
3. What changes would you make to the way you communicate in your romantic relationships, if any?
4. Relationships, especially in college, tend to be learning experiences, what would you say was the biggest thing you learned from this romantic relationship that you had?

Medical Liability Reform in the United States:
A Comprehensive Approach to Medical Error

James Gerard Muller

Medical liability, also referred to as "medical malpractice," is a form of professional negligence in which the treatment a health care provider administers falls below the accepted standard of practice in the medical community, thereby causing injury or death to a patient. In the United States, malpractice is resolved through the process of litigation. If a patient has reason to believe he or she is a victim of malpractice, the patient has the legal ability to file a lawsuit. The legal remedy patient-plaintiffs seek is monetary compensation, measured by "economic loss" or financial loss and damage suffered by a person (Kessler 93–110). For a physician-defendant to be convicted of malpractice, the plaintiff's counsel carries the burden of proving: (1) the relationship between the plaintiff (patient) and defendant (physician) gave rise to a duty, (2) the defendant was negligent—meaning his care fell below the standard expected, (3) the plaintiff suffered an injury, and (4) the injury was caused by the defendant's negligence (Mello, Studdert 11–12).

Medical liability insurance is a commodity health care providers purchase in the event that they need to compensate a patient. Medical liability crises are understood as the result of sharp rises in the cost of liability insurance premiums for health care providers. In recent years, roughly from the years 2002–2005, the United States has experienced a malpractice crisis (Sloan and Chepke). What exactly created the recent medical liability insurance crisis remains in dispute.

Medical malpractice and medical liability insurance have been the topic of American policy debate from time to time. The topic's prevalence in American politics and policy debate result from a unanimous agreement among policy makers, scholars, physicians, and the educated public, that the medical malpractice system as it stands, is operating inadequately. There has been a general agreement as to what the central flaws of the current medical liability system are: (1) heavy financial burdens associated with insurance and (2) a disregard for patient safety and

reducing medical error. Admittedly, these problems are prioritized differently—some view the financial complications associated with medical liability as the most important concern, whereas others view medical error and patient safety as what deserves the greatest attention. Before addressing these concerns with considerable detail, it is first imperative to gain insight on how the litigation process works.

Malpractice Claims and the Litigation Process

Only 1 in 8 people injured by medical negligence file a malpractice claim[a], and of those claims filed, only thirty percent reach court (Mello and Studdert 13). According to the landmark Harvard Medical Practice Study (1990), only 1 in 15 patients who are victims of malpractice receive compensation, and many cases awarded compensation have no evidence of negligence (Kessler 96). Understanding these facts, the problem is obvious: many victims of malpractice are uncompensated, and the money that is given out does not go to the correct beneficiaries. The purpose of medical malpractice litigation is to compensate victims, and to deter physicians from practicing medicine in a complacent manner (Kessler 94; Mello and Studdert 17). Litigation has failed its purpose of providing compensation to those whom deserve it. In addition, litigation does not have the deterrent effect many assume it does; as Mello and Studdert have noted, the deterrent effect of the tort system rests on too many assumptions (17). One of the largest assumptions, *litigation as a form of deterrence,* makes is that physicians are always rational actors, performing some sort of hedonic calculus (Mello and Studdert 18). The second assumption the deterrence claim makes is that physicians internalize the costs of their own negligence (18).

Realizing litigation fails in accomplishing its two core purposes, the question that needs to be asked is: Should the medical malpractice compensation system continue to rely on litigation? Medical malpractice, placed in the context of litigation, is an adversarial process. Is an adversarial process in everyone's best interest? These are important questions policy makers must ask themselves, and they are questions constituents ought to ask policy makers. As it stands, malpractice's confrontational nature is a concern that requires considerable attention, and will be addressed detail throughout this paper.

Common Misconceptions With Respect to Malpractice

---

[a] American Association for Justice, *Five Myths About Medical Negligence*

Before addressing the two concerns laid out thus far, it is essential to know what scholars widely regard as "misconceptions" with respect to medical malpractice. The first, and perhaps the principal misunderstanding, is that frivolous claims are abundant. "The epidemic is medical negligence, not lawsuits." [b] As already discussed, many victims of negligence do not sue. A study conducted by researchers at the Harvard School of Public Health found that of 1,400 closed medical negligence claims, 97 % were meritorious.[c] In many cases, patients are filing claims in hopes to discover what went wrong during surgery, because the information was not disclosed by the physician.

The presupposition has been made that malpractice claims drive up health care costs. The National Association of Insurance Commissioners (NAIC) finds such a claim baseless, with factual information leading to a contrary finding. "The total spent defending claims and compensating victims of medical negligence in 2007 was $7.1 billion—just 0.3% of health care costs," claims the NAIC.[d] A similar argument, that malpractice claims drive up doctors' premiums, has been widely criticized as well. Medical malpractice insurance is not strongly correlated with experience-rating (history of claims), and many academics and health policy economists see malpractice premiums influenced by insurance cycles more than anything (Mello and Studdert 13–14; Sloan and Chepke 27–31).[e]

Proponents of the Help Efficient, Accessible, Low-cost, Timely Healthcare (HEALTH) Act of 2012, also referred to as "H.R.5," support the Congressional Bill on the premise "tort reform will lower insurance rates, and improve patient access to health care services."[f] Ironically, the Bill's text seems to protect physicians' self-interest more than patients. Contrary to the notion of "improving the liability system for patients," expressed in the preface of the legislation's text, H.R.5 limits a patient's legal rights by putting damage caps on malpractice lawsuits. It is assumed a bill with such a high marginal cost, limiting patients' rights, would produce an even greater marginal benefit—tort reform lowering insurance rates. The causal relationship between tort reform and a lower insurance rate has been in serious question for many scholars in the health policy field. Obviously, if damage caps are in place, insurers will be paying out less money to victims of malpractice; however, the assumption that they will pass their savings along

[b] American Association for Justice, *Five Myths About Medical Negligence*
[c] Mello & Studdert, *Claims, Errors, and Compensation Payments in Medical Malpractice Litigation*
[d] American Association for Justice, *Five Myths About Medical Negligence*
[e] American Association for Justice, *Five Myths About Medical Negligence*
[f] (HEALTH) Act of 2012, Introduction

to doctors by lowering premiums is unsupported[g]. According to research compiled by the American Association for Justice, states that have cap damages and states that do not, have little variance in premiums. The statistics make it clear that the cap-on-damage approach has done nothing to reduce health care costs.

The final issue regarding misconceptions is the concept of "defensive medicine." Also referred to as assurance behavior and avoidance behavior, this term can be understood as the workups physicians do (and do not do), respectively, that are influenced by liability pressures. Some scholars in the field, those very adamant about patient safety and strongly in opposition of tort reform (in other words, the most liberal of health policy scholars), claim defensive medicine is the biggest myth and largest over-exaggeration in the malpractice policy debate. Another thought, one that some reports bring to our attention, speculate that what doctors claim to be *defensive medicine* may be a way physicians generate income.[h] Mello and Studdert, whom could still definitely be characterized as "opposed to tort reform," are a little more diplomatic in their approach to defensive medicine, calling it a "slippery concept" and claiming measuring it is "notoriously difficult" (Mello and Studdert 23). Mello and Studdert make a good point— defensive medicine is a very hard concept to operationalize; after all, when treating a patient, where does "high-quality care" end and "liability motivations" begin? Scholars like Daniel P. Kessler have contributed research on defensive medicine's prominence that has been largely disputed by a wide range of colleagues.[i] With such strong assertions on both ends of the political spectrum, and with no strong methodological process to measure assurance or avoidance behavior's prevalence in medicine, there is no way to truly understand the reality of the "defensive medicine" concept.

The Facts and Figures of Medical Liability

With no clear consensus on the legitimacy of defensive medicine, perhaps it makes more sense to draw our attention to the facts well understood about medical liability and defensive medicine. The fact is, assuming arguendo, that defensive medicine does exist, and it contributes to health expenditures by as much as $50 billion dollars (a very generous estimate), eliminating defensive medicine would be nearly impossible. How can policy makers seek eliminating something of such an abstract form? The supposition is that caps-on-damages will reduce health

---

[g] American Association for Justice, *Five Myths About Medical Negligence*
[h] Congressional Budget Office, 2004
[i] American Association for Justice, *The Truth About Defensive Medicine*

care providers' fear of medical malpractice claims. The flaw in this reasoning is that doctors fear more than merely economic sanctions; a caps-on-damages proposal does absolutely nothing for the noneconomic costs associated with a malpractice lawsuit (i.e. negative publicity associated with being involved in a malpractice lawsuit). Considering the fact that these risks cannot dissolve with a "caps-on-damages tort reform approach," the practice of defensive medicine will continue.

An undisputed fact of medical liability is that preventable medical errors contribute to billions of dollars in health care expenditures. In 2000, the Institute of Medicine disclosed research regarding the health care system in a report titled "To Err Is Human." The report claimed "at least 44,000 people, and perhaps as many as 98,000 people, die in hospitals each year as a result of medical errors that could have been prevented."[j] The Institute of Medicine reports that the costs associated with these errors are thought to be as high as $29 billion annually.[k] Because "preventable medical errors" is a more concrete, realistic concept than defensive medicine, policy makers ought to focus their attention more effectively on something known to exist, and on something that, if largely resolved, could have the same fiscal effect as "eliminating defensive medicine[l]."

In attempt to resolve any confusion up to this point, while at the same time not over-simplifying the situation and dichotomizing the diagnosis of the malpractice system, obvious points can be made. Those that see the system's flaws, including too many lawsuits and the inner-workings of defensive medicine as a result of "fear of liability," argue for a tort reform approach. Others, those viewing the malpractice system's flaws as disregarding patient safety and not working vigorously enough to limit medical error, advocate for more comprehensive reform approaches. More "comprehensive" reform approaches steer away from tort and litigation in general, viewing the adversarial process skeptically. They see the reform approaches they advocate for limiting liability costs in the long run; the argument goes: obviously, if you reduce the amount of medical error, the amount of liability cases will also subside. Before shifting focus to the two contending approaches of medical liability reform, and examining their differences,

---

[j]Kohn, L.T., Corrigan J.M., & Donaldson, *M.S.*, *To Err Is Human: Building a Safer Health Care System*
[k] American Association for Justice, *The Truth About Defensive Medicine*
[l] Reducing medical error could directly lower costs by about 30 billion, and indirectly have a large impact as well, way beyond the fiscal benefits of "eliminating" defensive medicine.

analyzing points of consensus among scholars, policy makers, and the educated public is essential to conceptually forming the policy debate.

The medical malpractice tort-run system, as it stands, is problematic—this fact has been established by anyone that has invested research in the current medical liability system. The system tries to meet its goals of compensation; deterrence; corrective justice; and efficiency, but fails miserably in most cases. In respect to corrective justice, advocates of both approaches note that medical liability fears inhibit quality of care improvements. Proponents of both reform approaches also agree that the recent medical liability insurance crisis is not a unique experience of the 21st century. Within the past four decades, experts assert that we have experienced three medical liability insurance crises: the first in the late 1960s, and early 1970s; the second during the mid-1980s; and the third crisis occurred approximately from 2002-2005 (Kersh 46–48).

First-Generation Reform Approaches to Medical Liability

Health policy expert Rogan Kersh discusses reform in terms of two categories: first-generation approaches and second-generation approaches. First-generation approaches include reform options such as caps-on-damage awards, whereas second-generation approaches include reform options like alternative dispute resolution (Kersh 43–44). As the title of the chapter implies, Kersh discusses the political dimensions of liability reform in his "Medical Malpractice and the New Politics of Healthcare." According to Kersh, first-generation approaches are supported predominantly by policy-makers, while the academic community generally advocate for second-generation approaches (Kersh 43–67).

The Medical Injury Compensation Reform Act (MICRA) of 1975 and the Help Efficient, Accessible, Low-cost, Timely Healthcare (HEALTH) Act of 2011[m] (hereinafter H.R.5) are two examples of "First-Generation Proposals." These policies seek to "reduce the excessive burden the liability system places on the health care delivery system."[n] H.R.5 draws on the cause-and-effect relationship medical liability has on the health care system. As the argument goes, medical liability holds physicians too accountable; as a result, physicians are forced to practice defensive medicine, and the quality of care improvement is inhibited.[o] H.R.5 plans to resolve the liability system's flaws through imposing limits on medical malpractice litigation through: capping

---

[m] Also referred to as the Help Efficient, Accessible, Low-cost, Timely Healthcare (HEALTH) Act of 2012, because the proposed bill is still under Congressional consideration and review
[n] (HEALTH) Act of 2012, Purpose and Summary
[o] (HEALTH) Act of 2012, Background and Need for Legislation

awards and attorney fees; modifying the statute of limitations and "collateral source" rule; and eliminating "joint and several liability." H.R.5 utilizes research conducted by the Pacific Institute, a free-market think tank, which estimates the cost of defensive medicine to be $191 billion dollars.[p] The validity of McQuillan and Abramyan's research and statistics on defensive medicine ought to be seriously questioned—not only is the estimate four times larger than some of the most generous estimates, but their research is based on earlier research that has been largely refuted and criticized by scholars in the field and the Congressional Budget Office (CBO). H.R.5 provides a weak refutation to the argument that defensive medicine may be practiced to increase income, saying that "if an internist orders a CAT scan, the radiologist gets paid, not the internist." This rebuttal does not assume income increases at a systemic level; in other words, hospitals, which are corporations (regardless of our lack of conceptualization as such)[q], may enforce policies that encourage increased income. H.R.5 then makes the assertion that medical liability fears inhibit quality of care improvements, going on to also say the liability system fails to deter medical errors or promote patient safety. These points are duly noted by advocates of second-generation reforms as well. The question is, how will measures taken by H.R.5, which allegedly lower malpractice premiums for physicians, simultaneously result in increased disclosure and less medical error? As discussed earlier, it is naïve to reason that simply lessening the economic consequences a physician will face when disclosing details to a patient will make the physician more prone to disclosure. After all, the non-economic damages a physician endures (i.e., negative publicity and image) are not disappearing once a cap-on-damage award is implemented.

Along with the problems already discussed, the congressional Committee on Energy and Commerce critique H.R.5 in terms of its federal nature. The committee recognizes that the current state-based systems for dealing with medical malpractice are far from perfect; however, the committee claims that modifications and improvements need to be made within the context of this general state-by-state framework. In other words, the national "one-size-fits-all" approach that H.R.5 endorses will not suffice. Within the dissenting views of a national approach, a valid point is made—such a new, "…untested legal structure…" has "…little regard for the potential

---

[p] McQuillan & Abramyan, *Jackpot Justice: The True Cost of America's Tort System*
[q] G. Annas, personal communication, March 9, 2012

consequences."[r] Furthermore, aside from its highly problematic implementation, attempting to resolve medical malpractice with national legislation is unprecedented—"the principles of medical malpractice liability and the procedures for the conduct of medical malpractice lawsuits have [always] been governed by state law."[s]

The Committee of Energy and Commerce also critique the causal nature of H.R.5's proposition. H.R.5's logical train-of-thought flows in the following way: implement damage caps. As a result of damage caps there will be a reduction in premiums, and as an indirect result, medical error will diminish and high quality care will flourish. Critics of H.R.5 point out the flaw in the H.R.5 proponents' reasoning by stating the following:

> Yet, data indicate[s] that today, the overall medical liability insurance market is not in crisis. They also show it is the direct regulation of insurance companies—and not a cap on non-economic damages (one of the core elements of H.R.5)—that is responsible for the reductions in insurance premiums that have been seen.[t]

In light of what an overwhelming amount of data indicates, in which the above cited data concurs, caps-on-damages resulting in premium reduction is not realistic, and caps having such an intended impact is *indeterminate* at best. Logically, it follows that the causal link between caps-on-damages and a reduction in medical error is equally weak, if not weaker. Taking this data into consideration, instead of relying on damage caps to indirectly solve the severe problems the United States faces with medical error, it makes more sense that reducing medical error should be legislation's first step, its direct intention, and reduction in premiums, if at all influenced by medical malpractice lawsuits, flow from a reduction in medical error.

Problems in terms of constitutionality and fairness arise with H.R.5's caps-on-damages approach also. Lawyers representing patient-plaintiffs have argued caps are not only unfair, they are also unconstitutional. The caps are claimed to be such because they "[discriminate] between patients with minor medical injuries and those with severe injuries." The reasoning here is simple: people with minor injuries have the potential to receive full compensation, whereas people with more serious injuries might get only a fraction of their owed compensation under capped damages.[u] Employing this argument, the caps-on-non-economic-damages have been

---

[r] (HEALTH) Act of 2012, Dissenting Views
[s] (HEALTH) Act of 2012, Dissenting Views
[t] (HEALTH) Act of 2012, Dissenting Views
[u] Gallegos, A., *Caps Under Fire: The Fight for Medical Liability Reform*

successfully declared unconstitutional by some states' Supreme Courts, due to the fact that the statutes are inherently discriminatory. Maxwell J. Mehlman, a prominent scholar in the health policy field, advances another argument discussing the statues' unfairness.

> … [It] seems unfair to limit the recoveries of one class of victims—those injured by malpractice—while permitting other tort claimants to recover more fully. Why should someone whose leg is mistakenly amputated, for example, get less than someone who loses a leg when they are run over by a negligent, insured motorist (153).

As Mehlman describes in a couple of sentences above, caps-on-damages in medical malpractice tort cases are unfair because they prevent a victim of injury from receiving the same compensation they might have otherwise gotten had the tort been of a different nature (e.g. as the example used above proposes, compensation resulting from the negligence of a motorist).

For the reasons listed and described above, first-generation reform approaches to medical malpractice are less promising than the latter, second-generation reform approaches. While the different first-generation approaches to reform are limited, most are simply different variations of damage cap proposals, second-generation reform approaches, which position medical error as their top priority, are large in number and diverse. Equally important to note is that second-generation approaches are not mutually exclusive; in fact, advocates of each approach list other reform policies that ought to be adopted in conjunction with what they are discussing. In the next section of this paper, contending second-generation reform options are described, and the section concludes with a proposal to policy makers of the best second-generation reform package.

Second-Generation Reform Proposals: An Abundance of Reform Proposals That Solve the Right Problems

Unlike the simple solution first-generation reform approach, damage caps, second-generation reforms vary a great deal specifically in terms of the way they aim to resolve their fundamental concern, which is medical error. The various proposals' degree of projected influence also varies significantly. This paper considers some of the most popular second-generation proposals; these include litigation reform; alternative-dispute resolution; and enterprise liability reform.

Litigation Reform

It makes sense to first discuss reform proposals that aim at altering the litigation system; before discussing more drastic measures, let us first discuss the least intrusive reform approach,

and the one most similar to the current system. Catherine T. Struve suggests the first measure that needs to be taken, one that at least seventeen states have already adopted, is implementing certificate-of-merit provisions (174). Such a provision would ensure that the case a plaintiff's lawyer files is potentially meritorious. The certificate to file would be provided once an expert approves of the case at hand.

Another litigation reform measure, also directed at experts, would change the nature of expert testimony. As Struve suggests, much expert testimony relies on medical experts describing what they would have done in a given situation, as opposed to what the general health care professional would have done; litigation reform needs to address this issue and provide modifications that would allow for a more robust systemic analysis.

Other litigation reform provisions take the importance of the judicial function into account. Emphasis on the judicial function is important, especially considering their important roles as gatekeepers throughout the course of a trial. If judges are to employ tests that assess technical or scientific expert testimony, the judge needs to understand the basic concepts of scientific evidence. Some people have advocated for specialized health courts that would only hear medical malpractice cases (Struve 176). However, the problem with this idea is that the selection and retention of judges may become too politicized, and interest groups' interests may coincide with judicial decision-making. Yes, the same argument could be made with regard to general courts, but by its nature a specialized court would be more prone to such influence. Instead, "specialized divisions within a state's trial court of general jurisdiction would be a better option," claims Struve (177). As judges rotate into this division of the trial court, they receive special training through a "continuing judicial education" program that states could implement.

Considering the above proposals, litigation reform, in general terms, attempts to address the claims process, the evidence process, and judicial decision-making, respectively. While reforms in these three areas are all necessary, these reforms alone are insufficient. The litigation reform measures laid out would aim to fulfill the reform goals of compensation and efficiency, but how would these policies significantly reduce medical error? Furthermore, this reform policy fails to consider the fact that a litigation process, in general, will discourage *real change* in medical liability from occurring; as discussed earlier, litigation, an adversarial process by nature, does not encourage disclosure—and disclosure is absolutely necessary if we are to see a significant reduction in medical error.

Struve's suggestions are thoughtful and well noted. Trial reform is attractive, because it is a "less invasive procedure," compared to other second-generation alternatives. Trial reform should be further researched and considered as a legitimate end to a better medical liability system. However, in light of the reservations already made about litigation, it is important not to presuppose that the court remains the predominant institution in instances of medical malpractice. In other words, litigation needs to be seen as a subsequent method of resolution, and not the primary approach to reform.

Alternative Dispute Resolution

The American Academy of Orthopedic Surgeons (AAOS) sees the medical liability system for what it is. The organization understands that a majority of the money medical professionals pay out in malpractice cases go to the costs of litigation. In light of their understanding, these physicians wish to see broad-based liability reforms; in their opinion, second-generation reforms have the ability to fairly compensate patients while reducing liability costs and increasing patient safety. The AAOS argues for alternative dispute resolution (ADR)— or "non-litigious means of resolving conflict."

In the article "The Limits of Apology Laws," published by the Hastings Center Report, Rebecca Dresser discusses how an initiative implemented to reduce medical error and apology laws have had a questionable impact. Medical error is reduced by coming to terms with, and discussing, mistakes (i.e. disclosure), and proposing ways to prevent those mistakes from occurring again in the future. Malpractice lawsuits deter physicians from pursuing disclosure. Theoretically, apology laws encourage physicians to disclose information to their patients, because the law prevents an apology or a sympathetic expression from being used as evidence of negligence in court. As Dresser points out, the majority of states that adopt apology laws only have provisions that "eliminate expressions of sympathy, not admissions of fault…" Dresser's point illustrates the limits of apology laws; her point also implies that disclosure and litigation are mutually exclusive concepts. If methods other than litigation were embraced, however, and apologies were encouraged, results might include increased disclosure and more compensated patients. One critique of these ADR methods is that widespread disclosure and compensation would result in too much money being paid out. This point is well noted; however, are we to assume disclosure will always result in large payouts?

Research conducted within the past decade has made the point clear that patients mostly sue for non-monetary reasons. In light of these findings, ADR has become more prominent and popular in recent years. Patients sue mostly for disclosure (they want to discover what exactly happened to them), desire for an apology, and prevention of future errors to other patients. Different forms of ADR make these desires of patients possible, while compensating patients for experienced injuries at a much lower rate than would be the case with litigation. With mediation, patients get the remedy sought—an explanation and apology—while physicians get the satisfaction of a clear conscience, and the learning experience. The nature of litigation's adversarial process is that there is a winner and a loser; someone at the end of the trial will be satisfied, while the other person is unsatisfied and will most likely become resentful of the legal system. ADR creates the possibility that both parties will walk away content with what was discussed, and what deal was obtained.

Enterprise Liability Reform

A well-known and frequently cited statistic released by the Institute of Medicine claims "as many as ninety-eight thousand people die in hospitals each year as a result of medical errors that could have been prevented." A less familiar fact: systematic malfunctions are responsible for ninety percent of that medical error.[v] The keyword "systematic" means that the medical errors are a result of the system, the institution, as opposed to the individual, the physician. On this basis alone, if we are to assume the Institute of Medicine's figure to be correct, the current medical liability system is wrong in targeting the physician.

Bovbjerg and Berenson claim "[the] most important legal rationale for imposing liability on physicians and other medical providers is that doing so deters substandard practice." In light of the fact that information already discussed refutes the deterrent effect of liability, imposing liability on physicians does not make much sense. It is particularly interesting that while a broad array of research mentions the predominance of systematic errors as opposed to individual errors in medical error, both first- and second- generation reform policies take the existence of the "physician"-defendant for granted— caps-on-damages, litigation reform, and alternative dispute resolution, are three examples of reforms that all fail to address the question of who the defendant is: an individual or an institution? A comprehensive approach to medical liability reform needs to, at the very least, reconsider what entity should be held accountable for medical

---

[v] Obama and Clinton, *Making Patient Safety the Centerpiece of Medical Liability Reform*

error. Enterprise liability, considering the high proportion of systematic error involved in medical error, argues for the liability system to make health care organizations (e.g. hospitals) the respondent.

According to Michelle Mello, professor of law and public health at Harvard University, "[p]hysician insurers consider it infeasible to individually experience rate their subscribers. Physicians are sued too infrequently, and their claims experience varies too much over each three- or five-year period, to make experience rating actuarially feasible" (14). Institutions that have adopted enterprise liability, however, have insurance premiums that are experience-rated, because at the institutional level, experience-rating is practical. Experience-rating is a very attractive feature of enterprise liability; a reduction in insurance premiums as a result of safe practice gives a health care institution just another reason to take patient safety very seriously— in this case, an economic incentive is present, enough to encourage any corporation (such as a hospital). Although the public does not like to consider them in such terms, hospitals are corporations, and their economic interests and perspectives regarding the administration of health care could facilitate great measures to improve patient safety. Hospitals with enterprise liability now employ workers to assess the degree of quality improvement occurring at the institution, and they implement policies that further encourage quality improvement.

Hospitals, the institutions themselves, need to lead campaigns for reducing medical error, because systems can do more than any individual can—"health care organizations are in a far better position than individual providers to see opportunities to improve patient safety and to act on those insights."[w] Disclosure needs to be enthusiastically encouraged if we are to see great reductions in medical error. With respect to enterprise liability, Philip G. Peters, Jr. says the following:

> [Enterprise liability] in malpractice law will help hospitals create 'blame free' cultures that encourage open discussion of errors. As long as physicians remain at risk of individual malpractice liability, they can legitimately scoff at the notion that disclosure will be 'blame free.'[x]

When mistakes are made, the institutions need to take the blame. The system consuming the responsibility for medical error prevents physicians and other health care professionals from having to claim responsibility themselves, making them reluctant to the discussion of errors.

---

[w] Peters, *Resuscitating Hospital Enterprise Liability*
[x] Peters, Resuscitating Hospital Enterprise Liability

A Comprehensive Reform Proposal

A policy proposal with a one-size-fits-all approach, assuming the liability concerns of all states could be treated with the same solution, would be just as flawed as first-generation reform's federal legislation H.R.5. With that being said, the policy recommendations provided here should be analyzed in a broad manner, understanding the intended goals each portion of the proposal is trying to uphold.

What conditions does the medical liability reform system need to fulfill in order to be considered "successful?" If the goals of reducing medical error, upholding a patient's right to safety and compensation, and resolving problems with liability insurance premiums are adequately met, the reform proposal is adequate—and may very well become a "success." Absolutely essential to reducing medical error is disclosure; therefore, unless reform measures encourage professionals to disclose error, reform will not be successful. Reform measures will be equally flawed if they attempt to resolve fiscal concerns and in the process, limit patients' legal rights (i.e., caps-on-damages). Adequate and efficient compensation is also important and needs to be included in proposals; compensation needs to be negotiated outside the court room if possible, because litigation costs increase the amount paid out significantly, which increases health expenditures and ultimately, insurance premiums.

In prescribing liability policy, it makes sense to begin with how health care professionals are being insured. Medical liability reform proposals that wish to assure success need to include an enterprise liability solution, with experience-rating whenever feasible. Adopting enterprise liability ensures physicians they will not have to deal with the heavy financial burdens of the liability process, and that their time will not be taken from them. With these two burdens taken into account and resolved, physicians will be less prone to keep from disclosing errors. Disclosing errors at a more frequent rate might not only reduce medical error, but also might decrease the number of malpractice suits being filed. After all, "malpractice suits often result when an unexpected adverse outcome is met with a lack of empathy from physicians and a withholding of essential information."[y] Adopting a liability system that could actuarially handle experience-rating, like enterprise liability, would also provide yet another reason for health care organizations to make patient safety its top priority.

---

[y] Obama & Clinton, *Making Patient Safety the Centerpiece of Medical Liability Reform*

For the reasons listed above, under the "Alternative Dispute Resolution" heading, it makes sense for litigation to be a secondary instrument in disputes, and for some form of ADR to be used first. A no-fault liability payout system may be a good approach to compensation under an alternative dispute resolution. The "no fault" approach to compensation utilizes an administrative agency or "health court"[z] to evaluate claims without reference to whether negligence occurred. Critics oppose the notion of no fault because they feel creating such low standards for compensation would not work—there is no way the number of payouts would be affordable. Advocates of the no-fault approach argue that even though it compensates a larger percentage of injured patients, it also generates lower overhead costs, and the money used to otherwise pay for litigation could go to compensating injury. Although a no-fault system is vastly different from the current liability compensation framework, the policy seems feasible and is worthy of consideration. All in all, no-fault liability (and systems similar to it), with quick and efficient payout systems, need to be endorsed by liability reform proposals. Admittedly, more experience with such administrative models is needed; more experimentation in general is necessary for no-fault liability and the alternative compensation approaches it competes with.

Predictable, reliable compensation for medical error depends on injury compensation tables and guidelines. Legislation enacted could make it that experts—including actuaries, economists, physicians, insurers, and judges—establish tables of injury cost data. Although every injury is different, general guidelines, regardless of whether they have a binding or non-binding effect, produce compensation that is more efficient and predictable, as opposed to arbitrary and slow-paced. Tables that illustrate the average an injured patient receives in relation to their injury, facilitates in the negotiation process and prevents patients from being under- or over-compensated. Considering no evidence suggests medical malpractice is malicious in nature, compensation tables would not consider "punitive damages" when appraising injuries. Compensation tables ensure, as George A. Huber puts it, that in most cases "settlement is only a matter of using predetermined information to establish the compensation necessary to cover damages resulting from injury" (Huber 41–43). Because the injury suffered is only one element in the compensation equation, and the other elements vary greatly from patient-to-patient, tables should be used as guidelines as opposed to being used as mandates.

---

[z] As opposed to a judicial court

The final policy recommendation this paper chooses to discuss is seldom talked over in medical liability reform proposals. Nonetheless, physician work hours, and the influence this has on avoidable errors committed, requires our attention. In Mark R. Mercurio's *A Day Too Long: Rethinking Physician Work Hours*, an article published by a bioethics institution, The Hastings Center in its *Hastings Report*, Mercurio considers the logic of long work hours for physicians. Mercurio claims there is ample evidence that long work hours can lead to increased chances of medical and surgical error.[aa] In light of this evidence, residents' and fellows' work hours have been significantly reduced. Mercurio explains that the current limits still seem inconsistent with "what is known about human physiology and performance," however, "[in] any case, the present limit is an improvement over the old system."[bb] Surprisingly to Mercurio, despite concerns and an eventual limit on residents' and fellows' hours in the hospital, there are no hour rules regarding attending physicians. Ironically, the workers that have no hour limitations are the oldest, and those with "final responsibility for patient management in the unit."[cc]

The rationale for no time limitation may be "'continuity of care'—trying to avoid passing a patient from one physician to another."[dd] Yet the same concerns with respect to residents' assignments to patients was trumped by the greater concern of declining performance and medical error; it seems difficult to understand why continuity of care would not be trumped by concerns of medical error for physician work hours as well. Mercurio makes his point clear when he concludes by saying, "if allowing a twenty-eight-year-old resident to work around the clock and beyond is unsafe for patients…then allowing a fifty- or sixty-year-old attending to do so makes little sense."[ee] All of the safe-practice standards and paper pushing in the world is not going to resolve medical error if physicians are incompetent due to exhaustion. Medical liability reform, in addition to resolving inefficiency at the systemic level, needs to seriously consider implementing legislation that limits physician work hours under certain circumstances.

The Hurdles of Implementation

Regardless of how great the intended prospects of a given liability reform proposal is, reforms obviously require implementation, a mechanism that puts the policy into effect. Some great reform proposals are disregarded and do not receive much appraise at all, because of their

---

[aa] Mercurio, *A Day too Long: Rethinking Physician Work Hours*
[bb] Mercurio, *A Day too Long: Rethinking Physician Work Hours*
[cc] Mercurio, *A Day too Long: Rethinking Physician Work Hours*
[dd] Mercurio, *A Day too Long: Rethinking Physician Work Hours*
[ee] Mercurio, *A Day Too Long: Rethinking Physician Work Hours*

inability to be implemented. Some reforms that have great prospects, and receive great reviews during experimentation, still have trouble with being implemented. The reasons that reform proposals have trouble being implemented are plentiful.

Second-generation reform proposals have difficulty receiving the attention they deserve because medical malpractice is chiefly understood as a "legal problem." Too often, medical malpractice is viewed primarily as a legal problem, and attempts to resolve the problems are aimed at methods of reducing litigation. For this reason, first-generation reform proposals, i.e. caps-on-damages, have been more successful than second-generation reform proposals. Reducing litigation as the intended goal, however, fails to recognize medical error in itself is intrinsically a problem, not just medical error's impact on the litigation system. Research that suggests up to 98,000 avoidable deaths occur is being ignored. Perhaps because they take the harsh realities of medical error for granted, policy-makers do not see reducing medical error as the top priority. Medical malpractice needs to be understood by policy makers as a "health policy problem" first-and-foremost. Focusing effort on the adversarial legal process will not help medical professionals learn from their mistakes; the only impact litigation will have on physicians is building more resentment towards the medical liability system.

To really understand the true difficulty of implementing malpractice reform proposals, the politics of malpractice need to be discussed. Rogan Kersh provides an interesting, insightful, and unconventional perspective of liability reform by discussing medical malpractice's political nature. Malpractice policies, like all policies, are influenced by interest groups. Some interest groups that benefit greatly from the current malpractice system are reluctant to endorse reform measures, and others go so far as using their fiscal capabilities to ensure reform measures will not be taken. President Bush made this obvious point during a 2004 speech, saying "trial lawyers pursuing their own agenda have continued to block this much-needed reform" (Kersh 60). The effects of malpractice reform on their profession are obviously a concern of considerable weight for plaintiff lawyers. First-generation reforms limit the amount some plaintiff-victims will be compensated, which therefore decreases the amount the plaintiff's lawyer can make off of the case. Second-generation reforms go as far as considering alternatives to litigation, which obviously could greatly impact medical malpractice injury lawyers across the country.

American policy-making, in general, also has something to do with the limits of implementation. "Incrementalism," a term adopted by Rogan Kersh, "explains most American

policy making—including health policy. Ambitious malpractice reforms, in this perspective, will almost inevitably be whittled down to minor changes, as has happened in most health care debates over the past thirty years" (65). Albeit very skeptical, this view of implementation makes a valid point: until the vitality of broad reform measures is understood, only minor changes to the current system will be implemented. First-generation reforms are very mild with respect to the changes they prescribe, and because of this, they are more likely than second-generation reforms to be implemented.

Because all states have different experiences with medical malpractice, it is important that malpractice reform be implemented at the state level. A one-size-fits-all approach to reform, whether it is first- or second-generation reform, is not the correct way to proceed. The Affordable Care Act, legislation the Obama Administration has been advocating for, will in effect nationalize health care in the United States. If the bill, which is under Supreme Court consideration now for allegedly being unconstitutional by nature, is upheld, nationalized medical liability reform might be worthy of reconsideration. But as of now, if the federal government were to create legislation that reflected some sort of second-generation reform proposal, judicial action would be likely to follow (Kersh 66). H.R.5, the HEALH Act of 2012, is one example of national legislation that, if it became law, may be struck down for going beyond the reach of the Commerce Clause (the Commerce Clause is how national malpractice reform would be justified) (66).

The final issue this paper addresses can be understood by asking the question: can we mandate compassion? Ron Paterson asks this exact question in *Can We Mandate Compassion*, an essay he wrote published in *The Hastings Report*. Paterson, a former Health and Disability Commissioner of New Zealand, wrestled with this question as a result of a New Zealand statute that demands compassion. Although his article is not directly related to the issue of medical malpractice, his inquiry forces us to reflect on, and perhaps come to terms with, the limits of the law. Yes, there is a chance, as small as it may be, the legislature could enact legislation that embraces all of the concerns and viewpoints of second-generation reform. Even if the best attempts to reduce medical error were implemented, does that ensure change will, in fact, occur?

Reform proposals that encourage disclosure are valuable. The thought is that an increased level of disclosure will result in less future error. In most cases this antecedent-consequent relationship holds true. Disclosure, however, does not simultaneously lead to less error;

disclosure enables physicians to learn, and from the experience, errors of that kind do not reoccur. Turning medical error into a learning experience for physicians is important. This learning experience, however, is capable of being hindered if physicians are not receptive to the mistakes being made. The "culture of medicine," as it stands, encourages physicians to rationalize and excuse their mistakes, not to embrace them and encourage improvement. If we are to see a serious reduction in medical error, mistakes need to be confronted and discussed in a passionate manner no legislation can ensure. Certainly, for medical liability's sake, the "culture of medicine" needs to change—medical schools, teaching hospitals, along with other educational institutions, need to emphasize the process of atonement through self-improvement.

Conclusion

The aim of this paper was to discuss the highly problematic nature of the current medical liability system, resolve some of the most common misconceptions medical malpractice is prone to, and suggest reform measures that address the key concerns of the current medical liability system. More specifically, this paper attempted to focus on what issues are important, and the mechanisms by which these concerns could be addressed, and perhaps even resolved. It cannot be stressed enough that reform proposals, regardless of how great their resolutions are, all face the reality of implementation.

Finally, the need for more experimentation has been discussed rather briefly. It should be made clear, that conducting more experiments on the various reform proposals is absolutely essential to correctly resolving the medical malpractice system. If proposals with little influence are implemented, it will be a waste of both money and time, and the flawed liability system will continue to operate in a dysfunctional manner. The empirical research that experimentation provides will contribute more insight and guidance as to what contending reform options are most suitable for the medical malpractice system.

Works Cited

American Association for Justice. *Five Myths About Medical Negligence.* Washington, D.C., 2009. Print.

American Association for Justice. *The Truth About Defensive Medicine*. Washington, D.C., 2009. Print.

Gallegos, Alicia. "Caps Under Fire: The Fight for Medical Liability Reform." *American Medical News*. 11 April 2012. Print.

Huber, G. A. "Creating a Safe Environment." *Health Affairs* 31.4 (2012): 41–43. Web. 27 March 2012

Kersh, R. "Medical Malpractice and the New Politics of Health Care." *Medical Malpractice and the U.S. Health Care System*. Ed. W. M. Sage & R. Kersh. New York, NY: Cambridge University Press, 2006. 43–67. Print.

Kessler, D. P. "Evaluating the Medical Malpractice System and Options for Reform." *Journal of Economic Perspectives*, 25 (2011): 93–110. Print.

Kohn, L. T., J. M. Corrigan, & M. S. Donaldson. *To Err Is Human: Building a Safer Health System*. Washington, D.C.: National Academy Press, 2000. Print.

McQuillan & Abramyan. *Jackpot Justice: The True Cost of America's Tort System*. Pacific Research Institute. 2007. Print.

Mehlman, M. J. "Promoting Fairness in the Medical Malpractice System." *Medical Malpractice and the U.S. Health Care System*. Ed. W. M. Sage & R. Kersh. New York, NY: Cambridge University Press, 2006. 137–153. Print.

Mello, M. M., & D. M. Studdert. "The Medical Malpractice System: Structure and Performance." *Medical Malpractice and the U.S. Health Care System.* Ed. W. M. Sage & R. Kersh. New York, NY: Cambridge University Press, 2006. 11–29. Print.

Mello, M. M., & D. M. Studdert. (2007). "Claims, Errors, and Compensation Payments in Medical Malpractice Litigation." *New England Journal of Medicine* 354.19 (2006): 2024–2033. Print.

Mercurio, M. "A Day Too Long: Rethinking Physician Work Hours." *Hastings Center Report* 38.4 (2008): 26–27. Print.

Obama, B., & H. R. Clinton. "Making Patient Safety the Centerpiece of Medical Liability Reform." *New England Journal of Medicine* 354.21 (2006): 2205–2208. Print.

Peters, P. "Resuscitating Hospital Enterprise Liability." *Missouri Law Review* 73 (2008): 48–62. Print.

Predictors of Discordant Outcomes in Medical Malpractice Litigation. *Journal of Legal Studies,* 36(S2). Print.

Sloan, F. A., & L. M. Chepke. *Medical Malpractice*. Cambridge, MA: Massachusetts Institute of Technology Press, 2008. Print.

Struve, C. T. "Expertise and the Legal Process." *Medical Malpractice and the U.S. Health Care System.* Ed. W. M. Sage & R. Kersh. New York, NY: Cambridge University Press, 2006. 173–190. Print.

United States Cong. House. *Help Efficient, Accessible, Low Cost, Timely Healthcare (HEALTH) Act of 2012*. 112th Cong., H.R.5. Washington: GPO 2012. Print.

Noël Coward:
A Celebration of A Style All His Own

Emily McDevitt

There are two spheres of Noël Coward critics: one that believes Coward's success relies purely on his fine-tuned personality and dandyish eccentricities to create characters that will entertain the audience, and a second sphere that argues it is this frivolity and playful style that manages to reach above and beyond itself, making many of Coward's plays notable for their content and their comment on how people truly are behind closed doors. While we might argue that he comments on the social and societal conventions of the then present day, such conclusions are drawn at a stretch. Coward himself admits to very little preparation and pre-meditation in the construction of his plays. In an interview with Peter Garland he confesses, "The motives [are] always far less noble… To give myself a whacking good part as a rule, or the main motive was to entertain the audience." It is valuable to note that Coward did a remarkable job of concealing his rumored homosexuality, and that in the same manner he might have concealed deeper societal commentary in his works, but to excavate such criticism would mean ignoring the daring frivolity and personality of his masterfully woven plays. Coward himself says he wrote his plays for the amusing celebration of people being themselves, and both spheres of criticism would agree that Coward expressed his own personality as vibrantly as his characters do. He did acknowledge in his work that the strict codes and moral righteousness once dominant in the drawing-room dramas had given way to "a new pace and a new people" as "talentocracy mixes with the aristocracy" (Lahr 42). His comedies begin by establishing the expected picture of life and just as quickly subside to reveal characters reveling in their own plot-less realities, rejoicing in cigarettes, parlor games, and self-expression. At the center of *Private Lives*, *Hay Fever*, and *Fumed Oak* is the struggle for individuals to be themselves despite the mundane and restrictive circumstances that fight to silence them. Noël Coward's own flippant and playful self-expression blossoms in his characters in a more naturalistic comedy of manners style: one in which the dandy gets the final witticism while dancing about and breaking the furniture.

Noël Coward's legendary personality has preceded him throughout his career, painting a picture of a self-inspired dandy swathed in luxury, the very definition of elite. John Osborne commented that, "Mr. Coward…is his own invention and contribution to this century…" and John Lahr goes on to describe him as, "England's solid-gold jazz-baby who later turned into an international glamour-puss" (2). Friends and critics alike have recognized the infectious flippancy of Coward's personality and its translation onto the page in his most acclaimed pieces. In short, "Coward was his own hero; and the parts he created for himself were, in general, slices of his legendary life" (Lahr 2). This is not to say that all of the parts he wrote were direct reflections of Coward, but rather recreations of the stage eccentric that he wanted to be remembered as. The playful qualities recognizable in Coward's works should not be justification for writing them off as mere trivialities, and playwrights who deal more obviously with social commentary, such as Wilde or Shaw, should not overshadow the successes of Coward's repertoire. Such a comparison is impossible because Coward's talent lies in the creation of a genre all his own, pulling from a tier in society that marveled at him as much as he marveled at himself. The beauty of Coward's work can be realized by examining the characters and their expressive lives; the ones who are disappointed by his plays are those who seek spectacle and intricate plot thrill. Garland perfectly describes Coward as, "A master who knows that what lasts in drama isn't high seriousness or great ideas but the energy that is released when you turn loose two actors to battle with wit and passion in that cage behind the footlights." It is this fierce and passionate display of self that is at the center of Coward's writing, devoting its time to entertainment and frivolity in a way that some find difficult to appreciate, and others find intoxicating.

Nowhere is this fierce and tantalizing fervor more extravagantly displayed than in *Private Lives*, a comedy about an ex-husband and ex-wife, Elyot and Amanda, who meet again several years later on honeymoons with their new and comparatively plain spouses, Sybil and Victor. Coward masterfully sets the vivacious Elyot and Amanda against the dull Sybil and Victor, constructing a sense of the moral and mundane for the characters to rebel against as the play progresses. "What [Elyot and Amanda] both share, even before they encounter each other, is for the Self to be extraordinary and omnipotent" (Lahr 62). As both Elyot and Amanda reflect on their fast-paced marriage, it is not hard to see that both are feigning relief at having retired to a calm and quiet new spouse. Elyot remarks, "Love is no use unless it's wise, and kind, and

undramatic. Something steady and sweet… Something tremendously cosy; and unflurried by scenes and jealousies. That's what I want, what I've always wanted really," but there is no confidence to his words, and they transparently reveal a longing for the colorful flare he shared with Amanda (188). It is not at all ironic that these 'dramatic' scenes and 'flurried jealousies' will rear their ugly heads later in the play.

On the other side of the balcony, Victor pressures Amanda to smudge Elyot's image with thoughts of how they argued both physically and violently. Instead Amanda remembers with a smirk of pleasure, "I struck him too. Once I broke four gramophone records over his head. It was very satisfying" and admits to loving Victor "much more calmly" than Elyot (191 and 193). There is an acknowledgment of the restrictions of the calm life she is about to subdue to in her line, "I don't believe I'm a bit like what you think I am" (192). It is a cry from behind the veil of quiet composure Amanda has adopted for Victor, a sign that the wild and individual side of her is clawing its way to the surface. In a classic comedy-of-manners style, their respective others make a conveniently timed exit, allowing Amanda and Elyot to reunite once more. In a fit of selfish passion and spontaneity the couple charges off to Paris, leaving Victor and Sybil to sip cocktails and stare off into the distance.

Coward's plays are often criticized for their lack of plot, but Coward himself argues that "generally the characters make the situations" and that is where the audience should focus their attention. Act Two of *Private Lives* takes place in Amanda's conveniently vacant apartment in Paris, and it is the affectionate banter, which mounts into an aggressive fanfare that forms the heart of the story. Coward establishes the private aspect of Amanda and Elyot's 'private lives' in that their discussion always revolves around themselves. Looking back on the desertion of her new husband on their honeymoon, Amanda only has selfish comments to make: "Personally I feel grateful for a miraculous escape. I know now that I should never have been happy with Victor. I was a fool to ever consider it" (216). In a sigh of fresh relief, Amanda has once again established herself next to the sidekick who will celebrate himself as selfishly and playfully as she does.

Lahr argues that it is in the second act of this play that Coward picks up his pace, playing on the subtext of the language that reveals a biting and glamorized self-image, one that both Amanda and Elyot will forever admire in each other despite the wears and tears they subject themselves to. Not only do they both celebrate themselves as individuals, they revel in their

ability, as Elyot puts it, "to manage each other" (216). It is the raw edge in their banter, the beautiful way that each competes with and challenges the ideas of the other, that allows for the fluid navigation of conversations in which imperfect views of the world and flippant dismissals are key.

But Milton Levin, a defender of Coward's careful construction and organization of content, argues that the increasing simmer of tension in this scene is a result of Coward purposefully turning up the heat. "A pattern is established: badinage and love-talk alternate with increasingly dangerous reminiscence and insult; and, each time a return is made to small talk, a larger residue of temper is left to smolder" (Levin 101). Their dialogue is "so oblique, so teasing, so bold in foreshadowing the new naturalism which the next generation of English spoken playwrights would call 'the drama of the unspoken'," that Elyot and Amanda must invent a word to censor themselves from attacking each other too harshly (Lahr 63–64). They decide on the ever-humorous phrase 'Solomon Issacs' eventually shortened to 'Sollocks' as an apt means to prevent resorting back to their violent insolence but, just as it was foreshadowed in Act One, such viciousness was imminent.

The new naturalism and re-formatting of drawing room drama ideals that Coward is classically known for strips the upright morality and sense of surrounding that was previously upheld. In its place, characters sit on the floor, hit each other, break records over one another's heads, and call each other names. Where Oscar Wilde's epigrams once reigned supreme, Coward decided that his dialogue took a more naturalistic approach: "Wilde had a special line in epigrams. I hardly ever make epigrams. It is a form of play dialogue that's rather gone out of fashion." The last traces of societal standards finally disintegrate at the end of Act Two, when Amanda calls Elyot a "disagreeable pig" and he fires back that she is, "a mean, evil-minded, little vampire." When words are no longer fit for expression, the couple's aggression can be found in the stage directions at the end of Act Two. "*She rushes at him, he grabs her hands and they sway about the room…they trip over a piece of carpet, and fall on the floor, rolling over and over in paroxysms of rage.*" Such brutality might have shocked certain theatergoers who expected a return to the catchphrase witticisms of Wilde's day, but Coward understood this shift away from drawing-room conventions and brought a new audience to the theatre to welcome it along with him. Since Coward prefaces this conflict with the exposition in Act One and at the beginning of Act Two, this convention of quarreling is not a striking reality but an accepted one.

No matter what passes between Elyot and Amanda, the recognition of themselves in the other is what keeps them together. It is hardly surprising that at the end of Act Three, after Sybil and Victor have made their timely appearance and taken up defending their respective others, that Elyot and Amanda make an unnoticed escape once again and "*go smilingly out of the door, with their suitcases, and—CURTAIN.*" It is obvious there is no room in the self-glorifying world they share for anyone else, nor could anyone but Amanda and Elyot survive in it. "Coward's great theme might be marriage subverted, normality endangered; but his charm is that he is never so obvious. What seems most important in life is to live it wittily" (Hoare 5). While we might draw many conclusions from the twisted nature of Amanda and Elyot's coupling, Coward urges us to sit back and enjoy the show, safely in the dark looking in on the passionate wrestling behind the proscenium arch.

Coward uses this same abandonment of conventionality in his play about parlor games and bad manners, *Hay Fever*. Before the outsiders arrive, Coward begins by establishing the reality of disconnection and willing delusion the Blisses keep themselves in. The family operates on a level all its own. "To be interesting, to abhor dullness, to disdain normality, to indulge in the self and its self-expression, to worship accomplishment are the rules which govern the Blisses' manners" (Lahr 45). The only character that attempts to revolt against this and welcome normality is the daughter, Sorel. Her struggle throughout the course of the play to put on airs of convention makes her the only possible link between the world of the family and the mundane reality of the guests the family invites for the weekend. She finally relinquishes this battle at the end of the play and withdraws back into her niche as a Bliss, making the final decision that to be flippant, eccentric, and celebrated is better than fitting in with the boring outside world. Simon is the first to notice his sister's changing mood: "Sorel's beginning to be ashamed of us all…—I don't altogether blame her; we are very slapdash" (118). In a sort of meta-theatrical way, Simon recognizes the family's eccentricities, but in contrast to his sister, he doesn't feel that the way the family interacts with outsiders is all that unreasonable. His means of upholding this proud celebration of self is a signifier of just how much their parents, Judith and David, have made their mark.

SOREL: We're so awfully bad mannered.

SIMON: Not to people we like.

SOREL: The people we like put up with it because they like us.

SIMON: What do you mean, exactly, by bad manners? Lack of social tricks and small talk?

SOREL: We never attempt to look after people when they come here…. We've never once asked anyone if they've slept well.

SIMON: I consider that an impertinence, anyhow.

SOREL: I'm going to try to improve. (118)

It is clear that Simon believes an inability to entertain one's guests is bad manners, while Sorel defends that it is *the way* the Bliss family entertains that is in itself appalling. The audience is queued in to watching for Sorel's decision to be moral and polite but, much like in *Private Lives* where Amanda and Elyot decide to be civilized to one another, we acknowledge that her attempt is doomed from the beginning.

When the mother finally does make an appearance it is not at all surprising that she is even more childish than her children. A self-proclaimed "elder sister," Judith refers to her children and they to her as 'darling,' a means of defiantly collapsing what was left of familial hierarchy within the first few moments of Act One. Her whole life is treated as an anticipation of opening night, and her air of constant performance shows exactly where Simon got his display of self. Although it is clear that Judith is aware of her outlandish tendencies, she goes to great lengths to justify them: "I've been morally an extremely nice woman all my life—more of less— and if dabbling gives me pleasure, I don't see why I shouldn't dabble" (121). When her daughter intervenes, warning that "flaunting about with young men" shouldn't be something a married woman does, Judith is quick to shut her down. But, in a great display of flippancy and with a dismissal of 'darling,' she wins her affections back instantly.

JUDITH: You know, Sorel, you grow more damnably feminine every day.

I wish I'd brought you up differently.

SOREL: I'm proud of being feminine.

JUDITH: [*kissing her*] You're a darling, and I adore you; and you're very pretty and I'm madly jealous of you.

SOREL: [*with her arms around her*] Are you really? How lovely. (121–122)

Judith beautifully exemplifies the Coward-esque ability of making her audience see what she wants them to see. The ease with which Judith manipulates and reconstructs coins her as a creation of Coward himself, and the translation of this theatricality and frivolity from writer to

character is not difficult to detect. In his play, and especially in the character of Judith, Coward captures the vibrancy of the times. "*Hay Fever* is about style: the affectation and irony of an age addicted to seeing itself perform" (Hoare 4). Once again the focus is not on worldly witticisms or complicated plot twists, but on the celebration of self.

The mundane representation of normality that enters into the whirlwind of the Bliss household comes in the form of the four guests each family member invites for the weekend. Not only do they make the family's disconnect from the real world clear, they also highlight the ignorance and lack of communication between the family members themselves. As each member argues about which guest will stay in the family's only guest room, the conflict is dismissed to working itself out. The lack of honesty in the family about their obvious dysfunction is what is most noticeable to outsiders. As aforementioned, Sorel is struck with the wisdom that something needs to change, and she tries to explain the absurdities of her family to Richard:

It's Father's and Mother's fault, really; you see, they're so vague—they've
spent their lives cultivating their arts and not devoting any time to ordinary
conventions and manners and things. I'm the only one who sees that, so
I'm trying to be better. I'd love to be beautifully poised and carry off
difficult situations with a lift of the eyebrows— (140).

But the ease with which Sorel wishes to manipulate the world around her is a skill that she might learn directly from her narcissistic and extravagant mother. Although Judith is not always successful at relieving tensions for her guests, she puts on a show that satisfies herself, and in the end, that's all that really matters. What's worse is that the other members of the family instigate her theatricality. "One always plays up to mother in this house…we none of us ever mean anything" (155). Throughout the weekend the Bliss family forces their guests into a game of word guessing which eventually results in them pairing off with the people they didn't invite. David and Judith end up kissing other people, Sorel hears confessions of love, and Simon announces his presumptuous engagement to an unwilling guest. Myra finally bursts with frustration, as any sensible onlooker might after the displays of the evening, "This house is a complete featherbed of false emotions—you're posing, self-centered egotists… Every time I opened my mouth I've been mowed down by theatrical effects" (164). But this cry of reality is quickly absorbed into the atmosphere the Blisses have constructed as they resort once more to their infamous theatricality. They finish off Act Two re-enacting a scene from Judith's hit show

"Love's Whirlwind" with the help of Simon and Sorel leaving their guests to "*look on, dazed and aghast."*

The guests make a rapid departure the next morning, unwilling to "face another day like yesterday" (172). Their swift exit from the Blisses' world of chaos is not at all unexpected, nor will they be missed much. As their guests slam the door behind them, David ironically remarks, "People really do behave in the most extraordinary manner these days—" and the family continues arguing as if nothing had happened (179). *Hay Fever* is about living in the moment, selfishly pursuing individuality and being flippant about all the rest. Coward comments in his interview with Garland that, "[*Hay Fever*] doesn't rely on any particular international or particular political situation. It's just a frivolity, it could take place at any time." The fact that it was written and produced in the late 1920s can speak to the show business and carefree style of the time, but Coward acknowledges that people will always be willing to watch characters completely invested in their own ridiculous realities.

It is a pity that Coward's shorter plays did not receive as much publicity as the longer ones, but the length is not reminiscent of a shortening of stylistic resonance; it might arguably make the play stronger. *Fumed Oak,* one of ten one-act plays in the collection entitled *Tonight at 8:30*, still struggles with an individual fighting for self-expression, but does so in a shorter time frame. Coward handles the establishment of the calm and mundane much the same as he does in *Private Lives*, setting up the world the main character is being subjected to before they manage to break free. Henry Gow suffers at the hands of a family that does not appreciate him. His nagging wife, dim-witted daughter, and complaining mother-in-law are ignorant to their mistreatment of him, and this is made clear to the audience in their dominance of the first scene. Henry speaks a total of five words in Scene One and his disconnect from the rest of his family subtly alludes to the break from the peaceful and everyday that the audience has come to expect from the second portion of Coward's plays. In Scene Two, the audience relishes in Henry's backlash almost as much as he does because they have served witness to his abuse at the hands of his family.

Stop ordering me about. What right have you got to nag at me and boss me?

No right at all. I'm the one that pays the rent and works for you and keeps you.

What do you give me in return, I'd like to know! Nothing!... You're too busy

being snappy and bad-tempered even to say good morning… Here see this ham?

> This is what I think of it! [*He throws it at her feet*] And the tomatoes and the A.I
>
> bloody sauce! [*He throws them, too*] (51)

In a triumphant display of self, Henry breaks with the structure imposed on him by his family. In Coward's characteristic manner, Henry symbolically shatters the 'normal' drawing room conventions by throwing his dinner to the floor. But the dismantlement of what is proper doesn't stop there. Henry goes as far as to slap his mother-in-law across the face and lock the family in the living room with him, demanding they hear him out for once in his life. "While [the slap across the face] may express the audience's aggressions, it also continually underlines the idea that being shocked is itself the surest mark of the outsider" (Levin 119). In *Fumed Oak* the outsiders are the other members of the Gow family who are shocked at what they feel to be Henry's sudden break with temperament. In actuality, with the audience as his witness, Henry's outburst has been bubbling under the surface for a long time. Levin notes that *Fumed Oak* is one of Coward's only comedies in which he shifts his attention from the slender portion of upper-class aristocrats to a family that is considerably less well off, and almost "desperate" for money. He goes on to say that *Fumed Oak* is still a notable contender of Coward's repertoire "for it draws heavily on one of his chief talents—that for insult" (98). But it is not just insult that characterizes this play as one of Coward's bigger successes; it is also the undeniable breaking with the normal in favor of vibrancy, self-expression, and being true to oneself.

> DORIS: A fine exhibition you're making of yourself, I must say.
>
> HENRY: Not bad, is it? As a matter of fact I'm rather pleased with it myself. (51)

Henry dominates the final portion of the play, and he holds the audience's attention with the same self-exertion and personality as any character from one of Coward's longer works. Yes, Henry Gow is no Elyot Chase, but in his own working-class way, he justifies his right "to see a little bit of real life" (57). With the same ease it takes for Henry to assert himself, he slams the door behind him. While his abandonment of his family is more startling than directly humorous, it will never have the effect on an audience like Nora's slamming of the door at the end of Ibsen's *A Doll's House*. Henry's desertion, while not pretty, is at least as justified as Nora's and speaks in the same way to the persistence of humanity against the codified systems that attempt to define and control us.

The name Noël Coward conjures up an image distinct and unique in itself. He was a man who enjoyed living, and this celebration of self-expression is represented in his plays. It is

tempting to pigeonhole his works, to categorize and compare them with other playwrights, but this is not possible. Coward's eccentric style, his frivolity, his charm, and his love of language break through in his writing demanding that his works stand alone in a genre all his own.

Works Cited

Coward, Noël. "Fumed Oak." *Twenty One-Act Plays: An Anthology for Amateur Performing Groups*. Ed. Stanley Richards. Garden City, NY: Doubleday &, 1978. Print.

Coward, Noel. "Hay Fever, Private Lives." *Three Plays: Blithe Spirit, Hay Fever, Private Lives*. New York: Vintage, 1999. 111+. Print.

Coward, Noël. Interview With Peter Garland. "Omnibus: Noël Coward." *The Noël Coward Collection.* BBC Video. 1969: DVD.

Hoare, Philip. Introduction. *Three Plays: Blithe Spirit, Hay Fever, Private Lives*. New York: Vintage, 1999. Print.

Lahr, John. *Coward, The Playwright*. London: Methuen, 1982. Print.

Levin, Milton. *Noël Coward, Updated Edition.* Boston: Twayne, 1989. Print.

Influential Figures of Dance:
Loïe Fuller and Isadora Duncan

Emma Jayne Porter

As dancers and people who study the artistry of movement, we need to take time to reflect upon where we would be in dance without the brave and creative souls who helped develop the art. Dance is an art form that originated centuries ago, but has evolved, developed, and expanded across the world. Beginning as a classical, codified set of movements, dance is now a complex art and an expression of emotion. Reaching their full potential and influence on the world of dance in the late 1800s and early 1900s were Loïe Fuller and Isadora Duncan. While Fuller focused on elaborate lighting, the use of costume, and the revolution of classical dance, Duncan developed a modernized form of movement that was natural and heart-felt. The effects of Fuller and Duncan's artistic additions to dance are still present in classes and on stages all around the world.

Before discussing the impact of Fuller and Duncan, it is important to understand the new era in dance that was beginning to unfold during their breakthroughs. Just at the turn of the century in 1898, Bernarr Macfadden formed the beginnings of the Physical Culture Movement (Lancos). Outlining the path to a good life, Macfadden stressed the importance of exercise, natural food, fresh-air, sunshine, and a quiet mind (Lancos). The American public tested Madfadden's theories of a healthy lifestyle and realized the benefits of following the steps to a "good life." Paired with the theories of art nouveau, dance became a popular pastime and source of entertainment for Americans of all ages. Art nouveau was a development of the oriental design elements and woodcut prints, ukiyo-e (Lancos). Sinuous curves, linear pattern, opposition, contrast, and asymmetry were design elements of art nouveau that fascinated the public. Between the movement aspects of the Physical Culture Movement and the design elements of art nouveau, the late 1800s and early 1900s called for a revolution in dance.

Loïe Fuller, born in 1862 in Fullersberg, Illinois, was a sensation to the world of dance, finding her peak at a time that called for change in the art of motion (Brown, Mindlin, and Woodford 11). Marie Louise Fuller, originally a child actress, found her passion in skirt dancing. "Skirt dancing consisted of graceful, somewhat balletic steps," in which voluminous skirts

enhanced movement (Kraus, Hilsendager, and Dixon 116). After realizing that she could not appeal to the American audience, Fuller traveled to Europe in an attempt to make a career of her passion. In 1892, Fuller took the stage at the Folies-Bergère in Paris (Brown, Mindlin, and Woodford 11). Unlike Fuller's countrymen, the Parisians greatly appreciated her innovative style and expressivity. Fuller, nicknamed "La Loïe" by the French, became an overnight sensation (Lancos). The European audience looked past Fuller's mediocre technique, her less than refined skill, and her full-figured body; they saw harmony between movement, light, and material. Fuller shared her fascination of the play between light and cloth with audiences, expanding the art of dance beyond movement.

"Often characterized as the embodiment of the Art Nouveau movement," Fuller embraced the ideals of creativity, design, and expressivity (Chitwood 64). What Fuller did for the dance and theater world reaches far beyond what many other notable dancers have contributed. Fuller had the courage to oppose the accepted classical technique of dance, adding artistry and lighting effects. Imagine a dancer who is wrapped in 100 yards of beautiful fabric, surrounded by darkness, yet illuminated by only a few colorful lights (Kraus, Hilsendager, and Dixon 116). Her maneuvering of the cloth with long bamboo sticks extending from her arms is natural, playful, and simply artistic. The light reflects upon the material and highlights the brilliance of dance. By incorporating these technical and design elements into performances, Fuller left audience members awestruck. As dancers and theater-goers, we still see the long lasting effects of Fuller. Modern dance done today is often natural, just as Fuller's dancing was. All stage performances use brilliant lighting and costuming to complement movement; this was Fuller's exact goal as an artist and dancer. Some say "her legacy is ascending," as technology is changing and Fuller's theories are becoming more relevant (Chitwood 66). From codified to natural, from plain to complex, and from dark to illuminated, there is no doubt that Loïe Fuller had a significant impact on dance.

In addition to revolutionizing dance as an art, Fuller influenced other legendary figures, like Isadora Duncan. Like Fuller, Duncan sought to oppose the accepted technique of dance in the 1800s. Born in 1877 in San Francisco, California, Duncan began studying classical ballet at a young age (Kraus, Hilsendager, and Dixon 116). Finding that ballet did not satisfy her needs as an artist, and realizing that classical dance seems to "separate the gymnastic movements of the body completely from the mind," Duncan sought to develop a self-expressive form of dance

(Kraus, Hilsendager, and Dixon 116). Dancing in concerts and revealing her true self in New York and Chicago, Duncan found that she could not appeal to the American audience. Like Fuller, Duncan traveled to Europe in search of inspiration and an appreciative audience. Duncan continued to develop her own personal expressive form of movement and was particularly inspired by her travel to Greece in 1899 (Lancos). Inspiration finally came to Duncan while observing the natural movement of waves on the beaches of Greece. The natural lightness, forward-backward motion, and flow of water brought a new image for Duncan to formulate movement around. Duncan's European performances became light-hearted and natural; her hair was let loose, she wore draping fabrics, and she danced with emotion and artistry. Soon a popular and appreciated dancer, Duncan performed in Budapest, Berlin, Italy, Greece, and Russia, in front of audiences who were captivated by Duncan's innovative style of dance (Kraus, Hilsendager, and Dixon 117). Duncan "liberated the spirit of dancing," and enlivened movement ("Trailblazers").

Today people can observe the significant impacts that Duncan had on the world of dance. Prior to Duncan, classical technique seemed to be the basis of all dance that captivated audiences in America. Duncan broke away from codified, disciplined dance and developed four movement theories. First, she believed that all movement originated in the solar plexus, the spiritual center of the body (Lancos). Frederick Ashton portrays Duncan's movement style in "Five Brahms Waltzes in the Manner of Isadora Duncan." By watching Ashton's piece, one can observe the importance of the solar plexus in Duncan's movements, as seen by a constant uplift in the chest and rising of the arms ("Trailblazers"). The second movement theory of Duncan's is breath rhythm, coinciding with the third theory of wave rhythm. Finally, Duncan theorized sequential movement, and emphasized flowing from one movement to another (Lancos). Although these four movement theories have allowed dance to evolve into art, the most lasting effect of Duncan is incorporating emotion, passion, and inspiration into dance. Where would we be today without artistic expression in dance? Modern dance, contemporary ballet, social dance, and artistic dance would not have evolved without Duncan's bravery to oppose tradition. By setting the theatricality of dance aside and emphasizing creativity, emotion, and expression, Duncan opened a new opportunity to dance.

Dance is complex and full of history. Today it is a recognized art form that has numerous branches, forms, and styles. Unafraid and firm in their beliefs, Loïe Fuller and Isadora Duncan

expanded dance beyond the accepted codified traditional movement. Fuller introduced the power of lighting and supported the design elements of art nouveau. Duncan contributed natural, flowing movement that was inspired by emotion and the environment. Both legendary dancers and choreographers propelled dance in a new direction. Because of Fuller's and Duncan's influences on dance, movement has become a complex art full of meaning and creativity.

Works Cited

Brown, Jean Morrison, Naomi Mindlin, and Charles H. Woodford. *The Vision of Modern Dance (Second Edition).* Hightstown, NJ: Princeton Book Company, Publishers, 1998. Print.

Chitwood, Susan. "Loïe Fuller." *Dance Teacher Magazine*. July 2009: 64–66. Print.

Kraus, Richard, Sarah Chapman Hilsendager, and Brenda Dixon. *History of the Dance in Art and Education (3rd Ed.).* Upper Saddle River, NJ: Prentice Hall, 1991. Print.

Lancos, Jonette. "Dance History Since 1900." State University of New York at Geneseo. Geneseo, New York. 26 January 2011. Lecture.

"Trailblazers of Modern Dance." *Dance in America*. PBS. WNET, New York. 22 June 1977. Television.

Originators—Influential Figures of Dance:
Loïe Fuller and Isadora Duncan

Stephanie Willmarth

As dancers and people who study the artistry of movement, we need to take time to reflect upon where we would be in dance without the brave and creative souls who helped develop the art. Dance is an art form that originated centuries ago, but has evolved, developed, and expanded across the world. Beginning as a classical, codified set of movements, dance is now a complex art and an expression of emotion. Reaching their full potential and influence on the world of dance in the late 1800s and early 1900s were Loïe Fuller and Isadora Duncan. While Fuller focused on elaborate lighting, the use of costume, and the revolution of classical dance, Duncan developed a modernized form of movement that was natural and heart-felt. The effects of Fuller's and Duncan's artistic additions to dance are still present in classes and on stages all around the world.

Before discussing the impact of Fuller and Duncan, it is important to understand the new era in dance that was beginning to unfold during their breakthroughs. Just at the turn of the century in 1898, Bernarr Macfadden formed the beginnings of the Physical Culture Movement (Lancos). Outlining the path to a good life, Macfadden stressed the importance of exercise, natural food, fresh-air, sunshine, and a quiet mind (Lancos). The American public tested Macfadden's theories of a healthy lifestyle and realized the benefits of following the steps to a "good life." Paired with the theories of art nouveau, dance became a popular pastime and source of entertainment for Americans of all ages. Art nouveau was a development of the oriental design elements and woodcut prints, ukiyo-e (Lancos). Sinuous curves, linear pattern, opposition, contrast, and asymmetry were design elements of art nouveau that fascinated the public. Between the movement aspects of the Physical Culture Movement and the design elements of art nouveau, the late 1800s and early 1900s called for a revolution in dance.

Loïe Fuller, born in 1862 in Fullersberg, Illinois, was a sensation to the world of dance, finding her peak at a time that called for change in the art of motion (Brown, Mindlin, and Woodford 11). Marie Louise Fuller, originally a child actress, found her passion in skirt dancing. "Skirt dancing consisted of graceful, somewhat balletic steps," in which voluminous skirts enhanced movement (Kraus, Hilsendager, and Dixon 116). After realizing that she could not appeal

to the American audience, Fuller traveled to Europe in an attempt to make a career of her passion. In 1892, Fuller took the stage at the Folies-Bergère in Paris (Brown, Mindlin, and Woodford 11). Unlike Fuller's countrymen, the Parisians greatly appreciated her innovative style and expressivity. Fuller, nicknamed "La Loïe" by the French, became an overnight sensation (Lancos). The European audience looked past Fuller's mediocre technique, her less than refined skill, and her full-figured body; they saw harmony between movement, light, and material. Fuller shared her fascination of the play between light and cloth with audiences, expanding the art of dance beyond movement.

"Often characterized as the embodiment of the Art Nouveau movement," Fuller embraced the ideals of creativity, design, and expressivity (Chitwood 64). What Fuller did for the dance and theater world reaches far beyond what many of the other notable dancers have contributed. Fuller had the courage to oppose the accepted classical technique of dance, adding artistry and lighting effects. Imagine a dancer who is wrapped in 100 yards of beautiful fabric, surrounded by darkness yet illuminated by only a few colorful lights (Kraus, Hilsendager, and Dixon 116). Her maneuvering of the cloth with long bamboo sticks extending from her arms is natural, playful, and simply artistic. The light reflects upon the material, and highlights the brilliance of dance. By incorporating these technical and design elements in to performances, Fuller left audience members awestruck. As dancers and theatergoers, we still see the long lasting effects of Fuller. Modern dance done today is often natural, just as Fuller's dancing was. All stage performances use brilliant lighting and costuming to complement movement; this was the exact goal of Fuller as an artist and dancer. Some say "her legacy is ascending," as technology is changing and Fuller's theories are becoming more relevant (Chitwood 66). From codified to natural, from plain to complex, and from dark to illuminated, there is no doubt that Loïe Fuller had a significant impact on dance.

In addition to revolutionizing dance as an art, Fuller influenced other legendary figures, like Isadora Duncan. Like Fuller, Duncan sought to oppose the accepted technique of dance in the 1800s. Born in 1877 in San Francisco, California, Duncan began studying classical ballet at a young age (Kraus, Hilsendager, and Dixon 116). Finding that ballet did not satisfy her needs as an artist, and realizing that classical dance seems to "separate the gymnastic movements of the body completely from the mind," Duncan sought to develop a self-expressive form of dance (Kraus, Hilsendager, and Dixon 116). Dancing in concerts and revealing her true self in New York and Chicago, Duncan found that she could not appeal to the American audience. Like Fuller, Duncan traveled to Europe in search of inspiration and an appreciative audience. Duncan continued to develop her own personal expressive form of movement, and was particularly inspired by her travel

to Greece in 1899 (Lancos). Inspiration finally came to Duncan while observing the natural movement of waves on the beaches of Greece. The natural lightness, forward-backward motion, and flow of water brought a new image for Duncan to formulate movement around. Duncan's European performances became light-hearted and natural; her hair was let loose, she wore draping fabrics, and she danced with emotion and artistry. Soon a popular and appreciated dancer, Duncan performed in Budapest, Berlin, Italy, Greece, and Russia in front of audiences who were captivated by Duncan's innovative style of dance (Kraus, Hilsendager, and Dixon 117). Duncan "liberated the spirit of dancing," and enlivened movement ("Trailblazers").

Today people can observe the significant impacts that Duncan had on the world of dance. Prior to Duncan, classical technique seemed to be basis of all dance that captivated audiences in America. Duncan broke away from codified, disciplined dance and developed four movement theories. First, she believed that all movement originated in the solar plexus, the spiritual center of the body (Lancos). Frederick Ashton portrays Duncan's movement style in "Five Brahms Waltzes in the Manner of Isadora Duncan." By watching Ashton's piece, one can observe the importance of the solar plexus in Duncan's movements, as seen by a constant uplift in the chest and rising of the arms ("Trailblazers"). The second movement theory of Duncan's is breath rhythm, coinciding with the third theory of wave rhythm. Finally, Duncan theorized sequential movement, and emphasized flowing from one movement to another (Lancos). Although these four movement theories have allowed dance to evolve in to art, the most lasting effect of Duncan is incorporating emotion, passion, and inspiration in to dance. Where would we be today without artistic expression in dance? Modern dance, contemporary ballet, social dance, and artistic dance would not have evolved without Duncan's bravery to oppose tradition. By setting the theatricality of dance aside and emphasizing creativity, emotion, and expression, Duncan opened new opportunity to dance.

Dance is complex and full of history. Today it is a recognized art form that has numerous branches, forms, and styles. Unafraid and firm in their beliefs, Loïe Fuller and Isadora Duncan expanded dance beyond the accepted codified traditional movement. Fuller introduced the power of lighting and supported the design elements of art nouveau. Duncan contributed natural, flowing movement that was inspired my emotion and the environment. Both legendary dancers and choreographers propelled dance in a new direction. Because of Fuller's and Duncan's influences on dance, movement has become a complex art full of meaning and creativity.

Works Cited

Brown, Jean Morrison, Naomi Mindlin, and Charles H. Woodford. *The Vision of Modern Dance (2nd ed.).* Hightstown, NJ: Princeton Book Company, Publishers, 1998. Print.

Chitwood, Susan. "Loie Fuller." *Dance Teacher Magazine*. July 2009: 64–66. Print.

Kraus, Richard, Sarah Chapman Hilsendager, and Brenda Dixon. *History of the Dance in Art and Education (3rd ed.)*. Upper Saddle River, NJ: Prentice Hall, 1991. Print.

Lancos, Jonette. *Dance History Since 1900*. State University of New York at Geneseo. Geneseo, New York. 26 January 2011. Lecture.

"Trailblazers of Modern Dance." *Dance in America*. PBS. WNET, New York. 22 June 1977. Television.

Paving the Way to Equality: Voice Part and Gender in the Operas of Mozart

Devon Borowski

Over the years since its inception in 1600, the genre of opera has seen the rise and fall of empires, changing social morés, and the evolution of musical styles. It seems that while the patrons of opera were finding their seats, everyone else was outside, either admiring or scoffing at them. Indeed, the world continues to change while the curious art form of opera has grown socially irrelevant and disconnected. Yet, search the major playhouses of the world and you will be hard-pressed to find something on stage from the eighteenth century, especially something filling seats. Major opera houses, on the other hand, would face uproar if even one perennial favorite from the eighteenth century were taken off the roster. The real genius of opera composers is their use of every aspect at their disposal, both musical and dramatic, to communicate ideas and emotions to their audience. The artful deployment of voice parts is one means of achieving these ends. The various voice parts lend themselves to harmonic and sonorous variety, but more importantly, dramaturgically speaking, they serve as means of characterization. In the first century of opera's existence, the voice parts came to be associated with specific dramatic roles, and these changed over time, reflecting changes in society.

Acknowledged differences among voice types, where training is concerned, stretch back to at least the seventeenth century and for the most part have stayed the same. The 19th century Italian pedagogue Francesco Lamperti recognized the specific needs of the soprano voice (Baker, Slonimsky 338). In opera, the assignment of voice part to character may be used as a societal barometer of the day. While the dramatic purpose of each voice part has changed over time, the most intense developments have taken place in the mezzo-soprano and baritone roles.

The attitude toward the high voices, soprano and tenor, has traditionally been that the higher the voice, the closer to God. These two are models of virtue sent on a quest from heaven, destined to end tragically and be rewarded in the afterlife. They are the

perfect embodiment of courtly love, led by an intense sexual desire manifested musically in climactic cadenzas, a desire seldom consummated. This can be seen, for example, in the relationship between the title characters in Wagner's *Tristan und Isolde.* While the argument could be made that society today has outgrown such archaic and stereotypical views on sexuality, effectively outdating these characters, it would be rather naïve to think that a royal bastard or a corporate affair would not cause controversy and considerable media coverage, even today. The Judeo-Christian worldview may be dismissed in post-modern thought, but still has quite the hold on Western society and therefore what is portrayed on the opera stage remains relevant. These voices often embody qualities of innocence, even naiveté, in the face of danger, often trusting those that they should perhaps not. In Donizetti's *L'elisir d'amore*, for example, the tenor, Nemorino, trusts a traveling salesman to sell him a love potion.

If high voices are celestial, it follows that low voices must be terrestrial. Least often featured in opera, the contralto and bass voices are as rare as the wisdom that they so often symbolize. And yet, their arias are often the most coveted. At times, they may reveal the crux of the plot, as in Verdi's *Il Trovatore*, when Azucena describes her son's fraternal connection to the Count he wishes to fight in a duel in her aria "Stride la vampa." At other times, they signify the defeat of evil, as in Mozart's *Die Zauberflöte*, when Sarastro sings "O Isis und Osiris," a prayer to help the warriors win their impending battle. They are usually the oldest characters in the story, their ages ranging from old to antique and, occasionally, to primordial. In Wagner's *Das Rheingold*, for example, Erda is the goddess of wisdom and the earth. She is so old and in tune with the earth, that she exists only as a torso jutting out of the ground. These characters are as ancient as the world itself and possess the wisdom of the ages, wisdom they have earned through having survived tragedies and trials. Sometimes they seek out a tyrant and sometimes they themselves are sought out for counsel by brave heroes, as seen with the character Gurnemanz in Wagner's *Parsifal*. In this sense, they act as equals and opposites to the high voiced characters, both after the same ends but with very different jobs to achieve them.

It may seem odd that these wise characters may be found in the lowest register of the human voice. Depths are, after all, associated with the underworld, playground to the

demons so feared in the Western world-view. To fully understand, the Judeo-Christian view must be seen through a purely historical lens. The proliferation of Catholicism, of monotheism itself, depended upon gaining followers from polytheistic cultures, which ultimately meant allowing pagan imagery and thought into the sphere of Catholicism (Halliday 318). For those people, were they Greco-Roman, Arabic, or pre-Colombian, a considerable focus was placed on the earth, as a wise maternal figure, and the underworld was viewed not as a torturous place but as destination of mortal souls after death. Certainly it is no accident that their deep voices are connected with the literal depth of the earth. Whether composers were aware or not, there seem to be strong pagan influences hidden within their deeply Catholic art form.

The medium voices, the mezzo-soprano and baritone, remain. Unlike the extremes of heaven and the earth, the roles assigned to these voice parts have changed over time. A look at the characters from various operas will show different directions, new patterns of thought emerging. These characters often serve as the volatile force that fuels the actions of the plot. The baritone voice may represent a character who is the conniving helper, as in Donizetti's *Don Pasquale*, the anti-hero, like in Menotti's *The Old Maid and the Thief*, the lover, or even the fighter, responsible for goosing the story forward. The mezzo-soprano voice may characterize a complex heroine, such as the title role in Bizet's *Carmen*, a temptress, like the lead in Saint-Saëns' *Sampson and Delilah*, or a jilted lover. These medium voice parts provide the human element in the opera. It then seems natural, that they have become the most common voice type. The boundaries of heaven and earth have already been established and the domain between them needs to be populated, after all, and ordinary people, caught up in ordinary human dilemmas, make the most relatable population. In many respects the mid-range is the perfect location for these characters.

Not too high or too low, not in touch with the divine in either way, nor on a mission from God, they are not out to help anyone but themselves. This is not to say they are evil, but they put their own interests before those of other people. Even characters that are traditionally seen as wicked, such as Iago in Verdi's *Otello*, upon closer dramatic examination, are interested in harming others only as a way to fulfill their own desires. It is a perceived slight that causes Iago to fly into a rage and plot revenge, and though he

chooses to be vengeful, the cause is far more important to him than the effect. These characters tread the fine line between good and evil that ordinary people walk each day. Their allegiance is not to a malevolent organization, nor even a benevolent one. They fight for themselves or those that they love, be it their family, as is the case of Germont in Verdi's *La Traviata*, or the people of a nation, like Le Comte de Saint-Bris, fighting for the Catholics in *Les Huguenots*. While the characters change from story to story, they all fundamentally have the same interests in mind, driven by their own desire for fame, liberty, sex, or wealth. Their self-interest often makes them the most interesting characters, who act as the dramatic catalyst for a story.

The roles for the medium voice, then, are typically the most malleable and have no set archetype, instead affording the composer a range of possibilities. Consequently these characters often reflect the spirit of the opera most intensely. What is commonly known as the *tinta*, or atmosphere of the opera is not only set by musical style, thematic material, orchestration and the like, but by the personality of the medium-voiced characters. As the most fundamentally human characters, they closely reflect the society of the day, its morés and politics.

Though the various voice parts have long existed in human history, their first dramatic codification appears in the *Commedia dell'arte* tradition, beginning in mid–sixteenth-century Italy. The six main character types, improvising on a few fundamental story lines, form the basis of the theatrical tradition to come. The six main characters, each wearing a specific mask, serve three basic roles: the servant, the master, and the lover, one of each gender. Over time, these stories were written down for reuse, and the characters expanded upon. The master developed into the miserly bass merchant, as in Pergolesi's *La Serva Padrona*, and the wise contralto, into the matriarch. The lovers became the amorous tenor and soprano couple, as in Donizetti's *Don Pasquale*. The servants grew into the mischievous baritone butler and the coquettish mezzo-soprano maid. Societal changes affected the details of their personalities and predicaments, but their basic dispositions have remained unchanged over the years.

Much has been written about "the musical shot-heard-around-the-world," the premiere of Jacopo Peri's *Dafne*, in 1597, which gave birth to opera as an art form. More enduring though, is the work that cemented opera as a legitimate genre, Claudio

Monteverdi's seminal work of 1607, *L'Orfeo* (Kelly 4)*. Given the relatively limited resources at the composer's disposal, there is no question he achieved a great deal in a single piece. For our purposes, however, the most important aspect of the work is what is *not* there: the medium voice is completely absent, also in his subsequent works. As the genre of opera evolved, high voices, especially *castrati*, were favored. Vocal spectacle was more important than realism in the portrayal of character, and the complex lines of instrumental compositions were transferred to the voice. The suspension of disbelief required of opera audiences during the Baroque era may seem ridiculous today, but as the general public grew interested in the genre, tastes began to change, and the music evolved to meet the new audience's expectations.

Baroque *Opera Seria* depended on certain strict conventions. George Frederick Handel composed over forty operas with rather rigid rules governing style, aria type, and of course, voice part (Burrows 121). The tenor and soprano voices were considered the height of operatic majesty, and the focus of the works as a whole. The strict gender conventions of the time were, perhaps counter intuitively, reinforced by these voice parts. Characters representing the height of Baroque masculine and feminine ideals, such as Cleopatra and Caesar in *Giulio Cesare*, sing in the same octave. This gender confusion was further muddled by classical plots extolling traditional masculine and feminine traits—for example, aggression and subservience, respectively—as virtues (Bokina 47). Themes of duty and honor trumping individual desire served to ensure the continuation of the aristocracy, as their very existence was intrinsically tied to these and other traditional values.

By the mid-eighteenth century, the extreme passions of the Baroque era were no longer in vogue. Instead, cool wit combined with understated elegance prevailed over the art and society of the time. Enlightenment ideals were sweeping through Europe, and reason governed science, literature, and even music. The catchwords of the day were liberty, fraternity, and, with the rise of the new middle-class, equality. No more perfect embodiment of these ideals in music can be found than in the operas of Wolfgang Amadeus Mozart. One of the pioneers of the medium voice, Mozart cast some of his most beloved characters as baritones. And yet, while Mozart pushed musical equality for men, he withheld this same right for women. In his operas, there are few mezzo-soprano

characters in the same context as the baritones, a lack of vocal equity afforded to the female characters. While Mozart's male characters are free to sing in a high, low, or medium voice, the women are by and large only given the high option. This reflects the newly burgeoning middle class society of the day, which afforded more opportunities for men, but still few for women.

Mozart wrote four great baritone roles at the zenith of his career, in the three Da Ponte operas and his last *Singspiel* with Emmanuel Schikaneder. The first of these, *Le nozze di Figaro,* contains a wide cast of characters encompassing the various *commedia dell'arte* types. This opera is more concerned with domestic affairs than those of the state. Here, Mozart takes the roles of servant, master, and lover, and blends and flips the traditional hierarchy of characters inherited from the Baroque period, but with disposition acting as a constant. The servant now becomes a gallant lover as well, and the blundering master is infused with a good deal more cunning. This means that not only may the lover outwit an authority figure, but a servant may directly defy and undermine his own master. Given the domestic focus of the work, there is considerably more room for a variety of characters and situations with these innovations in mind.

The most significant exception to Mozart's neglect of the mezzo-soprano voice is the britches role of Cherubino, in *Le Nozze di Figaro*. This character, though starkly different than Count Almaviva, his fellow medium-voiced character, still has much in common with his master. Cherubino is a young servant, a dandy, and most importantly, he's naïve about the world. It must be noted that Mozart is highlighting Cherubino's effeminacy by using a female voice (Brett, Wood, and Thomas 41). This is in no way a ploy for female vocal equality, but just an old trick of *Opera buffa*, employed for the sake of characterization. What is clear through his text is that Cherubino is as interested in the opposite sex as his master. Once again, their commonality on the surface seems trivial, but the human desire for sex, a self-centered pursuit, gives them a realistic quality. While the Count is viewed as the villain for this reason, Cherubino is excused as a hormonal teenager, something with which audiences have long been familiar. This also accounts for his being a mezzo-soprano, and not a high voice. He is far more strongly identified with the overactive hormones of his age than with its innocence.

The other mezzo-soprano role in *Le Nozze di Figaro* is that of Marcellina, the old maid looking for a husband in Figaro. Like her fellow medium-voice counterparts, she too is looking out for her own carnal needs through any means possible, even if it involves taking legal action. While the quest of an old woman seeking companionship may look purer on paper, her attempt at coercion through invoking the contract, as well as her heated argument with Susanna, shows her fierce determination to gain Figaro at all costs, with no evidence of real feeling for him. It is reasonable to assume that Marcellina is, like Cherubino and the Count, looking to fill an immediate physical need rather than seeking a lifelong partner. She too, as both older and female, is reviled for not knowing better, and for flaunting her sexual appetite in ways unbefitting of a lady. It seems, once again, that Mozart's democratic ideals stopped at women; sexual freedom is reserved for the male character.

By contrast, the other characters involved in the opera's romantic shenanigans are focused on marriage, on the union more than the physical act. Like Cherubino, Figaro also takes on a new role for his voice type, as a bass playing the lover, a role traditionally seen as the tenor's purview. Immediately, it is clear that this is not a typical love story. With the couple already betrothed and nothing to vanquish, it takes less divine inspiration and more cunning to unravel the plot. A clever man, Figaro is able to outwit the Count with knowledge of human nature that he presumably gained from generations before him. In his cavatina, "Se Vuol ballare," he demonstrates his deep understanding of the various classes by appropriating their dance structures, employing formal elements of both upper and lower classes dances, as well as a dotted rhythm (Hunter 257). With all of this knowledge, he sounds a lot like a deific personage, much like Sarastro, but thrust into a different situation. He, too, wishes to make everyone happy; while looking to marry his betrothed, Figaro watches out for Cherubino and the Countess, and eventually, ensures the betterment of the whole estate.

Among those under Figaro's protective watch is Susanna, his betrothed, who is firmly set in the realm of the lyric soprano. For the female love interest, this range is not surprising, but her situation is more complex than just a girl in love. Though the Count attempts to take advantage of her, she is not the weak-willed and loose-moraled servant, but a powerful woman able to take care of herself. She not only fends off the Count's

attacks, but also keeps Figaro calm, while simultaneously acting as confidante to the broken-hearted countess. Despite having to balance these various actions, she remains calm, her musical lines steady and *legato*, not frantic and excited like many others' throughout the opera. Yet, as a high soprano her power to effect change is limited; she is unable to completely take charge and get things done herself, like a mezzo-soprano would.

As befits her noble station, the Countess Almaviva is a high voice part, though, like Susanna, she too, is far wiser than her type might suggest. Her wisdom comes from harsh experience with her husband, but her soaring lines characterize her ability to transcend her hard lot in life. Of all the characters, she is the most closely linked with the divine, through the cool acceptance of her fate and her generous bestowal of forgiveness in the finale. She is not unlike a biblical hero being tested, and while enduring a cruel life with her husband, she never ceases to take care of those around her. Cherubino not only identifies her as his godmother, but also worships her and the divine trials she puts him through in song. At the end, her simple, floating melody doles out heavenly mercy, while reminding her husband, of course, that she is far nobler than he.

While Mozart did not create these archetypes, he did develop them, especially the male roles. By giving the bass the role of the lover, and allowing the servant to outwit his master, he demonstrated that the old order was on its way out. Unfortunately, the women's roles were not developed to the same degree. The mezzo-sopranos are either the wrong sex, in the case of Cherubino, or virtually sexless, as in the case of the aging Marcellina. Susanna, a strong, proto-feminist character, is consigned to the high soprano range, her stereotype as a beautiful lover trumping her potential for change.

The Opera seria/buffa hybrid, *Don Giovanni*, extends the dramatic range of the male voice parts, even as it, once again, denies such developments to women. At the heart of the story is the titular Don Juan character, a powerful noble with an insatiable lust for women. With this opera, Mozart demonstrated that a medium voice—for Don Giovanni is a baritone—could successfully lead the story of an opera, without audiences feeling they had missed out on a tenor. As in his other works, this vocal freedom is not extended to the female characters, a state that is intrinsically linked to the prevailing

social inequality of the day. Of course, it also comes as no surprise in a story about one of history's most famous rapist (Waldoff 165).

This opera is about how ultimate power corrupts absolutely. The audience is offered no insight into why the Don does what he does; only that he always has and always will be a sociopath. There is no credible defense offered for him in the story, and even his own servant condemns him mercilessly. He does not even attempt to defend himself and, given his social station, feels no need to. At first, this seems surprising, given the meticulous characterization that Mozart usually provides. The simple answer to this is that Mozart did not want the Don to be a sympathetic character in the opera. This however would be a departure from his other characters, as one of Mozart's strengths as an opera composer is his understanding of human nature. It seems that even the anti-hero deserves some pity, especially from a composer personally familiar with sexual promiscuity (Stafford 126).

The scene in the graveyard is the key to understanding the character of Don Giovanni. There we see his uncontrollable downward spiral to his disastrous fate. In the end, he is an addict that everyone is constantly enabling. The Don cannot possibly see where his choices are taking him, and is incapable of censoring himself. In this, he shows his complete disrespect for the newly arising social order. His adversary, the *Commendatore* has a militaristic position that must have been earned through hard work and education. Giovanni is willing to take on the *Commendatore* who, though perhaps a lesser noble, is certainly more deserving of respect than he. Don Giovanni's lack of respect demonstrates that he takes the old order of birthright for granted, and doesn't understand the newly changing social order. Mozart characterizes the *Commendatore* as the wiser and more respected of the two by giving him a bass role. The Don disregards the traditional connotations of this voice part in favor of his birthright, the source of the power that ultimately corrupts him (Bokina 42).

Don Giovanni's inability to see and unwillingness to accept the social changes afoot leads him to his horrible fate. He is punished not only for this, but for his crimes against gentile society. Not only is he a rapist, but the very embodiment of sexual chaos. His target is women, with no more specific delineation beyond that aim, and his voice part reflects his lack of preference regarding social class. As a baritone, he is able to sing

in the upper or lower echelons of the voice, taking on whatever timbre and character is needed to seduce the woman of the moment. With his vocal flexibility, he is an "equal opportunity" seducer, attracting women in a way that no bass or tenor could. He does not stratify women based upon their appearance like other men, because he is open to any kind of female. What he lacks in an understanding of political equality, he makes up for in his total lack of discrimination regarding his female prey.

The women of the opera, already dehumanized by the Don, are likewise stripped of their gender equality through their voice parts. Each of the three principal women is, once again, restricted to the soprano range, though there are sharp differences among the three characters. In this way, the voice distribution takes its cue from *Opera seria*, rather than *Opera buffa*, where women are given more range, albeit within the domestic sphere. Through the three characters, the audience is introduced to three classes (a peasant girl, and two noblewomen of seemingly different rank) and three different types of victims (one who must be forced, one who is mildly coerced, and one who gladly succumbs). As high sopranos, their vocal homogeneity reminds us that, to the Don and the society in which they live, these women are all the same, despite their individual differences.

Dramatically, each woman is stereotyped differently, based on her own experience. The Don views Zerlina, the poor peasant girl, as an easy target, and assumes her to be sexually free because of her confidence and lower class. On the bottom rung of the social ladder, she must take charge of her own life and, even once wed, will have to manage a household. Her self-assurance is viewed as tantamount to sexual promiscuity, especially in a world where women are taught to be silent onlookers. In her aria, "Batti, batti, o bel Masetto," the audience gets to see her ability to take charge of a situation. In this aria, she begs her new husband to beat her for her crimes against him. Not actually wanting or expecting to be beaten, she is simply using an alternative tactic to gain his forgiveness (Smart 62). She does not intend to supplicate herself before him as punishment, but is in fact using her body, the mere idea of touching her, as a way to arouse him until he can no longer remain upset with her. By the standards of the eighteenth century, such a capable woman is still not deserving of praise. Though Don Giovanni is the anti-hero of the story, he is not a villain on account of his sexuality, but

only because of sexuality expressed to excess. Zerlina is not afforded the same leniency or vocal equality, even though she is merely using sexuality as a means to an end.

Donna Anna, the victim the audience meets first, is immediately perceived as weak. For the entire opera, in fact, men surround her, offering to help, which only seems to take away from her own abilities in equal measure. She is viewed as a "victim" incapable of helping herself, someone who has been permanently damaged. As a soprano, she looks to her tenor fiancé, Don Ottavio, for support, and cannot function independently, like a mezzo-soprano would. Donna Elvira comes off as even more pathetic, following Don Giovanni around like a puppy dog. A clear victim of the Stockholm Syndrome, she cannot help loving him despite his crimes against her, and is repeatedly willing to take him back. She is easily fooled by the Don's empty promises, but once again, as a soprano, does not possess the strength of will to go her own way. As a high soprano, more concerned with romantic love than anything else, she is the ultimate victim.

While Don Giovanni is ruled by his animalistic passions, we see his foil in Don Ottavio, Donna Anna's fiancé and protector. Anna is flanked on the one side by him and on the other by her father the *Commendatore*, both men, at either end of the vocal spectrum, willing to give their lives for her. At the low end, her father, the bass, represents the wisdom acquired over the long and illustrious career that earned him his title. At the high end, the tenor Don Ottavio, is primary concerned with his fiancée's welfare and their romantic love. Though of noble birth, his high tessitura is less indicative of his station than of his polished attitude, even in revenge. And though the text of his "Il mio Tesoro" speaks of protecting his beloved, his delicate *fiortura* suggests less an action-oriented soldier than a high-minded dandy. While Don Giovanni easily gives in to his lustful nature, Don Ottavio carefully composes himself in preparation for his complicated and clearly doomed stratagem.

Meanwhile, Don Giovanni has perfected his manipulation of the outside world through his servant, the bass, Leporello. The most useful tool for Don Giovanni, in fact, is that his home key of D-minor is the relative minor of, and therefore related to, Leporello's home key of F major. This related-key device, that allows the pair to change places smoothly and escape detection, reveals more than just cowardice. The fact that

Leporello is a bass role serves two purposes. First, as a smart servant, he has grown wise through protecting his master. In the same way that a personal assistant will know tricks of the trade better than their employer, Leporello knows Don Giovanni better than he knows himself, and is weary of his selfish schemes. This is relayed to the audience at the beginning of the opera in his "Notte e giorno faticar." In fact, when he starts to sing, he is almost too tired to move out of a two-note motif, and is only motivated to begin a more legato phrase by his increasing frustration with his master. The other purpose is when Don Giovanni wishes to switch roles with him for his own safety. Not only does the crowd get a different target to chase, but one who is earthier and more grounded, wiser, in his deep, bass voice. Whether this distracts, intimidates or placates them, it serves Leporello well, so that he manages to escape each time he is hotly pursued.

In *Don Giovanni*, the victimization of the three women is mirrored in their vocal range, the soprano register a musical parallel to their dramatic oppression. Though they react differently to their seductions, none of them are given the vocal freedom afforded the men of the opera. The title character is perhaps the ultimate representation of the baritone as a self-serving figure, going so far as to derive pleasure from controlling others through rape. At the same time, he demonstrates a political system in decline, challenging the *Commendatore*, the representation of the new political order within the story. Here, dark themes serve as a backdrop to a struggle for equal rights, a struggle that once again excludes women.

The least known, and last, of the Da Ponte operas, the perennially charming *Così fan tutte*, involves one of Mozart's most fascinating vocal schemes. Of course the themes of fiancé swapping and the inevitable disloyalty of women are problematic in a modern context, but audiences of the twenty-first century are naturally more inclined to see these as problems than the original spectators (Waldoff 189). Throughout the nineteenth and early twentieth century, the opera's partner-swapping motif was considered too salacious for the stage, and was not performed often, but today, it is an issue that we must confront. What was originally intended as a harmless farce actually reveals much useful information about the opinions on women in society at that time.

The opera's characters naturally divide into three pairs; the soldiers, the fiancés, and the troublemakers: Guglielmo and Ferrando, the baritone and tenor duo, Dorabella

and Fiordiligi, their soprano and mezzo-soprano lovers, and Don Alfonso and Despina, the bass and soprano facilitators of mischief. The façade created by these troublemakers also serves to divide the six into two groups: the three men versus the three women, creating a war between the sexes, a common *Opera buffa* theme. Given this division, it might seem as though the distinction between tenor and baritone, as well as soprano and mezzo-soprano is superficial and inconsequential, but closer examination reveals important differences.

Though initially the two soldiers may appear to be a package deal, it should come as no surprise that Mozart created a unique personality for each of them, best exhibited in their arias. After the initial introduction of the Albanian gentlemen, it is Guglielmo who cannot help but push the envelope to gain the women's affections. Though he is on the end of the bet that stands to lose should the women succumb, as the baritone he has other motivations.

The gender-centric themes of the opera were common for the time, and still are, as they broach, but do not explicitly mention, the topic of sex. In the cosmopolitan cities of the day, sex was just as much of a universal preoccupation as it is today, especially in a society where an entire class of people had no profession nor need of a day job. Guglielmo, as high society's representative, is much more interested in broaching the subject of sex for his own amusement than concerned with winning a bet. In his aria following their introductions, "Non siate ritrosi," he points out to the women very specifically what they have to physically offer. He is more than happy to see them blush at his descriptions. With his words, he is planting the first seeds of infidelity in the women's heads.

When the women finally acquiesce, it is first Dorabella that falls for Guglielmo. Naturally, he is overjoyed, even though it means that they are losing the bet. Not only is his sexual prowess rewarded, the prize is his best friend's fiancée. He has the admiration of his target, but has also outdone his friend. Furthermore, his betrothed, to this point, has remained faithful to him, and after initial sympathy, he cannot help but revel. While the audience may feel sorry for Ferrando, it is the mixture of *Schadenfreude* and intrigue that Guglielmo displays that the audience most identifies with.

In a society with as large a noble class as Vienna's at the time of the opera's premiere, the somewhat dubious preoccupation with courtly intrigue was to be expected. It is natural then, to have the baritone provoking such mischief. When Fiordiligi eventually succumbs to Ferrando's repeated attempts, instead of expressing anger, Guglielmo is genuinely agitated, as if he cannot understand the infidelity committed against him. The vain blindness he exhibits demonstrates his human fallibility once more—a state in which, again, the audience can see themselves.

The tenor comrade, Ferrando, has a much different outlook on the situation. His fiery reaction to his fiancée's infidelity is indicative of a much purer nature. Though he entered into the bargain with Guglielmo, when he sees firsthand how losing the bet looks, his constitution is revealed to be much different than that of the other men. Not only does he lack the stomach to watch his betrothed be unfaithful, he is enraged that his own picture has been given to Guglielmo (enclosed within the heart locket). It is the lack of honor in the situation that triggers his anger, and Guglielmo's gloating that causes him to then press on and succeed in seducing Fiordiligi, something he would surely have preferred not to do. Not only can we see Ferrando's unease with the scandalous situation, but also Guglielmo again acting as a catalyst, causing Ferrando to try his seduction once more after being teased so much.

Though Guglielmo continues to push his luck, ultimately Don Alfonso, the wise philosopher, is to blame for the whole situation. As the bass, he is presented as the wisest character of all, as well as years older than the rest of the characters. In fact, it is he who understands the ways of women down to their movements, and can impart that knowledge to the two young soldiers. Not only does he know that the women will falter, but he can play on the actions of each character, anticipating that the men will act as they do.

The women in this opera, though perhaps less physically harmed than in the others, suffer by far the most direct plight. The famous line and title, *Così fan tutte*, translates literally as "so do they all," directly referring to the infidelity and heartlessness of women. Don Alfonso, already shown as the wise man, knows for a fact that this is the way of all women, and he is proven correct. The soldiers feel no guilt about entrapping their fiancées, and then consider themselves magnanimous to forgive such dishonest and

undeserving women. It is not surprising then, that all three women have the tessitura of sopranos, restricted to the high range, when women of the day were viewed as more alike than different (Brown-Montesano 218).

When it comes to Despina, however, there are a number of significant distinctions. What is important to remember is that she is not in fact, paired with the other two women over the course of the opera, but with Don Alfonso, the bass. The first difference is that she is a maid and as a low-class woman, by definition is assumed to have loose morals. More significantly though, she goes along with the joke being played not only on two fellow females, but her own employers, no less. She is not just participating in this culture of female shame, but is helping to lead it. As a result she is "one of the boys" and has a certain level of status that goes with that. Not surprisingly then, Despina has the lowest tessitura of all three women. Given her soprano voice part, this means she's torn between two loyalties. She wishes to be loyal to her mistresses, but the other option earns her more favor and attention from the men than she would normally get. Despina, though just a maid, has her own set of advice for her ladies with which, ultimately, she genuinely wants to help them. Her association with Don Alfonso, and their role in the mischief, brings her tessitura down, though she fundamentally remains a soprano (Ford 170). Her relatively low tessitura also serves to reinforce her role as a notary when she must trick the couples into signing a mock marriage contract.

All of these musical personalities work together, in conjunction and opposition, in the Act I Sextet, "Alla bella Despinetta." Here we hear the three groups of two divided into polyphonic duets in the case of the soldiers and fiancés, and in the case of Don Alfonso and Despina, their own roguish line. For the duration of the number, Guglielmo and Ferrando sing homophonically together, usually in thirds. They are working in complete conjunction, and have gone so far as to plan out their entire speech. The homophonic texture is primarily true for Fiordiligi and Dorabella as well, though at times they break off from one another and toss the line between them. At that point, the two women go from being on the same page to speaking for one another entirely, completely trusting in the other's judgment.

Don Alfonso has little to do on his own but introduce the two gentlemen. After that, he leaves them to their own devices, but what lines he does have afterwards are all

directed towards Despina, either instructing her or laughing with her. Despina's first homophonic line is with the soldiers in the three-part hymn-like section, in which they beg the women for forgiveness. The beauty of this section is Mozart's ability to work out how the thoughts of each character intersect at any given moment. Though Despina has been reprimanded for allowing strange men into the house, she has the same musical sentiment as the men who are trying to get into the women's good graces.

Eventually, everyone sings simultaneously, and after a long polyphonic section, they all end in agreement, singing only two different sets of words between them. While the ladies say that "*Innocente è questo cor*," that their hearts are pure, the rest speak of "*Quella rabbia e quel furor*," the impending anger and furor to come. Though they say different words, everyone agrees musically, singing in unison until the very last authentic cadence. While they have different takes on the situation, they all understand the situation through the same sharply chromatic phrase.

In *Così fan tutte*, women fare even less well than in the other two Da Ponte operas, if that is possible. Not only are Fiordligi and Dorabella seen as a single entity to be manipulated, but their mutual disloyalty is assured from the very beginning. This lack of fidelity is then encouraged by the rest of the characters, most of whom are in the position to care for their well-being, but don't. Despina, the character given the greatest vocal equality of the three women, through her relatively low tessitura, is already seen in a more masculine light because of her lowly station in life. As a mere maid, she cannot afford the feminine languor the other two women enjoy. What appears to be equality, then, is actually a punishment for her lack of femininity. Despina is, ultimately, a recipient of the same inverted equality that Don Giovanni offers his victims.

As far as equality for the sexes, Mozart's penultimate opera, *Die Zauberflöte,* goes one step further in each direction. While the baritone is afforded vocal equality, and the least outwardly selfish character, the highest soprano is actually condemned for having any power of her own. The work is not unlike a fairy tale with its archetypal characters and underlying moral structure. Papageno the happy bird-catcher, cares not for affairs of state, nor would he be much good at them if he were to. Nonetheless, he is given musical equality, that is, music to sing within the baritone range, sitting in his own tessitura. He is not expected to sing the highs of the tenor or the lows of the bass, but is

recognized as deserving his own legitimate voice part. His simple, *buffa* style arias supply the audience with tunes that they can hum long after the performance has finished.

Right away he is relatable, not being held to the extremes of the other characters, but out for simple happiness through his own song. More than just the comic relief that he provides so well, Papageno represents everyone that Sarastro is fighting for, the citizen that has long been oppressed by the powers that be. This oppressive power manifests itself in the character of the Queen of the Night, whose vocally taxing, stratospheric coloratura demonstrates her unquestionable power over the citizens of her territory. Just as she struggles and completes complicated musical lines, so must they struggle and complete her tasks for her. The complicated structure of her aria "Die Hölle Rache" demonstrates her immense intellectual capacity, which, for a woman at the time, immediately made her dangerous. The high religious and behavioral standard that she expects her subjects to meet is equivalent to the high tessitura that she must maintain.

Though Mozart gave Papageno his own tessitura, only Tamino can truly free him from the Queen's rule. It is the medium voice that truly wins alongside Tamino and Sarastro. Papageno is no longer subjugated as a direct employee of the ruling class; he is not held to the same standards of honor that the high voices are. No more will he be punished for telling lies—as he was by the Three Ladies of the Queen—but is free do to as he pleases, governing himself without the threat of moral retribution.

This new order, brought about by Sarastro, was heavily influenced by Mozart's ties with Freemasonry and its democratic ideals (Donington 71). More important than his position as leader of the council of Priests of Isis and Osiris, is Sarastro's humanity and wisdom. He is not a god but, as a bass, he is of the earth, and therefore is able to free the citizens from the Queen's heavenly oppression. In this ancient setting, it is clear that heaven does not imply benevolence and the earth is far from impotent. Sarastro's famously low notes reinforce his strong connection with the wise earth. Not concerned with complicated rules that the Queen decrees from heaven he, as the wisest character, is far more interested in reason and equality. Like a good democratic leader, Sarastro can see beyond his own needs and to what people need for themselves, such as Papageno and his family.

At the other end of the spectrum, Tamino and Pamina, the tenor and soprano, are given the almost unattainable goals of goodness and divine love. That Tamino falls hopelessly in love with Pamina does not imply a lack of understanding of these goals on his part, but sets him up as a kind of super-human being, made for the kind of platonic love that people look for in great leader-couples. By contrast, Papageno and Papagena's final duet gives the impression that they leave the stage to consummate their love. Given that Papageno is a baritone, and therefore more concerned with himself and his own corner of the world, this makes sense. This sort of corporeal focus is what society needs in its citizens, but not its leaders. For political couples, it has always been true that they are partners in both their private and public lives. Justinian and Theodora, William and Mary, Barack and Michelle Obama all embody this idea, that society needs more than one to lead a country, and a married couple is ideal for their balanced relationship. As the two high voices, Tamino and Pamina's love represents the ultimate romantic relationship and cannot be consummated before a proper conclusion in the form of a wedding.

In reality, though a gender-balanced duo is good for Western society, it is unlikely that the female would actually have much public power. It is fitting then, that this new sort of musical equality does not extend to women in Mozart's opera. The principal mezzo-soprano role is rather noticeably missing from this opera. Pamina, for example, while expected to rule equally with Tamino, has not been subjected to tests of her own, but instead follows him. She acts as the "great woman" behind the "great man," but is not great in her own right. Furthermore, the source of evil in the story, the one making selfish decisions for her citizens, is the Queen of the Night. More than just a lyric soprano subjected to a high tessitura like Pamina, her infamous *fioratura* sends her literally into the heavens, giving her celestial power. At the same time, these notes relegate her to the most feminine reaches of the voice, making her inescapably inferior in the eyes of Western Society. Her femininity, displayed through her soprano coloratura, is most closely linked with her malevolence for the same reason.

A good example of these voice parts working in conjunction is the Act I Quintet "Hm hm hm hm," in which the main plot of the opera is unveiled. It begins with Papageno unable to speak, as he is being punished for lying, with a lock over his mouth. His lie was not malicious, but opportunistic, attempting to take credit for saving

Tamino's life from the serpent he fought at the start of the action. Of course, the Three Ladies who punished him for taking their credit eventually set him free, but not before Tamino attempts to first, showing his continuing kindness. The women spend the rest of the ensemble more or less in subservience to the men, supplying them with tools for their journey.

The ensemble begins with Papageno's introduction of the infectious melodic line. Tamino then enters with a contrapuntal answer to the melody, clearly placing Papageno in the center of the action. When the Three Ladies eventually free him, there is no tangible difference between what one says and what all three say at once. When they sing together it is as one voice; the harmonies are so tight that it is difficult to tell them apart. Thus, the first women we see in the opera are a completely homogeneous trio. Tamino and Papageno usually have their own line apart from one another, though they occasionally sing together when they are on the same page. The point at which there is more of a dialogue between the Three Ladies and Papageno, he is always speaking, while the three of them sing as a single unit in tight harmony.

Tamino attempts to free Papageno from the lock, which here represents the oppression of the extreme vocal ranges. When he fails, the Three Ladies must do it for him. He has not yet been tested, and must first prove himself. The scene is important thematically, as it introduces the magic flute and chimes, the instruments that will protect and guide them throughout the story.

The chorale sections in which all five characters sing as one are perfect examples of the various voice parts and character roles working harmoniously together, an ensemble technique for which Mozart is especially famous. While the various solo lines are important and provide information, when they come together in a homophonic unity, the texture demonstrates the new Enlightenment ideals of political equality.

From this discussion, we may conclude that the baritone roles in Mozart's operas are the most malleable, and hence progressive, of the voice parts. Among the various baritone examples presented, there are several correlations. Each character is somehow focused on the fairer sex. For example, while they differ in station, both Count Almaviva and Papageno are on the prowl, in constant search of physical love. At first glance, Papageno seems to be the more admirable lover, in the noble pursuit of a wife, while

Almaviva is the false husband, seeking pleasure outside the sacred bonds of matrimony. Further examination reveals, however, that neither is after a woman in order to improve society, nor, conversely, to break anyone's heart. They are merely out for themselves, sometimes causing pain, but not maliciously scheming to do so. These two medium-voiced characters are self-centered, an intrinsic aspect of the humanity that they represent. Likewise, Don Giovanni, though undoubtedly a villain, is still primarily concerned with his own needs. In order for him to satiate his lust as a rapist, he must cause pain and exert power over others. Guglielmo is mainly causing mischief with his plans, though he could not accomplish these goals without the façade of an amorous Turk. Therefore, within the context of the soldiers' prank, he is no different than the other baritones in his search for physical love, with no concern for the well-being of others.

More important, though, is what we don't find in Mozart's operas. While the male voice is allowed to expand and develop its dramatic roles, reveling in newfound vocal equality, the female characters are all confined to the soprano range, crowded into a narrow band of sound. The innovations in the baritone voice parts reflect the strides that democratic ideals were making during the time that Mozart was writing. The continued subjugation of women, with their lack of variety in range or dramatic role, stands out in stark relief against such progress. In the Mozart's operas discussed above, if women are not downright interchangeable, their musical and dramatic potential is restricted to their archetype, or in the broader society of his day, to their appearance and function. Even when women might fend for themselves, they are restrained from acting; high voices awaiting the support of the men beneath them. When the mezzo-soprano voice is used, it is primarily as a device to indicate age, whether that of a young man, or an old woman— in either case, not a sexual threat to the social status quo.

Nevertheless, Mozart led the way for subsequent composers to expand their own image of the society beyond soprano and bass. Through his inventive use of the medium voice, as well as new uses for the high and low voices, he was able to accurately reflect a changing time. The lack of female equality, while not a positive thing, is an accurate representation of his world. The period during which he was composing saw the American and French Revolutions, and with them, a burgeoning middle class, in which he played an active part. Beyond his freelance work, Mozart's Freemason connections

show that he was a strong proponent of democratic equality. The freeing of the male medium voice in his operas was a subtle way of flashing enlightened ideals in the faces of the upper class. With all of that in mind, the lack of freedom that he provides for the female voice gives fascinating insight into the limits of even the most forward thinking of Mozart's day. Fortunately, the brilliance of his music knows no such bounds.

Works Cited

Baker, Theodore, and Nicolas Slonimsky. *Biographical Dictionary of Musicians.* New York: G. Schirmer, 1971. Print.

Bokina, John. *Opera and Politics: From Monteverdi to Henze*. New Haven, CT: Yale UP, 1997. Print.

Brett, Philip, Elizabeth Wood, and Gary C. Thomas. *Queering the Pitch: The New Gay and Lesbian Musicology*. New York: Routledge, 1994. Print.

Brown-Montesano, Kristi. *Understanding the Women of Mozart's Operas*. Berkeley: University of California, 2007. Print.

Burrows, Donald. *The Cambridge Companion to Handel*. Cambridge: Cambridge UP, 1997. Print.

Donington, Robert. *Opera and Its Symbols: The Unity of Words, Music, and Staging*. New Haven: Yale UP, 1990. Print.

Ford, Charles. *Così?: Sexual Politics in Mozart's Operas*. Manchester [England: Manchester UP, 1991. Print.

Halliday, W. R. *The Pagan Background of Early Christianity.* New York: Cooper Square, 1970. Print.

Hunter, Mary Kathleen. *The Culture of Opera Buffa in Mozart's Vienna: A Poetics of Entertainment*. Princeton, NJ: Princeton UP, 1999. Print.

Kelly, Thomas Forrest. *First Nights: Five Musical Premieres*. New Haven: Yale UP, 2000. Print.

Smart, Mary Ann. *Siren Songs: Representations of Gender and Sexuality in Opera*. Princeton, NJ: Princeton UP, 2000. Print.

Stafford, William. *The Mozart Myths: A Critical Reassessment*. Stanford, CA: Stanford UP, 1991. Print.

Waldoff, Jessica Pauline. *Recognition in Mozart's Operas*. New York: Oxford UP, 2006. Print.

Pension Reforms in Haiti: Retirement Savings Plan Along With a Credit Plan

Vishal Rajput

FOREWORD

Over the course of a 10-week long internship in Haiti, I had the opportunity to work on pension statements for the employees at Manutech Inc., an electronics manufacturing company. The mandatory private sector pension plan administered by Office Nationale de Assurance (ONA) is inadequate and does not seem to be promising for workers after retirement. Through the company, I had the privilege to meet Tom Adamson who has been in Haiti for the past 30 years and is an expert on pensions. My discussions with Tom on the need to reform pensions in Haiti got me interested in pensions and prompted me to do more research on pension plans in Haiti.

I will discuss the need for pension reforms in Haiti in order to create financial security for workers after retirement. The paper starts with a general introduction to the need for financial security for the elderly. It gives a brief description of the current economic and political situation in Haiti. The two main pension plans, the Civil Service Pension Plan (public sector pension plan) and ONA (private sector plan) are described with supporting schedules. I will discuss the problems with these pension plans and the need for reforms.

I will introduce the Personal Retirement Savings Plan (PRSP) as an alternative to the current plans. Chile switched to a defined contribution plan in the 1980s. Around the same time, three counties in Texas opted out of United States Social Security and created personal retirement accounts for public employees. The paper evaluates these two examples to support the PRSP. Tom Adamson has implemented this plan at his company, SAFICO, on a short scale. The PRSP has a credit plan attached to it which enhances it in the short term as well as in the long term. SAFICO workers under this plan have overwhelmingly accepted it. The paper describes the plan in detail and commends it. The conclusion reiterates the need for and steps toward

pension reforms in Haiti.

Pension Reforms in Haiti

Introduction

Security for the old aged

Declining fertility and the rising life expectancy as a result of advancement in modern medicine has increased the proportion of old people in most societies. By 2050, the proportion of people over 60 to those aged 15–59 would climb from about 30% now to over 60% (The Economist). Financial support for the old aged will be a critical issue in coming years. Old[1] people have traditionally depended on their families for support, but such systems are weakening. On the other hand, government pensions are proving unsustainable and many countries, including developed nations like the United States and nations in Europe, are facing severe problems funding pensions for retirees. The World Bank's *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth*, focuses on the issue of financial security for the old aged. The report identifies three functions of old age financial security systems: redistribution, saving and insurance, and suggests that governments should develop three pillars of old age security—a publicly managed mandatory system, a privately managed mandatory savings system, and a voluntary savings.

Different types of security for the old age people

People have traditionally depended on their families to help them in their old age. More than half of the world's population in 1994 was estimated to reply exclusively on informal and traditional arrangements for income security (The World Bank). However, this has changed over the years due to lifestyle changes throughout the world. Young generations are moving into the cities due to urbanization. Their parents are forced to stay behind at the houses or end up in nursing homes. Over 100 Haitian workers at Manutech Inc. filled out a short survey that asked questions on informal means of financial security (see *Box 1*). Only 16 out of 100 workers said that their grandparents were living with them, if alive. More workers who were 26 or younger said that they do not expect financial help from their children when they retire.

Most old people in developed countries have some type of retirement benefits. There are government implemented pension plans in almost every developed country. There are two main types of pension plans: defined benefit and defined contribution. A defined benefit pension plan

---

[1] "Old" people in this paper are people that are over 50 years or older and have retired from their jobs.

promises a specified monthly benefit on retirement that is predetermined by a formula. The formula is usually based on the employee's earnings history, age, and duration of service. The employers fund defined benefit pensions by saving money over time. In a defined contribution plan, the amount of the employer's annual contribution is set in advance. The payout depends on investment returns. Most of the current pension plans are defined benefit plans. However, recent pension reforms have tried to address the underfunding of defined benefit plans and provide incentives to save via defined contribution plans (Cocco).

Haiti

Economy

Despite optimistic investment and revenue growth in the first quarter of FY 2010, the 7.0 magnitude earthquake set the economy back, and it contracted by 5.3% in FY 2010 (CIA). The International Monetary Fund (IMF) expected growth to rebound strongly in FY 2011, to about 9%, driven mostly by reconstruction and foreign investments (globalEdge). However, Haiti remains the poorest country in the Americas and one of the poorest in the world (with a GDP per capita of US$ 671 in 2010), with significant deficits in basic services (IMF). Over half of its population of 10 million lives on less than US$1 per day, and approximately 80% live on less than US$2 per day, and it also has a high gap of income and unequal wealth distribution, with a Gini coefficient of 0.59 as of 2009 (The World Bank). Some other key economic facts about Haiti can be found in *Appendix A.*

> A study conducted by the CIA reports that Haiti's relative economic stagnation is a result of inappropriate economic policies, political instability, a shortage of good arable land, environmental deterioration, continued reliance on traditional technologies, under-capitalization and lack of public investment in human resources, migration of large portions of the skilled population, a weak national savings rate, and the lack of a functioning judicial system (CIA).

Why pension reforms?

Developing and underdeveloped countries are often helped by developed countries and international non-profit organizations through monetary donations, food aid and other services. While such efforts have helped the poor economies in the short run, they prove unsustainable in the long run. Many times, the aid does not reach the people who need it the most, due to corrupt governments. Sometimes, the aid provided with good intentions result in deterioration of local

businesses. For example, in the 1990s, President Clinton pushed the Haitian government to remove tariffs on imported rice which led to great losses for the Haitian rice farmers. President Clinton admitted, "It was a mistake. ... I had to live every day with the consequences of the lost capacity to produce a rice crop in Haiti to feed those people because of what I did" (Dupuy). Haiti needs to develop financial tools that will prove beneficial in the long run. Pensions will serve as a source of income for workers in Haiti to support themselves when they retire. Small financial measures such as these would develop a sustainable economy and positive economic growth in the long term.

Current Pension Plans in Haiti

Introduction

The pension system in Haiti includes a civil service pension mandatory plan, a private sector mandatory plan, around seven state owned enterprise mandatory plans, and around 35 private sector corporate voluntary plans. They all are pure defined benefit plans with the exception of the private sector corporate plans, which are mainly defined contribution plans. The corporate plans work mainly as tax advantage savings plans that corporations set up on behalf of their employees so that they can access credit against their own and their employer's contributions at better rates than are currently afforded by micro credit and banking institutions (IMF). There is no public information available on the corporate plans. Thus, they are not covered in this paper.

Civil Service Pension Plan (CSPP)[2]

The Civil Service Pension Plan is administered by a Directorate in the Ministry of Finance and it covers civil servants in the central government, doctor and nurses, teachers at all level of public education, police, and judiciary. Civil servants in local authorities have no pension arrangements while the CSPP also administers separate plans for the Parliament and the Armed Forces. Around 50,000 individuals contribute 8% of their base salary into the CSPP. The State stopped matching contributions in the mid-nineties and has not yet resumed 100% matching. Around 11.5 thousand individuals received retirement benefits in 2007, generating a

---

[2] Most of the data presented here can be found in the 2007 *IMF Report:* Financial Sector Assessment Program in Haiti.

system dependency ratio of 4.35. The vast majority of retirement benefits are represented by old age benefits. Nevertheless, the CSPP also offers a survivor pension for spouses (50% of old age benefits) and for children younger than 25 years of age (the other 50%, if alive and entitled). Civil servants also benefit from non-contributory health and life insurance coverage fully self-insured by the State (IMF).

Contribution income amounted to around HGD40 [3] million per month in 2007, while benefits amounted to around HGD37 million per month generating an average monthly surplus of HGD3 million (IMF). Benefits are exclusively paid via checks (an extremely costly and error prone system) that individuals need to collect in person or through delegation at one of the eight departmental offices of the CSPP in the country. Interestingly, beneficiaries have to re-register every fiscal year in person to be entitled to benefits for the following year. This drastically reduces the likelihood that the CSPP pays ghost pensions. Surpluses accumulate in a non-remunerated account in the BRH (Banque de la République d'Haïti). Only recently was the CSPP authorized to manage and invest reserves. As a consequence, in 2007, HGD500 million were transferred into a foreign exchange deposit also in the BRH (IMF). It is expected that the treasury will define the strategic asset allocation for such reserves with BNC (Bank Nationale de Crédit) being the implementing agency.

Members can retire according to the schedule in *Appendix B*. There are some exceptions to this table for nurses, doctors and teachers, for whom no minimum retirement age and lower ceilings apply (Social Security Administration). Old age pension benefits are generous but they are not indexed to inflation or real wage growth rate; this quickly reduces their purchasing power over time. The inflation was 9% in 2011 and is expected to increase throughout 2012 (The Economic Intelligence Unit). Increase in food and commodity prices can be related to an increasing foreign population (U.N. peacekeepers, non-profit organizations) in major cities like Port-au-Prince and Jacmel. Effects of pensions without indexation have been noticed in other countries too. In Venezuela, real pensions fell 60 percent during the 1980s largely because of inflation without indexation (World Bank). In addition, there is no mortality table for civil servants and no actuarial evaluation of the plan has been conducted.

---

[3] The exchange rate on 5/1/2012 was 1USD = HGD42.35. The exchange rate fluctuated widely in 2011 (See *Appendix J*).

The Directorate of the CSPP could, in theory, create very accurate benefit, age and gender distributions of retirees, but it lacks the information technology capacity to do so. The database underpinning the payroll system should be assessed to establish its ability to produce the necessary data on contributors (age, gender and wage distributions over time), in order to construct a mortality table and to conduct actuarial evaluations. No financial statements are published by the Ministry of Finance on the actuarial or financial viability of the plan. In summary, the Directorate administering the CSPP is merely used as a benefit payment agency, while it could be effectively and cheaply transformed into a civil service pension analysis office to advise the government on public pension, staffing and remuneration policies.

Private sector social security scheme (ONA)[4]

The Office Nationale d'Assurance-Vielliesse (ONA) was established in 1965 as the reincarnation of the old Institut d'Assurances Sociales d'Haiti (IDASH) established in 1949 (SSA). Its regulatory framework comprises the establishing law of 1965 amended in 1967 and several complementary laws that have progressively transformed the mandate of the institution from a pure retirement income plan to a housing and savings plan. These include: in 1968, the enactment of implementing provisions; in 1971, the introduction of personal loans; in 1974, the introduction of mortgage loans; in 1976, the introduction of disabled benefits provisions and the relaxation of mortgage loan provisions; and in 1983, the reform of its governance with the establishment of a joint board ("Conseil d'Administration des Organisms de Securite Sociale" or CAOSS) for the ONA, OFATMA and two other noncontributory social assistance programs (SSA).

The ONA is a technical and administrative division of the Ministry of Labor and Social Affairs (MAST) with its own jurisdiction and financial autonomy. The ONA covers all workers in the private and public sectors not covered by the CSPP. Salaried individuals are required to participate while self-employed and workers in the informal sector can participate voluntarily. The coverage rate is relatively small: in 2007, only 234,000 individuals were active contributors representing 2.8% of the population, 8.6% of the labor force, 12.1% of the formal sector labor force, and 12.4% of the formal sector labor force not covered by the CSPP (The World Bank).

---

[4] The numbers presented here can be found on U.S. Social Security Administration website under Caribbean nations. In addition, many of the statements made here are from my direct experience dealing with ONA during my internship.

ONA reported that 12,000 beneficiaries currently receive a pension from the ONA producing a highly favorable system dependency ratio of 119.

However, this high ratio does not correctly reflect the unhealthy financial situation of the ONA, given the erratic collection capacity of the institution and its high administrative expenses. Operational expenses amounted to HGD380 million—around 36% of contribution income (see *Appendix C*). This includes HGD55 million (1/5 of total expenses) in various non-identifiable sponsorships. Old age pension benefits are not as generous as old age benefits in the CSPP as they replace a maximum of 30% of average best ten years salary. Members can retire according to the schedule in *Appendix 3*. The formula is that ONA receives 12% of payroll (half from the worker and half from the employer) when they have contributed for 20 years and attained the ripe old age of 55 (current life expectancy in Haiti is about 50–60) they can draw a pension which is the equivalent of 1/3 of the average of the last 10 years. There is no provision for cost of living adjustments; either in calculating the base average or in the pension benefits itself, in spite of 20 years of double digit inflation.

The ONA also has a lending program. Mortgage loans are by far the largest share of loans granted at an average rate of 14% to 17% (ONA Office). Such high interest rates result in high default risk. Given the low earnings of an average Haitian worker, interest payments are not made on time. This results in workers losing their collateral, the remaining and future pension savings with ONA. Personal loans are also granted up to 50% of accumulated individual account balances at a rate of 6%. Salary advance loans up to 4 to 5 month salary are another category and finally, advances on pension benefits are also granted. The maturity of these loans does not exceed 18 months with the exception of mortgage loans, the maturity of which can be up to 20 years. A minimum of 5 years of contributions is necessary for members to qualify for a loan. Additional credit scoring is conducted by the ONA staff but the scoring procedure appears to have been rather ineffective in the past, at least as far as mortgage loans are concerned. The ONA admits to lacking investment and credit analysis skills to conduct mortgage loan operations and suspended all new mortgage loans in 2007.

The dire reserve management situation is the consequence of a developmental general mandate for the ONA, lack of investment skills, inadequate governance structure and possibly undue political interference. Since inception the ONA was supposed to have a tripartite board of directors, but this was never appointed. The 1983 Law establishes the CAOSS and in 2005 the

transition government requested nominations from the employer's association and unions but the President of the Republic has yet to appoint the nominees (The World Bank). Tom Adamson was nominated to be on the board two times since 2008. It is believed that the appointment of a board of directors, even if only tripartite, would improve transparency in the management of the ONA.

Recommendation—Personal Retirement Savings Plan

Example—Chilean Reformed Pension System

In the late 1980s, Chile faced a huge problem with their pay-as-you-go defined benefit pension plan. The increasing number of retirees, high evasion rates, and immense benefits promised by earlier political leaders led to unbearable stress on the social security system. The pension system was paying more benefits than it was receiving in contributions, and the projected actuarial balance was greater than the country's GDP (Soto). Chile overhauled their pay-as-you-go pension system into personal retirement accounts. The retirement age was raised to 65 for men and 60 for women. Chilean workers contribute about 12.5 percent of their wages to the old aged, disability and survivorship. Nearly 80 percent of the employee's contributions go towards personal accounts which are administered by privately owned pension management institutions; the remainder pays for disability and survivorship insurance (7 percent), administrative fees, and commissions (13 percent) (Soto).

The introduction of a privately managed pension system in Chile, based on individual capital accounts, has attracted world-wide attention. "This reform—as well as other market oriented structural changes—and the significant improvement in Chilean economic performance has led many observers to conclude a direct link, especially through the rise in private domestic savings generated through the new pension system" (OECD).

Example—Texas Model for Fixing Social Security[5]

In September 2011, Texas Governor Rick Perry pointed to three counties in Texas, Galveston, Matagorda, and Brazoria, as a solution to fix the United States Social Security problems. These three counties opted out of social security decades ago and created personal retirement accounts for public employees. County workers in these three counties have reaped higher benefits under this plan. These counties—unlike almost all others in the United States— face no long-term unfunded pension liabilities (Merrill).

---

[5] This can be found in a Wall Street Journal article "Perry Is Right: There Is a Texas Model for Fixing Social Security"

The Alternate Plan is different from a traditional defined benefit or defined contribution plan. Employees contribute 6.2% of their income with a 100% matching by the counties (Merrill). After the contribution is made, the county is not financially obligated to manage the funds, thus avoiding long term unfunded pension liabilities. The contributions are managed by financial institutions that bid on the money and guarantee a minimum interest rate. First Financial Benefits, Inc. currently manages the contributions. During the last ten years, the accounts have earned between 3.75% and 5.75% every year, with the average around 5% (Merrill). When the market is positive, the contributions get higher returns, and when the market goes down, contributions still make something.

The Alternate Plan is potentially a good substitute for social security. However, if everyone was to opt out of Social Security today, the shortfall of money for the current pension receivers would be immense and create more problems. A reform plan, such as the Alternate Plan would need to address the transition costs.

Personal Retirement Savings Plan coupled with Credit Plan[6]

Why Implement Personal Retirement Savings Plan in Haiti

As mentioned in a previous section, ONA does not provide enough benefit for a decent retirement. The pension one receives from ONA represents only about 22% of the last salary and there is no provision for increases in the annual cost of living. In most countries a comfortable pension represents at least 50–60% of the last salary with regular cost of living increases. The government or companies should set up a system so that their loyal employees do not end up living in poverty in their old age after working for many years. It can be very complicated to put together a company pension plan similar to ONA (defined benefit); such a plan has many drawbacks. Many companies in the United States have made news for going bankrupt because of their obligations to their retired workers through open-ended pension plans. Retirement Savings Plans are becoming widely accepted with no open-ended obligations. Under such plans, employees are able to build up capital, with the employer's help, that will help finance their retirement. This plan allows employees to know what their savings are at any moment. In addition, most banks in Haiti offer special accounts for Retirement Savings Plans.

---

[6] This plan has already been implemented at Tom Adamson's company, SAFICO Haiti, on a small scale. This paper evaluates this plan and uses data from SAFICO to support arguments.

Importance of Credit Plan attached to the Retirement Savings Plan

If a retirement savings plan (RSP) already exists, or is going to be set up, it costs very little to add a credit plan. The employer will be providing a big additional benefit for their employees at relatively little cost. Access to credit is the short term benefit that guarantees that the employees willingly participate in the retirement savings plan. RSPs with a credit plan have a greater annual rate of return compared to depositing them in a savings account at a bank. If the objective of the RSP is to enable employees to accumulate the equivalent of 6 years' salary, then the credit plan allows the employee to reach this objective much more easily (see *Appendix G*). With access to a line of credit the employee can more easily face up to many recurring expenses: school fees, rent, health problems, funerals, etc. (see *Appendix H*). The line of credit gets bigger each year and can eventually enable them to solve their major problems and might even help them invest in a business to create another source of income.

How does interest calculation work?

At the end of each month, all of the money in the RSP is either out in loans to the workers or in the account at a bank. During the month, an account will receive contributions and loan payments and give out new loans. At the end of the month, the company will calculate how much principal is out in loans and how much is in each individual RSP account. Money sent to the bank cannot be withdrawn until the employee retires or leaves the company. Thus, if 30% of the total funds are out in loans, each employee's bank account will be 70% of what he or she has in the RSP regardless of how much he or she owes on the loan. At the end of the month, their account balances will be determined and deposits will be made to bring them up to this level. The employees do not have direct access to the accounts, except through the company, so they won't borrow all their savings, quit the company, then go and clean out their account. The interest earned, at the bank and on the loans, is calculated on the last day of the month and distributed amongst the workers according to their percentage of the total at the end of the previous month.

How do the loans work—SAFICO example

Once a worker has accumulated a certain amount (say $50) in her account she is allowed to borrow up to 100% of her account balance. At SAFICO, Tom Adamson sits down with his employees to figure out how much they are allowed to borrow, how much they want to borrow and in how many months they want to pay back the loan. Tom does not allow repayment to go

beyond 12 months. However, he is considering allowing longer term loans for home construction or investment in a business since he has seen a very strong repayment rate. SAFICO charges 2% a month on their loans and this seems very reasonable for workers because they don't borrow too much and they must pay back within a year. However, the annual return on the loan is quite favorable due to this scheme (see *Appendix I*). The setup for such a system is quick and it can be easily managed. At the end of the month, all employees receive a pension statement, which informs them of their total contribution to date, the amount of interest they have accumulated, and the amount of outstanding loans, if any.

Benefits of Credit Plan attached to the Retirement Savings Plan

In the short term, the workers have access to a line of credit that gets bigger every year. Well managed, this enables them to deal more easily with many expenses, both foreseeable and unforeseeable. At SAFICO most workers now have between five and six months' savings in the RSP with a 6% employee contribution and 100% employer matching. If this continues, according to Tom Adamson, in ten years the workers should have the equivalent of two years' salary. In the long term, since the workers are paying commercial interest rates of 2% per month on their loans, the annual return will be raised on the fund as a whole. Since 2004, SAFICO has started lending anywhere from 25% to 45% of the total out in loans at any one time. This has meant an annual rate of growth of between 11 and 12% (see *Appendix J*).

In the short term, companies can avoid lending their own capital to the workers once the credit plan has reached a certain size. If a company already has a RSP, it costs very little to establish a credit plan which is perceived as an important benefit by the workers. In the long term, companies can benefit by preventing older employees from becoming a burden by receiving unfunded pensions out of cash flow or by staying on payroll beyond their productive years.

There has been a wave of reform in Latin America moving away from Defined Benefit towards Defined Contribution Plans (with individual accounts). However, none of these plans lend money to subscribers; most of the money is parked in government bonds or other local financial instruments. The criticisms of these new plans are that they are still plagued by a fair amount of delinquency and that they cannot attract subscribers from the informal sector. This leaves many people in the society with inadequate or no coverage when they reach retirement age. Both of these problems might stem from the fact that the retirement benefits are so far into

the future that without effective compulsion, many people are hesitant to subscribe. The credit plan offers a tangible immediate benefit which could reduce this problem. Tom Adamson recently reported that workers at SAFICO unanimously asked to increase their contributions to the plan twice because they know that the money is available for them when they need it.

Conclusions

Underdeveloped countries are often helped financially through donations and funding. Such help might solve problems in the short term, but will not help the country develop in the long term. Microcredit gives people access to money but does not incorporate saving mechanisms in poorer society. For underdeveloped nations to be self-sustainable, improvements in financial institutions and awareness of savings among people are necessary. Poor countries like Haiti can start reforming pension plans for long term growth and better living conditions. The poorest groups in such societies can still get social assistance, such as cash, through social pensions[7].

The current pension plans in Haiti need to be reassessed for higher efficiency. An actuarial evaluation for these plans should be conducted to better assess the contributions under these plans. The amount received after retirement should be indexed for factors such as inflation and wage growth. The IT technology at the Civil Service Pension Plan office and ONA needs to be developed to reduce administrative expenses and increase accurate bookkeeping. Implementing personal retirement savings plans, similar to Tom Adamson's SAFICO, through companies might be a good substitute for contributions to ONA. The pension system in Haiti is still in its early stages with a favorable dependency ratio. Switching to personal accounts would incur relatively low transition costs. Personal Retirement Savings Plan could develop the weak financial institutions in Haiti for competitive bidding by institutions.

---

[7] Social pensions are non-contributory cash payments to older people. They are helpful in situations where older people have not contributed to a formal pension plan.

Box 1: Demographic Survey

Data Collection: Workers at Manutech Inc. in Port-au-Price were asked to fill out a survey that asked questions about dependence on families for financial support. The survey asked for information such as gender, age, the number of people in the family and their relationships with each other, who provided financial support for the family, and if they expect their children or family to help them in when they retire or are of old age. Over four hundred surveys were filled out but only one hundred completed surveys were analyzed. 74 females and 26 males filled out the survey. Manutech workers consist of proportionally more females than males.

Findings: The age of the workers ranged from 20 to 61. The average age of the group was 33 years. Only 16 out of 100 workers said that their grandparents live with them (if alive). Twenty of them said that they live with both their parents and their children. 68% of the workers said that they were the main financial source of the family. 47% of them said that they do not expect their family or children to financially help them when they retire. The average age of the workers who do not expect financial help after retirement was 26 compared to 34 for the ones who expect help. 49% of females who filled out the survey do not expect financial help when they retire compared to 42% of males.

Conclusion: The traditional informal financial means of support for the old aged is weakening in Haiti. The survey shows that the younger generation is less likely to depend on their children or family for financial support during their old age. Thus, other formal means of support such as a strong pension plan needs to be present to support workers when they retire.

Limitations: One of the officers translated the survey from English to Haitian Creole for the workers at the company. The translation might not have been fully accurate. Some workers might have filled out wrong information such as age or family size. The size of the sample analyzed was limited to one hundred.

Works Cited

"Books and Arts: Matters of Life and Death; the History and Future of Pensions." *The Economist,* 31 Aug 2002.

Central America and Caribbean: Haiti. Central Intelligence Agency. *CIA*. Web. 15 Apr. 2012. <https://www.cia.gov/library/publications/the-world-factbook/geos/ha.html>.

Cocco, Joao F., and Paula Lopes. "Defined Benefit Or Defined Contribution? A Study of Pension Choices." *Journal of Risk and Insurance* 78.4 (2011): 931–60. ABI/INFORM Complete; ProQuest Health Management. 30 Mar. 2012.

Dupuy, Alex. "Disaster Capitalism to the Rescue: The International Community and Haiti After the Earthquake." *NACLA Report on the Americas* 43.4 (2010): 14–9. ProQuest Research Library; ProQuest Social Science Journals. 1 April 2012.

Friedberg, Leora, and Michael T. Owyang. "Not Your Father's Pension Plan: The Rise of 401(k) and Other Defined Contribution Plans." *Review—Federal Reserve Bank of St. Louis* 84.1 (2002): 23–34. ABI/INFORM Complete; ProQuest Research Library. 29 Mar. 2012.

Haiti Economy: Inflation on the Rise. New York, United States, New York: 2011. ABI/INFORM Complete. 6 May 2012.

Haiti: Country Brief. The World Bank. Web. 22 Feb. 2012.<http://web.worldbank.org/WBSITE/EXTERNAL/COUNTRIES/LACEXT/HAITI EXTN/0,,contentMDK:22251393~pagePK:1497618~piPK:217854~theSitePK:338165,00 .html>.

"Haiti State Pension Agency Staff Accused of Money Laundering." *BBC Monitoring Americas* (2008). Web.

"Haiti: Resources." *GlobalEDGE*. Michigan State University. Web. 12 Apr. 2012. <http://globaledge.msu.edu/Countries/Haiti/resources>.

"International Monetary Fund—Financial Sector Assessment Program. Tech. *World Bank*, 2007. Print.

Matthews, Merrill. "Perry Is Right: There Is a Texas Model for Fixing Social Security." *The Wall Street Journal* 24 Sept. 2011. Print.

Rosenberg, Richard. 2010. "Does Microcredit Really Help Poor People?" Focus Note 59.Washington, D.C.: CGAP

"Social Protection in Older Age." *Why Social Pensions?* Help Age. Web. 30 Mar. 2012. <http://www.pension-watch.net/pensions/about-social-pensions/about-social-pensions/why-social-pensions/>.

"Social Security Programs Throughout the World: The Americas, 2009—Haiti." *Ssa.gov*.
        Web.15 Apr. 2012. <http://www.ssa.gov/policy/docs/progdesc/ssptw/2008-
        2009/americas/haiti.html>.

Soto, Mauricio. *Chilean Pension Reform: The Good, the Bad, and the In Between*. Issue brief.
        Boston College: Center for Retirement Research, 2005. Print.

The World Bank. *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth.*
   Oxford UP, 1994. Print.

Appendix

Appendix A: **Key Economic Facts**

Haiti

Key Economic Facts

Budget Expenditures:	$.00 trillion (2011)
Budget revenues:	$.00 trillion (2009)
Debt - external:	$.00 trillion (2011)
Distribution Of Family Income - GINI Index:	59.20% (2001)
Exports:	$.69 billion (2011)
GDP per capita, PPP:	$1,110.70 (2010)
GDP, PPP:	$.01 trillion (2010)
Imports:	$3.28 billion (2011)
Industrial production growth rate:	-4.80% (2010)
Inflation rate (consumer prices):	9.30% (2011)
Labor force:	4.81 million (2010)
Population:	9.80 million (2012)
Unemployment rate:	40.60% (2010)

GDP Composition %

Agriculture	Industry	Manufacturing	Services
28	17	8	55
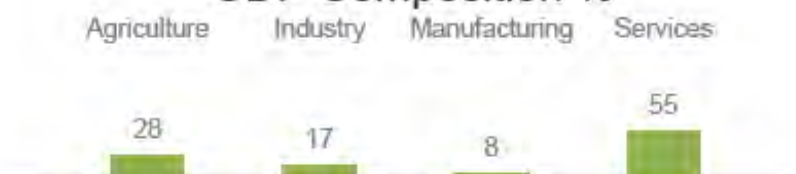
Rankings

Index	Rank
Corruption Perceptions Index	146/178
Ease of Doing Business Rankings	154/181
Freedom of the Press	100/196
Index of Economic Freedom	141/179
International Logistics Performance Index (LPI)	98/155
Inward FDI Potential Index	140/141
KOF Index of Globalization	172/208

Source: globaledge.msu.edu   *Numbers rounded to the nearest trillion dollars*

Appendix B: **CSPP – Benefit formula**

Age	Years of contributions	Replacement Rate	Monthly Ceiling
	Less than 25	**Restitution of Contributions**	
55	**25 – 29**	**60% of best salary**	**HGD15 Thousand**
55	**30 – 39**	**60% of best salary**	**HGD25 Thousand**
55	**40**	**100% of best salary**	**HGD30 Thousand**

Source: The World Bank

Appendix C: **O.N.A. Income Statement – June 30, 2007**

(Exprimé en gourdes haïtiennes)

REVENUS
Intérêts Gagnés	26,238,652.85
Escomptes gagnés	6,211,989.73
Revenu de Placement	15,219,090.64
Autres revenus	4,049,599.41

TOTAL REVENUS   **51,719,332.63**

CHARGES

Salaires & charges sociales	244,664,907.98
V Honoraires professionnels	5,992,023.36
Autres dépenses du personnel	6,901,016.20
Publicité	17,563,200.98
Frais de déplacement	2,736,438.74
Amortiss. Des immobilisations	11,370,857.66
Eau, Electricité, Téléphone	1,069,619.26
Entretien & réparation	5,926,622.02
Carburants & lubrifiants	5,038,577.50
Frais financiers	12,835.00
Subvention	54,734,704.00
Dons & Assistance sociale.	4,906,380.50
Fourniture utilisée & impres.	9,003,200.25
Loyer	407,500.00
Assurance Immeuble	517,054.66
Assurance Véhicule	110,310.59
Autres dépenses	8,460,225.39
TOTAL CHARGES	**379,415,474.09**

DEFICIT OU EXCEDENT   **(327,696,141.46)**

Source: The World Bank

Appendix D: **ONA. – Contribution schedule**

Monthly Salary Bracket	Employer	Employee
0 – 200	2%	2%
200 – 500	3%	3%
500 – 1000	4%	4%
1000 – more	6%	6%

Source: ssa.gov

Appendix E. **ONA. – Pension benefit formula**

Age	Years of contributions	Replacement Rate	Monthly Ceiling
	Less than 10	Restitution of contributions with immediate vesting of employers' contributions	
55	10 – 20	1.6% accrual rate x avg. salary of best 10 years x no. of yrs. of contribution	HGD15 Thousand
55	20 and over	Same	1/3 average salary of best 10 years

Source: The World Bank

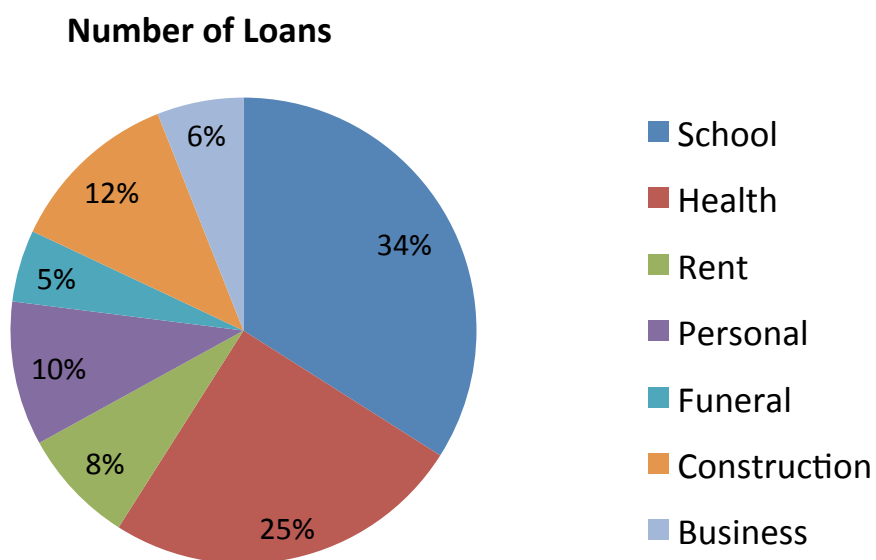*Appendix F***: TELECO. – Pension benefit formula**

Age	Min Years of contributions	Formula	Ceiling
	<15	Restitution of contributions with immediate vesting	
45	15	Choice of deferred pension at age 55 (see formula below)	70% of highest salary
55	15	2% accrual rate x arithmetic avg. best 5 year salary x no. of years in service	70% of highest salary

Source: The World Bank

Appendix G: **Interest Rate Comparison**

FV	6,000
PV	-
PMT	
Years	25
Rate 1	2.5%
PMT 1	176
Rate 2	9%
PMT 2	71

Bank Deposits

Credit Plan

Appendix H: **Analysis of Loan Motives** – January 05 – Dec 07

Number of Loans

- School
- Health
- Rent
- Personal
- Funeral
- Construction
- Business

34%
25%
8%
10%
5%
12%
6%

Source: SAFICO, Haiti

Appendix I:

Debt Principal	100.00
Interest Rate	0.01
Number of Periods	24.00
FV	-
Payment	4.71

	Payment	Interest	Principal	Remaining Principal
1st Payment	4.71	1.00	3.71	96.29
2nd Payment	4.71	0.96	3.74	92.55
3rd Payment	4.71	0.93	3.78	88.77
4th Payment	4.71	0.89	3.82	84.95
5th Payment	4.71	0.85	3.86	81.09
6th Payment	4.71	0.81	3.90	77.19
7th Payment	4.71	0.77	3.94	73.26
8th Payment	4.71	0.73	3.97	69.28
9th Payment	4.71	0.69	4.01	65.27
10th Payment	4.71	0.65	4.05	61.21
11th Payment	4.71	0.61	4.10	57.12
12th Payment	4.71	0.57	4.14	52.98
13th Payment	4.71	0.53	4.18	48.80
14th Payment	4.71	0.49	4.22	44.58
15th Payment	4.71	0.45	4.26	40.32
16th Payment	4.71	0.40	4.30	36.02
17th Payment	4.71	0.36	4.35	31.67
18th Payment	4.71	0.32	4.39	27.28
19th Payment	4.71	0.27	4.43	22.85
20th Payment	4.71	0.23	4.48	18.37
21st Payment	4.71	0.18	4.52	13.84
22nd Payment	4.71	0.14	4.57	9.28
23rd Payment	4.71	0.09	4.61	4.66
24th Payment	4.71	0.05	4.66	0.00
	112.98			

Annual rate of reurn	13%

Appendix J: **USD/HTG Exchange Rate**

Source: Yahoo Finance

Poohdism

Jeff Handy

In Benjamin Hoff's book *The Tao of Pooh*, he uses Winnie the Pooh as an allegorical tool to illustrate the basic principles of the enigmatic Eastern belief system called Taoism. A central Taoist belief holds that if you follow the path, or Tao, that resonates with your inner nature, you will be content with your life. Good living for a Taoist, then, can be sought by what has now become somewhat of a modern cliché—"going with the flow." In the Winnie the Pooh stories, Pooh rarely seems to think twice about his actions; he simply goes about doing whatever it is he has the propensity to do. Because of this fluid behavior, Hoff feels justified in advertising Pooh as an ideal Taoist, but I am inclined to question whether his metaphor holds when comparing Pooh to a human being. Perhaps Pooh can't help but act according to his inner nature. Ignoring the handful of personified traits he possesses, Pooh maintains the same "shallow" free will that a bear has in real life. The possession of a "deep" or "advanced" free will, which Carlo Filice describes in his book *The Purpose of Life*, makes "going with the flow" a much trickier undertaking for human beings. The possession of a deep free will allows a human to reconsider his default nature and make changes if he so wishes. The Buddhists make use of this distinctly human ability when they put forward their Four Noble Truths, which serve as guidelines for rising above the wheel of suffering they call Samsara. The Buddhists claim that if we wish to be blissful and to live the best life, we need to alter our inner nature, specifically our natural inclination to seek gratification, to desire. The Buddhists advocate the use of a deep free will to reprogram, or self-shape our inner nature, whereas the Taoists, at least as Hoff characterizes them, seek only to recognize our inner nature and to live according to it.

I think it's easy to form a few general misunderstandings about Buddhism and Taoism. On the surface, these problems seem to beg us to say that either one or both of these philosophies fail to adequately capture the human existence. Of course one could hold that opinion, but I think that if you look at the alleged shortcomings of each philosophy, you'll see that the second one provides a remedy to patch up whatever problems there are in the first. Put more clearly,

Buddhism does a lot to reveal and remedy some of the problems of Taoism, and vice versa. I think that both philosophies as stand-alone entities are actually respectable in the sense that they get a lot of what it is to exist as a human right. So, I'm not saying that the only way for Buddhism and Taoism to save face is to form a new hybrid, to join forces and create one super-philosophy. I'm saying simply that both would do well to reflect on bits and pieces from the other if they wish to clarify some of their ambiguities and promote a philosophy that accurately reflects the human condition. With that said, when considering the title of my paper—Poohdism—think of it less as the name of a new philosophy and more as a methodology for arriving at a clearer vision of what Taoism and Buddhism are.

So, let us make use of this method: Benjamin Hoff uses Pooh as an emblematic Taoist living a good life. Pooh's life is simple. The quotation on the back cover of *The Tao of Pooh* sums up Pooh's tranquil manner. It reads as follows: "While Eeyore frets…and Piglet hesitates…and Rabbit calculates…and Owl pontificates…Pooh just is." Phrases like "just is" and "going with the flow" are some of the many ways people try to wrap their heads around the outwardly mysterious, inexplicable principles of Taoism. In fact, Taoism is quite simple, like Pooh himself. This is why Hoff chooses to employ Pooh as his representative Taoist. Taoism claims that the universe works harmoniously in accordance with its own ways. To keep this harmony intact, one must align one's will with the way, or Tao, of the universe. Each being has an inner nature that, upon its birth, is harmonized with the universe, and that harmony can be upset only by aggressive acts of one's will. This avoidance of aggressive action is captured in the central Taoist concept of *Wu Wei*. *Wu Wei*, as Hoff describes it, is a "principle of minimal effort," where things happen naturally, without forceful human exertion, simply because "the natural world doesn't make mistakes" and so does not need the aid of aggressive human action. Pooh typifies the Taoist concept of *Wu Wei* because, with what seems like hardly any effort whatsoever, he manages to simply get things done. Essentially, whatever needs to be done will be done naturally (that is, it will be done in accordance with the natural way). Hoff quite capably describes Pooh as a great Taoist master, but in my opinion, the demonstrations he uses with Pooh fail to translate to the human experience. Since the general purpose of Hoff's book is to provide an introduction by way of allegory to Taoism, which can briefly be described as a philosophy aimed at encapsulating the right way for a *human* to live, it must be considered nothing less than

a glaring weakness to fail to have the main symbolic figure, Pooh, accurately represent how a human ought to live.

The allegorical failure in *The Tao of Pooh* stems from the differences in human free will from Pooh's free will. Pooh possesses what Carlo Filice would call a "shallow" free will. All beings are born with an initial set of wants, urges, and inclinations. Beings with a shallow free will are free to go about pursuing those wants and urges, but they cannot change what those wants and urges are. A deep or advanced free will, on the other hand, can change the set of wants and urges that it was born with initially. Humans are possessors of deep free will, but Pooh has only a shallow free will. Even with the array of personified characteristics that Pooh is given, he still can't help but to act in accordance with the nature of a bear—his love of honey being the most obvious example. Humans, on the other hand, often partake in the reprogramming, or self-shaping, of desires after reflecting on values and goals and sometimes deciding, freely, to change their natural urges so that they become conducive to the pursuit of those new goals and values. An example of this, which conveniently contrasts with Pooh's inability to stop desiring a certain food (honey), might be a human who researches the health and/or moral reasons for not eating meat and is then convinced that he will be better off if he chooses to not eat meat. In doing so, the person has replaced the omnivorous diet that was present at his birth with a vegetarian diet that reflects his new values. The human examples are ubiquitous. Indeed, most truly meaningful experiences in our life affect what we value and what we strive for—perhaps such an effect is even necessary in some conceptions of the term "meaningful experience"—and we change our desires accordingly. *The Tao of Pooh* does not seem to consider the significant differences between Pooh's free will and human free wills, which, in my opinion, makes its attempt at allegorically capturing the human experience fall short.

Imagining how Hoff could have dealt with the disparity between the free wills of humans and Pooh is difficult. Deep free will doesn't seem to resonate with the Taoist ideal of *Wu Wei*, where a good life is lived by acting according to one's inner nature, because the inner nature of a human is not fixed; it is constantly in flux due to our ability to change our set of wants. Would this not be a fundamental problem of Taoism if it truly struggles to apply to human agents? It is a *human* philosophy, after all. The Buddhists, however, are a group that do acknowledge our deep sense of free will and use it to promote a strategy to escape from the torturous confines of Samsara. Maybe they can come to the rescue.

The Buddhist Four Noble Truths, to be brief, describe that life is suffering, suffering is caused by our desires, removing desire ends suffering, and following the Eightfold Path is the strategy to remove desire and thus end suffering. Never mind the specifics; what's important here is that implicit in the Four Noble Truths is the fact that we have the ability to change our desires if we wish. Put simply: we have deep free wills. This conscious removal of one of our human traits (i.e. our capacity to desire) seems completely at odds with the Taoist idea of acting in accordance with one's inner nature. Buddhists directly and deliberately tamper with their inner nature to remove suffering. What's interesting is that the Taoists and Buddhists are both seeking the same thing—to live a "good life," a life with minimal suffering—but the ways in which they go about reaching that goal are nearly opposite. Perhaps a tactful convergence of their principles can offer a more practical and reasonable way of minimizing our suffering and living a good life.

I want to try to point out more explicitly some areas where both the Buddhist and Taoist philosophies appear to come up short of capturing the human condition. Fortunately, as we'll see, the weaknesses of the one are mended by the strengths of the other. The first mistake is the more obvious one that the discussion thus far has led up to—Taoists don't acknowledge the idea that perhaps a human's nature *is* to change his inner nature—that self-malleability is essential to the notion of deep free will. In my own less technical language, the Taoists, when characterized this way, remind me of a group of children doing arts and crafts. When someone steps in to offer help, to say "Hey, do you think it would be better if you did such and such instead?" or "Do you need any more popsicle sticks? What about more uncooked macaroni? Don't you think you're going a little heavy on the Glitter Glue?" The children will protectively hover over their projects, alleging, "No, no, everything is exactly how it's supposed to be! Don't mess with anything!" Yes, from Hoff's depiction the Taoists seem to be a lot like that. They treat the universe like it's a brittle house of cards, or a child's Popsicle stick creation, arranged in the best and only way. But what if, within the grand scheme of the universe, the changeability of human inner natures was a *part* of "how it's supposed to be?" What if the cards that built this big house are a little more robust than the Taoists thought? Maybe they only sway instead of collapse under the weight of human volition. I think the Taoists, who really aren't as cranky as the kids I compared them to for effect would accept these ideas, finding that they actually are not contradictory to their concept of *Wu Wei.*

Each change in a human's nature, specifically each change in a human's desires, occurs precisely because it feels like the right thing to do for that person. Consider this—would it not be going against my intuition, my inner nature, to stifle my desire to become a vegetarian once I had been convinced of the moral and health-related benefits of the change in diet? To resist a change like this in my inner nature would be to take the very aggressive action that the Taoists so openly condemn. "Going with the flow" for a human is not like floating down the lazy river at your local water park; it is more like riding turbulent ocean waves. Our goals and values change almost daily and to accommodate that variability we simply cannot ignore those changes. We have to harmonize our actions with our values, we have to lean and paddle in a certain way depending on how large and from what angle the waves are crashing around us. To do otherwise is to deceive yourself into thinking you can sail on calmly in rough waters—you are bound to capsize! Thus, I think the Taoists must acknowledge humanity's possession of a deep free will, lest they advocate a philosophy that encourages us to thwart our uniquely human ability to alter our inner natures. Indeed, for a philosophy to hold a central tenet like *Wu Wei,* which asks us to act in accordance with our inner nature, only to admonish the foremost characteristic of our nature, just seems nonsensical to me. But, since it would not be contradictory for *Wu Wei* to incorporate a changeable human nature, maybe Taoism *does* acknowledge our deep free will. The problem might not be with Taoism itself. Maybe the problem is simply in Hoff's failed allegory with Pooh.

The Buddhist misunderstanding, to switch gears here, rests in their reasoning for wanting to remove desires. Buddhists profess that, from the time of our birth, our nature is to follow our desires, and it is in the following of those desires that we ultimately end in suffering. I want to make clear that the suffering does not come about from the act of desiring itself, but by the consumption and subsequent disappearance of pleasure once a desire has been satisfied. If the pleasure felt at the moment of reaching a desire were eternal, perhaps desiring wouldn't be bad at all. It is the *impermanence* of pleasure, the exposure to pleasure followed by a fall back to normalcy that creates suffering. Buddhists may very well agree with this clarification, but nevertheless I don't think they follow it to its conclusion in their philosophy. I believe that if the problem of suffering is not found in the act of desiring itself but in the impermanence of the pleasure, then perhaps not all desires should be detested. If there exist certain pleasures that last a

lifetime, or even eternally, I think the Buddhists would be okay with the pursuit of, and the desire for those pleasures—for in what else would Nirvana consist?

I have begged the question of whether or not any eternal pleasures exist. I could say that these more noble pleasures might consist in the pursuit of knowledge, or in the act of contemplation, but I think an eternal pleasure might be even closer to home. The act of forming values and shaping your desires to match those values is a human capacity that can never be taken away (barring any form of mental disability). I think this ability is something to be desired and something that we should fully embrace. Furthermore, as I've shown, this ability seems to be consistent with the Taoist concept of *Wu Wei*, since it is innate in every human to self-shape, to remold the predetermined cast in which we were born. If it is in our nature to *change* our nature, I see no discrepancy with the Taoists. *Wu Wei* asks us simply to live according to our nature; it says nothing about what the nature of our nature must be. So, Taoists, you can take a sigh of relief. And, so long as they agree with my previous characterization of the problem of desiring, the Buddhists too can wipe their brow of nervous sweat, as self-shaping is a desire that leads in the long run, only to a better state of being. In other words, we self-shape to avoid suffering, which seems to be the main objective of both Buddhism and Taoism. So, I conclude, Taoism and Buddhism, although it takes a little wrestling with their respective ideas, certainly capture the human condition. Pooh might always find his head stuck at the bottom of the honey jar, victim to his inborn desires, but a human, in accordance with both Taoist and Buddhist principles, can stick his head in whatever jars he chooses.

Redrawing Borders in Africa: Norms and Exceptions

Andrew Kuntz

ABSTRACT

Stability is a central component in the operation of the international system. Reflecting its necessity, the principles of territorial integrity and state sovereignty are deeply engrained in international law. Secession challenges these tenets through the right to self-determination. However, the right to self-determination is not a generalized right to secede, and secession remains a rare occurrence. This has held uniquely true to Africa, where colonialism left a legacy of non-viable states and ripe conditions for secession, yet South Sudan and Eritrea have been the only cases in the half of a century since decolonization. Existing research on secession has primarily sought to understand where it occurs and why. A deficit in the literature of secession exists as to what differentiates successful cases of secession from failures.

Drawing local, regional and international factors of significance from case study analysis of Eritrea, South Sudan, Katanga, Biafra and Somaliland, this research finds that the presence of high-intensity, long duration civil war that generates regional instability strongly correlates with the outcome of a secession. The presence of physical or material resources does not necessarily enhance the viability of a secession, and the strength of a separatist legal claim as well as their fulfillment of the qualifications of statehood are dependent on other factors.

Introduction

Stability is central to the existence of the international state system, and the framework of international law reflects this (Shaw 443–4). The need for stability is reflected in formal international legislation, and in the existence of international regimes on issues of border alteration, state recognition, and the destruction of states (Krasner 185–205). In addition, the majority of domestic and international case law sets sovereignty and territorial integrity above self-determination (Shaw 444–5, 449). The principles of state sovereignty and territorial integrity form the structural foundation of

that stability. These principles protect the status quo by institutionalizing the permanence of existing states, thereby discouraging separatism within states and rejecting the potential for annexation by other states (390–1; Jackson 23).[1] However, these deeply rooted aspects of the international system conflict with the right to self-determination and the broader regime on human rights (Shaw 443). Though the implications of this conflict remain a point of contention, its existence represents a diversification of the means that actors in the international system believe are necessary to achieve stability, security, and power (Shaw 1036).

The end of the Cold War marked a shift from the broad challenges of nuclear warfare and proxy conflicts to delicate issues of state failure, ethnic and sectarian violence, and trans-national terrorism. Conventional solutions to communal conflict such as sanctions, economic and humanitarian aid, and peacekeeping interventions have proven to be only marginally effective and generally come at significant cost. As a result, alternative methods of resolution, such as secession, have taken a more prominent position among scholars and policy makers (Johnson 140–1, 143). Unsurprisingly, the discourse on secession has failed to translate into accepted practice. International laws and norms regarding state sovereignty and territorial integrity makes secession a rare occurrence in the international system (143, 148, 165).[2] Nevertheless, amidst an interconnected international economy and concern for human rights, the destabilizing effects of internal conflicts as well as fears of providing bases for terrorism have ensured that secession remains a possibility.[3] When it does occur, cases of secession offer unique insight into the framework of the international system and its components because they take place at a crossroads between local, regional and international actors, laws, regimes, and interests.

Compared to other regions of the world, Sub-Saharan Africa has witnessed high levels of post-decolonization political violence (Gleditsch 616). This violence can be attributed to a colonial history that left African states with weak government institutions,

---

[1] By state sovereignty, I am referring to negative state sovereignty, or the freedom from outside interference and intervention. Further, I acknowledge that this principle is far from inviolable, or in other words always obeyed, however its existence does limit both the extent of such interference, and the necessary level of power a state must possess to undermine or ignore the sovereignty of another state.
[2] Secession briefly became quite common in the 1990s when the Soviet Union Disintegrated.
[3] Ibid.

arbitrary borders, and undiversified commodity driven economies that are conducive to fueling conflict. These traits strongly correlate with separatism, which has led predictive models to regard Sub-Saharan Africa as uniquely prone to secession (Englebert, Hummel 403).

In contrast to these expectations, empirical analysis demonstrates Sub-Saharan Africa to have a deficit of secession (403). In the 56 years since the start of decolonization, there have only been two successful secessions in Sub-Saharan Africa: that of Eritrea in 1994, and South Sudan in 2011. Existing research has explored this deficit and the two exceptions in the interest of understanding why separatist sentiment arises where it does (403; Hannum, Babbitt 3). However, the current literature has failed to address the question as to what differentiates Eritrea and South Sudan from other major cases of separatism in Africa, such as Biafra in Nigeria, Katanga in the Democratic Republic of Congo, and Somaliland in Somalia.

Recognizing the importance of secession and the opportunity it affords to learn about the international system, this thesis seeks to address this question. Case study analysis of Eritrea, South Sudan, Katanga, Biafra, and Somaliland is employed to identify significant local, regional, and international factors that serve as determinants in the outcome of a secession. This thesis finds that high-intensity, long duration civil wars that generate external instability have the strongest positive correlation with the outcome of a secessionist movement. Research also indicates that the presence of geographical resources or commodities do not necessarily enhance the viability of a secession, which disputes existing theoretical work on the connection between resources and secession. Likewise, the strength of the legal claim of a secessionist movement and their fulfillment of the qualifications of statehood do not necessarily boost the viability of a secession. The impact of these factors may be significant but are dependent on other factors.

The fate of a secession is ultimately dependent on the decision of existing states as to whether or not extend recognition. As recognition of a secession leaves the recognizing state vulnerable to external consequences (unless the host state has already dissolved or agreed to the secession) and internal consequences, the cost of inaction must outweigh the potential price of recognition.

Understanding Secession

Secession is a process aimed at creating a new sovereign state on territory internationally recognized as forming part of an existing state (Englebert, Hummel; Radan 20). The process of secession is completed with the acquisition of sufficient official international recognition as a sovereign state, and the obtainment of membership as such in the United Nations (Englebert, Hummel 403; Radan 20). U.N. membership is necessary to ensure the juridical sovereignty of a state, without which the secession remains vulnerable to elimination by the host state.[4]

What Motivates Secession?

Discourse on the motivations behind secession can be divided into two interrelated categories: overarching arguments as to the decision-making structure of a secession, and substantive motivational factors. Opportunity structures, also known as rational-choice models, dominate the majority of explanatory literature, which conceptualizes nationalism and secessionist movements as constructs of cost-benefit analysis by rational actors (Pavkovic, Radan 41–2). For example, Paul Collier and Anne Hoeffler argue that secession is driven by economic gains, and that violations of human rights do not play a significant role (Hannum, Babbitt 3). A small portion of the literature asserts that non-quantifiable factors that are discounted by opportunity structures, such as history and culture, contribute to the probability of secession (Pavkovic, Radan 42). They argue that existing opportunity models fail to account for the capacity of individuals to diverge from the group (42, 48). They also believe that factors external to a conflict may alter the cost-benefit analysis of the actors involved, such as technological advancements or shifts in the values of the international system (42, 48).

The rational choice model holds obvious weight for understanding secession. However, individuals are not entirely rational actors. Their perspectives, and thus the elements analyzed in a cost-benefit analysis, are not limited to those that are quantifiable. Considered in this way, and because quantifiable and non-quantifiable factors do not

---

[4] Given this, secession as a method of state creation favors the Constitutive Theory of state recognition.

inherently conflict, I propose an alternative superstructure that takes opportunity structures, non-quantifiable factors, and external shifts into account.[5]

As with any other situation where a group is involved in a conflict or dispute, two potentially separate but interactive mechanisms are at work: an organizing factor and a mobilizing factor (Pavlovic, Radan 42; Englebert, Hummel 403). This does not discredit existing research, but instead provides a simplified structure through which concepts may be applied and examined (42; 403). The structure of any secession may be analyzed through identifying one or more elements that could serve as one or both of these factors, and may be broadly categorized into four bases:

1) Identity

2) Resources and Sovereignty

3) State Policies and Behavior

4) Viability Modifiers

Each of these factors may serve as both the organizer and motivator of a secession movement, however in most cases a dynamic relationship exists between them. Factors may reinforce one another, and the motivations behind secession may shift over time. This can occur for a variety of reasons, for example in reaction to the positions of existing states, or if a Diaspora or other another external support base is involved. Such groups may offer a distorted or fictional explanation to justify further support for the conflict.

Identity and Secession

Relevant literature on the construction and manipulation of identity is centered on the concept that ethnic and national identities are endogenous to "social phenomena such as mass politics and capitalism, both in how and when they were created and in the nature of their existence" (Green 4). As Elliot Green highlights in his work "On the Endogeneity of Ethnic Secessionist Groups," there are two schools of thought regarding the way in which ethnically motivated political action may function: the elitist and the mass (5). The work of Paul Collier, Anke Hoeffler and others represent the elitist school, which asserts that elites manipulate concepts of ethnicity, creating and shaping feuds and hatreds that

---

[5] The idea behind the need for this structure was to open the cost-benefit analysis theories to external factors within the scope of this paper. I grant that this may be considered unnecessary, or irrelevant, as such a structure may already exist in an alternate form.

shape the groups themselves (5). In contrast, the mass school "ascribes power to all members of a given ethnic group, who collectively choose to associate based on the material benefits they can potentially possess as a group" (Bates 152–171).

In either case, individuals possess a repertoire of potential identities that they can utilize, and through which others attempt to appeal to them, depending on the situation (Posner 2). The basis for identity as an interactive variable lies with the principles of identity construction and identity choice (2). Identity construction defines the array of political identities that may be mobilized within a society (2). Sub-national, national, and international markets, events, and political institutions determine traits that are viable for mobilization (2). [6] These elements create the bases of identity by shaping the competitive arena as well as the incentives for individuals to participate. These choices achieve society-wide impact as individuals coordinate their decisions in response to these bases (3).

The array of identity options available to an individual depends on their background and the dynamics that determine their political saliency. Individuals then utilize those identities that they believe best enable them to achieve their goals (Posner 2). The fault lines of identity that emerge then represent an aggregation of individuals' decisions as to what coalition will be the most political and economically beneficial to join or create (2–3). Benefits are not solely positive in nature; people may align themselves with one another for the purpose of protection or the mitigation of detrimental policies and events. In other words, identity operates as an organizing factor for secession in the pursuit of positive benefits, and economic and political incentives operate as factors of motivation. On the other hand, identity may operate as both the point of organization and motivation where negative benefits, like protection and survival are concerned (the notions of positive and negative are the same as those used with aspects of state sovereignty).

Understanding the structure and operation of identity coalitions is important to understanding secession. When individuals are unable to manipulate their situation to obtain access to the resources, power, and protection of the sovereign state, or are

---

[6] Theoretically any characteristic could serve as a viable identity basis for mobilization; however, in practice greater variation among those possessing such traits effectively rules out more minor features, such as hair color or favorite colors.

otherwise unable to participate in or protect against the government, they may seek to secede. This may be in part due to the way in which identity can modify the cost-benefit analysis of a potential or ongoing secessionist campaign. Notions of ethnic or religious solidarity give support to separatists through Diasporas, and other state and non-state actors. Such groups may contribute to secessionist sentiment and fuel the conflict, as "they tend to keep grievances alive, offer irredentist support, magnify beliefs in ethnic purity, and provide funding to local organizations" (Englebert, Hummel 404). For example, the Tamil Tigers in Sri Lanka and government of Somaliland have received significant support from their respective Diasporas (404).

Ethnicity, Ideology and Regionalism

Ethnicity and ideology have dynamic relationships with processes of identity formation. Nationalist sentiment may be built upon one or both of these, or it may conflict with them. In relation to secessionist movements, the two may be distinct or overlap as unifying factors. For example, in South Sudan, concepts of blackness aligned with Christianity and Animism within a number of different separatist groups (Hannum 308–9). However, regionalism can serve as another basis for identity, as the case of Eritrea demonstrates (Pavkovic, Radan 497).

By nature of their mechanics, ethnicity, ideology, and regionalism differ in effect as the foundations of secessionist movements. Ethnicity provides a strong basis for separatist movements, as it is perceived to be relatively exclusive. Similarities in language, culture, history, and more importantly the way in which others perceive and interact with potential members of an ethnic group, make it a prudent base for secessionist claims. Concepts of ethnic unity and purity, which may be utilized by people on the ground, or by members of a Diaspora, encourage the construction of support networks (Posner 3).

Ideology, particularly religion, being more expansive and inclusive lacks the rallying strength of ethnicity or culture. That being said, ideology may motivate secessionist movements by garnering external support from states and non-state actors such as religious institutions. Ideology may also skew the cost-benefit analysis of potential separatists, given the intangible elements involved.

Regionalism is a form of Segmentary Opposition, and exists as a reactionary sub-nationalist identity (Barth 5).[7] Though not necessarily possessing the unifying strength of ethnicity and the support networks of religion, regionalism represents a critical basis for mobilization as it presents fewer points of internal dispute amongst separatists. Regionalism represents a form of opposition coalescence, which develops when common opposition to a group or entity induces the alignment of multiple identities to form a new base. This may also occur among ideological and ethnic identities.

Resources and Sovereignty[8]

Economic incentives and material benefits are generally regarded as the dominant determinants and explanations of civil war and secession (Hannum, Babbitt 4; Englebert, Hummel 404). Existing research has shown a strong correlation between the presence of natural resources and instances of separatism (Ross 337–8). This work has produced a range of explanations for this relationship; however, they tend to center on the control of resources as a motivation for two groups: coalitions that would benefit from a secession and the existing sovereign state (337–8). When a region or group considers itself economically disadvantaged by its relations with the sovereign state, it may "reassess the relative costs and benefits of belonging to a national union, and in turn contemplate seceding" (Bookman 39). Interestingly, such a conclusion is not unique to regions that may be considered rich or poor (39).

The opportunity to control and profit from natural resources offers a strong incentive for local actors, though their course of action may depend on the type of resources. Resources are not limited to material goods, alternatives such as coastal access, or state sovereignty may also be viewed as desirable to control (39). Groups may launch a secessionist movement if they are not being provided sufficient services or benefits by the central government and seceding would allow them to create or retain necessary agricultural and material goods (Englebert). Where a group or region is deprived of their autonomy and authority, they may be motivated to secede to reacquire

---

[7] Segmentary Opposition refers to the tendency of groups to put aside their own conflicts and ally when faxing a greater external threat.

[8] I base my perspective of natural resources on the broad definition used by Michael Ross in his work "Natural Resources and Civil War," pg. 19. These include oil, gas, precious stones, non-fuel minerals, lumber, agricultural commodities and drugs.

their power in a new state or force concessions by the central government (Hannum, Babbitt 3).

Resources requiring significant infrastructure development, stability, and foreign investment, such as oil, are more likely to motivate secessionist movements (Englebert, Hummel 405). As the central government is capable of contesting agreements within its territory, local actors desire sovereignty of their own (405). [9] In contrast, resources that require little to no capital or foreign infrastructure and investment, such as diamonds, do not require sovereignty to be extracted. In these cases, local actors back a local power, rather than trying to secede (405; Nabudere 296). [10]

Sovereignty is an incentive in and of itself. Sovereign states can receive development, food, arms aid, loans, and they have a domestic monopoly on violence. Beyond these positive incentives, seceding states that leave a rump state behind tend to emerge without the burden of debt, though that depends on the circumstances and investments in the area seceding. That being said, where natural resources converge with these considerations (potentially offering a means by which groups may fund opposition) there is an increased change of foreign support and intervention (405; Hannum, Babbitt 34).

State Policies and Behavior

On a passive level, the condition of a state has the potential to motivate or discourage a secessionist movement (Englebert, Hummel 406). The style of government is less important than whether or not a state is able to combat a separatist movement. Similarly, levels of political change or democratic institutions only matter in how they affect the strength of the state. Weak or dysfunctional states may motivate secessionist movements as groups seek to provide themselves with services that the state has failed to provide. Somaliland is one such example, though it was initially a distinct state and

---

[9]The perceived necessity for local sovereignty is due in part to the ability of the existing sovereign government to oppose any such development or contracts, as well as the limitations of significant development aid and governmental business entities regarding dealings with non-states.

[10]It should be noted that in dealing with resources that are easily extracted and transported, it is plausible that state policies and behavior may incline the people to seek to secede, though this may be counter-acted in both cases by directly dealing with corporations and side- stepping the central or regional government. An example of such behavior is found in the eastern provinces of the D.R.C. where various mining companies such as Anglo-American and Banro-Gold have bartered with local populations and on-site rebels to have access to mining sites and mineral rights.

represents a relatively unique case (Paquin 153). Weak or dysfunctional states rely more heavily on political violence, which may spill over into secessionist sentiment if grievances are created in the process. States in transition, which are not necessarily weak or dysfunctional, also may also provide the impetus for secession, as periods of democratization and economic transition have an impact on identity dynamics and may intensify identities, and in the process heighten security dilemmas (Englebert, Hummel 406).

The state actively constructs the cost-benefit analysis for potential separatists, as it is able to influence which identity elements are relevant or politically viable. The usefulness of any potential or existing coalition depends on the boundaries and identity bases of the political arena. Policy shifts among state institutions alter these bases and the shape of the political arena, and change the fault lines of identity. More directly, the state is capable of creating or nullifying grievances amongst the population through its actions and policies. These policies may be entirely politically aimed, such as the exclusion of a region or group to solidify a base of power, though they have also been found to overlap with the location of natural resources (Ross 34). In cases where valuable natural resources reside in an outlying region of a country or where a relative minority exists, states may engage in 'pre-emptive repression', or rather, campaigns of terror and suppression (34).[11]

Viability Modifiers

Shifts at the system level, be it regional or international, alter the conditions of possibility for secession. In other words, factors and events beyond the immediate context of a potential secessionist effort impact the viability of an effort and the cost-benefit analysis of a group considering separatism. I define these as 'Viability Modifiers.' A viability modifier is any domestic, regional, or international factor that impacts the cost-benefit analysis of a potential or existing separatist group. Thus, the judgment of a state and a separatist movement may shift over time and take changes in the international system into account.

---

[11]The research of Michael Ross has shown such antagonistic activities and behavior to be relatively unique to combating separatist efforts.

A Viability Modifier may be separate from the previously listed factors, or may represent a shift among them; for example, the price of a commodity in the international market. Broadly speaking though, Viability Modifiers encompass everything from market prices to shifts in norms and regimes relevant to secession: the right to self-determination, territorial integrity, positions on partition and secession as policy options, and others. For example, when the army in Mali launched a coup against the government, this was a Viability Modifier for separatists in the northern part of the country, who announced their independence and the creation of Azawad (Palus).

Differentiating these factors and understanding how they interact with one another is critical to a broader discussion of secession. This is in part due to the deficit of material on the subject. More importantly, understanding how factors can interact allows a more thorough examination and analysis of different cases of secession. The ability to recognize the organization and motivation factors of a secession is necessary to enable a better understanding of how internal and external actors interact and why.

Why Africa: Africa as a Unique Setting for Secession

Africa is a unique setting for the study of secession, and the exploration and identification of significant factors involved in determining the outcome of secessionist efforts. As a region, Africa has a distinct relationship with secession, one that rests with the nature and processes of African state formation, as well as the contemporary conditions and characteristics of African states.

On the Nature of African State: Common Sense Secession

International state borders are the recognized delimitations of states' sovereignty and territory, as well as the area over which states are expected to exercise effective control (Okhonmina 177–179). States maintain a dynamic relationship with their borders, where the capacity of the state impacts the implications of its borders, yet its borders in part determine the capacity of a state. However, state formation and territorial consolidation have not been a uniform process throughout the world, and borders in different regions have differing implications and effects.

The contemporary borders of African states exemplify this point, as they follow the administrative boundaries of the European Colonial powers and which were transformed into international borders through the principle of *Uti Possidetis* during

decolonization (Ratner 595). These boundaries were designed arbitrarily in relation to the reality on the ground, being drawn with the simplest methods to reduce conflict among the European powers, without concern for demographic, geographic, or ethnographic factors, or the construction of viable states (Herbst 25). In some cases nascent power balances that had developed under colonial rule were disrupted as administrative zones and were either absorbed or consolidated. As such, this change left them severely disadvantaged within their new area and generally excluded from participation in discussions and decision-making during decolonization (effectively denying them independence). South Sudan and Cabinda are both examples of this practice (Southall 156, 158).

Decolonization raised the question as to whether the colonial borders should be restructured in an effort to resolve their arbitrary basis. However, participating African elites opted to maintain their respective borders as they were laid out prior to independence (Ratner 595). African leaders argued that opening the borders up to question could lead to violent and destabilizing conflict. European states backed this decision as it allowed them to rapidly depart with little obligation to arbitrate or mediate between different groups. Opposition to any efforts to alter or undermine the borders was then engrained into the Charter of the Organization of African Unity, declaring the existing boundaries inalterable (Org. of African Unity 17).

On the whole, African states differ from those of many other regions as they do not represent the molding of identity through conflict, nor are they a reflection of the dynamics of exercisable economic, military, and administrative power and competency (Southall 153–156). Their boundaries were designed with neither ethnic identity nor the future viability of the area in mind. Given these factors, both causes and means, Africa is potentially endowed with not only a greater number of separatists, but separatists who have a stronger argument than separatists from many other areas of the world.

Contemporary Conditions: The Secessionist Deficit as a Legitimizing Force

Decolonization enabled the formal transfer of control to African leaders. However, they inherited states more in name than in capacity, as they had weak institutions, societal fissures from divisive colonial policies, and vulnerable economic structures centered on the exportation of raw resources and singular commodities (Herbst

97, 118, 135, 140). As one might expect given these circumstances, in the 56 years since decolonization began, Sub-Saharan Africa has witnessed consistently high levels of political violence and internal conflict (Englebert, Hummel 399).

The states of Sub-Saharan Africa possess a host of social, economic, geographic, and state characteristics correlated with civil war and separatism (403–499). African states are regarded as being relatively young and still undergoing the throes of national integration and construction of a common identity (404). They consist of socially heterogeneous and polarized populations whose groups commonly engage in zero-sum interactions for control of the state (Englebert, Hummel 400). Significant reserves of natural resources exist that may act as incentives for the initiation of conflict for other internal and external actors, as well as being exported and sold by civil war combatants and separatists to fund existing conflict (400). African states are commonly 'culturally alien' to their populations, in contrast to states in other parts of the world, as a result of their borders, capital placement, and the colonial indoctrination of African elites (400). African states widely have weak records of evenly providing services and goods for their citizens, and control over these institutions is commonly the basis for conflict and neo-patrimonial loyalty systems. The logical derivative of these factors is that a significant proportion of the political violence should be related to separatism.[12]

Yet as Pierre Englebert and Rebecca Hummel illuminate in their work "Let's Stick Together: Understanding Africa's Secessionist Deficit", secessionist conflict in Africa is well below predicted levels (401). Between 1960 and 2005, separatism only accounted for 27% of all domestic conflicts in Africa (400). Similarly, the levels of separatism in Africa are notably lower than many other regions of the world. In the same period of time, based on recorded instances of separatism, the average annual probability of a secessionist conflict occurring in Africa was only .06%, in comparison to .16 % in Asia, and .08% in the Middle East (and was matched by .06% probability of secession in Europe) (402). Though the deficit of secessionist activity offers fewer case studies to analyze, it deepens the gains of studying separatism in Africa, as those that do occur

---

[12] As mentioned previously, civil war and secession are differentiated by the intent of the groups, however I acknowledge the difficulty in determining a 'critical date' upon which to judge the true or dominant motives of a group.

should be more dire situations where internal actors have a more viable opportunity or are more committed to seceding.

Examination of International Law: Viability of Secession

The international community does not recognize secession as a right. (Hannum 13). Yet "international law does not prohibit secession, whether voluntary or violent, but it has neither recognized a right to secede nor identified even tentatively the conditions that might give rise to such a right in the future" (14). In the place of such a right, the viability of secession is reliant on a framework of distinct but interrelated aspects of international legal principles: state creation, *Uti Posessidetis*, self-determination, recognition, territory, and the privileged role of the state.

The Framework of Statehood and Recognition

Acquisition of a sovereign state necessitates secessionist movements to fulfill the basic requirements of statehood as laid out in the 1993 Montevideo Convention on the Rights and Duties of States: a permanent population, a defined territory, a government, and the capacity to enter into relations with others states (Montevideo Convention). However, these requirements are broadly defined and subjectively interpreted, which makes fulfillment of these measurements of statehood relatively easy. This has in part been enabled by the development of the right to self-determination, which has modified the traditional circumstances and forms of these requirements (Shaw 183). No minimum population number has been established; moreover, defined and settled boundaries are not entirely necessary at the time of statehood. Instead, the area in question must be contiguous and undeniably controlled by the government aspiring to acquire statehood (183). The government has no necessary capacities or level of functionality beyond being able to engage in legal relations as it sees fit, being able to legislate and enforce the agreements it makes with states (181). This is the most crucial requirement of statehood, as it necessitates not only effective control but the legality of exercising control in so far as other states accept that legality and have a willingness to engage in relations by it.

The Montevideo Convention asserts that the existence of a state, and its possession of statehood, is independent of recognition by other states, and as such is able to defend its 'integrity and independence' regardless of its status as a recognized or

unrecognized state (Montevideo). The Convention further states that the rights of other states under international law form a limitation to the exercise of this right (Montevideo).

Legal theory on state recognition is divided between Declaratory Theory and Constitutive Theory (Shaw 363). In accordance with the view of the Montevideo Convention, Declaratory Theory argues that recognition is an acceptance by 'existing states' of extant factual realities and circumstances (368). A new state, a category that includes secession under this theory, will be legally constituted by its own efforts and does not have to wait for recognition by other states, or the virtue of their consent to do so (368). Opposing Declaratory Theory and contrasting the Montevideo Convention, Constitutive Theory is set in the belief that states acquire "legal personality" and are established as full subjects of international law through the act of recognition by other states and by virtue of the will and consent of the international body of existing states (368–9). The practice of recognition has shown to take a middle road between these two theories. In concept, recognition indicates that the new state is capable of conforming to the basic requirements of international law and satisfies the qualifications for statehood. In practice, this rhetoric justifies highly politicized judgments of self-interest (372–3).

The Framework of Territory and Sovereignty

The international system is premised on the stability and supremacy of territorially based states. The centrality of this concept is reflected in the framework of international laws and norms, particularly in the principle of territorial integrity, which prohibits any type of interference in the domestic affairs and territory recognized as being under the jurisdiction of a state (443). The state itself rests on the notions of positive and negative sovereignty (409). Positive sovereignty refers to the exclusive right of a state to exercise authority over its territory (412). Negative sovereignty refers to the right of a state to be free from external interference (412). At its core, the relationship between sovereignty and territory rests upon the concept of title, in other words, the factual and legal conditions that deem the ownership of territory, a relationship that while absolute in domestic territorial questions, is relative concerning international territory (412).

Self-Determination as a Vector to Statehood

Separatists and advocates of secession as a policy cite the right to self-determination to justify and enable their claims; particularly where the borders have been

formed by relatively artificial means (200). Self-determination evolved from a means of protecting minority populations, and is defined as the right of 'all peoples…to freely determine their political status and freely pursue their economic, social and cultural development' (Basak 198). Self-determination was initially instated in international law through the Treaty of Versailles and the Fourteen Points speech, and then integrated into the encompassing framework of the U.N. Charter (198). However, the concept of self-determination as a right was further legalized in the 1996 International Covenant of Civil and Political Rights (ICCPR) and the International Covenant of Economic, Social and Cultural Rights (ICESCR).

However, self-determination is not a generalized right of secession; international law does not hold secession to be a right, nor does it regard self-determination as intended to or able to dissolve the sovereignty and territorial integrity of a state. The International Court of Justice has noted that all states have a right to territorial integrity and sovereignty regardless of the process of their formation (198).

In line with this, the reach of such arguments are limited as self-determination is engrained as being subservient to other principles of international law and its realization must conform to the Charter of the United Nations, as well as other sources of international law (228). Thus, though self-determination exists as part of an expanding regime on human rights, its scope and applicability to secession is limited (230, 269).[13]

The Viability of Secession

As this examination has shown, the legal and logistical framework of the international system and its norms oppose secession, and deny its viability. Yet, international law remains an evolving subject, and the internal denial of the right to self-determination may induce the development of an acceptable and legally recognized principle of external self-determination, or secession, through a series of exceptions.

---

[13]Denial of the viability of secession through the right of self-determination was initially legalized in 1970 with the passage by the U.N. of the Declaration on Principles of International Law Concerning Friendly Relations and Co-operation among States in Accordance with the Charter of the United Nations. The Declaration of Friendly Relations defined self-determination employing similar language to the ICCPR and the ICESCR before stating that 'Nothing in the foregoing paragraphs shall be construed as authorizing or encouraging any action which would dismember or impair, totally, or in part, the territorial integrity or political unity of sovereign and independent states conducting themselves in compliance with the principle of equal rights and self-determination of peoples as described above and thus possessed of a government representing the whole people belonging to the territory without distinction as to race, creed, or colour.'

For now though, without significant incentives for major powers to extend recognition, secession remains dependent on the agreement of the host state (whether voluntarily or through the inability to obtain military victory over a secessionist movements that obtains control over a significant percentage of territory) and the cost-benefit analysis of external states to act or influence the negotiations.

Case Study Analysis: Tracing Secession and Identifying Differentiating Variables

The aim of this analysis is to draw out the factors that seem to be the significant determinants of whether or not a secession will achieve statehood. Case study analysis is employed to identify these determinants and, through examination of their contextual roles, to understand how they impact the fate of a separatist movement. The case studies are ordered by their similarity, progressing from the successful cases, to the failures, to the exception. Each case study is broken into five sections. First, a brief overview of the secession is given. Second, the origin of the conflict is identified and explored. Third, the domestic, regional and international course of the secession is reviewed. Fourth, the arguments and demands of the secessionist movement are examined in conjunction with the basis of the conflict to determine their legality. The last section examines the costs of the conflict, the presence of resources, and the involvement of external actors.

These case studies enable the compilation and discussion of potentially significant determinants. These determinants are identified and judged before being used to construct an investigative framework that is applied to all of the cases to see if any trends or patterns are identifiable as to what differentiates failure from success.

Selection of Case Studies

Eritrea and South Sudan were selected for case analysis because in both cases, separatists achieved statehood. Katanga and Biafra were both significant separatist conflicts, in terms of the costs and the actors involved, yet failed to acquire statehood. Lacking statehood despite possessing the most acceptable legal justification, Somaliland offers a unique opportunity for analysis.

Eritrea: Overview

Eritrea emerged as the first and, as of then, only exception to the inviolability of post-colonial African state borders. Eritrea was neither ethnically nor religiously homogenous, but was bound by the experience of Italian colonialism and persecution by

the Ethiopian State (Pavkovic, Radan 497). After 30 years of conflict, Eritrea attained recognition as an independent sovereign state in 1993 after a U.N. monitored referendum on independence, in which 98.5% of Eritreans supported independence (497). Ethiopia, the host state, agreed to the referendum in 1991, as it lacked either the will or capacity to oppose the measure, or the resources potentially necessary to maintain unity (Paquin 128). Ethiopian consent to the process and subsequent recognition of Eritrea enabled O.A.U. support, and uncontested international recognition of Eritrea (Pavkovic, Radan 497).

Origin of Secession

Though Ethiopia rebuffed Italian colonial efforts it ceded the territory of Eritrea to Italy in 1889 (497). From then until the Second World War, Eritrea existed as an Italian colony, briefly being united with Ethiopia in 1935 under Fascist Italy until its defeat by local and British forces, which allowed Ethiopia to reclaim its independence (497). In contrast, Eritrea became a British trusteeship for the remainder of the war until a decision could be made as to its future (497). In 1953, the U.N. made the contentious choice of placing Eritrea in a federation with Ethiopia through the Contractual Federal Act. This act structured the federation such that Eritrea was separate and autonomous, with its own constitution and elected government (498). Despite the designs of the U.N. Ethiopia soon began to violate and undermine Eritrean autonomy (Paquin 130). The culmination of this pattern of actions occurred in 1962, when Ethiopia revoked the federal status and nullified its autonomy through annexation, transforming it into the 14th province of Ethiopia (130). The violation of Eritrean autonomy extended into the linguistic realm, when contrary to the constitution of Eritrea, Amharic replaced Tigrinya and Arabic as the official languages (Pavkovic, Radan 46). Though Eritreans appealed to the U.N., it avoided any responsibility by claiming the matter to be domestic in nature (498).

Review of Secession

These actions served as the catalyst of the secessionist conflict, causing the emergence of Eritrean secessionist insurgencies and the launch of guerilla warfare campaigns (Paquin 130). In 1974, the Emperor of Ethiopia, Haile Selassie, was overthrown by a military coup, which formed a military regime, also known as the Derg, led by Menguistu Haile Mariam. The regime proceeded to use brutal military force to

silence and severely suppress all forms of dissent and ethnic claims (Spears 46). This effort backfired as it galvanized secessionist movements, induced the emergence of internal non-exclusive insurgencies of marginalized groups that sought to overthrow the state, and "drove Ethiopians in increasing numbers toward the rebel fronts" (Spears 46, 71; Paquin 131).

Ethiopia sought to address these insurgencies through the use of military force, posturing for U.S.S.R. aid as their relationship with the U.S. and the wider regional context made it unwilling to supply arms (Paquin 130). Previously the U.S. had maintained support for Ethiopia due to the importance of a communications station in Eritrea, and its desire for stability in that area (Spears 55). However, technological advances and relocation of the station removed those incentives. As the U.S. withdrew, the U.S.S.R. moved in, committing $1,000,000 worth of armaments, hundreds of military advisors and 11,000 Cuban troops (55). Prior to this, Eritrean separatists had pushed Ethiopian forces out of Eritrean territory, denying Ethiopian sea access, which necessitated a restructuring of the army in the view of the regime (57). The shift was primarily from quality, with 40,000 British trained soldiers, to quantity, with several hundred thousand poorly trained and poorly disciplined but heavily armed soldiers (55).

In 1982 Ethiopia launched Operation Red Star, a massive effort utilizing its restructured army and the support of the U.S.S.R., and erased almost of the gains Eritrean separatists had made, regaining significant control over Eritrean territory (61). Yet it failed to capture key Eritrean cities, and provided Eritrean separatists with massive amounts of modern weaponry and equipment whenever the haphazard Ethiopian soldiers withdrew or surrendered (Spears 61). Soviet support pushed back the gains of the separatists, however the failure to establish effective control of major cities made these gains hollow. The dominant separatist group, the Eritrean Peoples Liberation Front, was able to construct "structures and institutions that made Eritrea a virtual state within a state" (61). By setting up government administration, farms, hospitals, and weapon factories, separatists were able to satisfy their logistical needs at the same time as fulfilling the qualifiers for statehood (61).

Expanding rebel capacity met with the fall of the Soviet Union and the withdrawal and discontinuation of support for Ethiopia caused the fall of the government

(Paquin 130). The transitional government ratified a new constitution allowing for secession and worked with the EPLF to set the date for a referendum on independence. The EPLF postponed the referendum until the new government could stabilize (81). Then in the wake of the U.N. monitored referendum on independence, Eritrea declared its independence on May 24th 1994, being recognized by Ethiopia, the O.A.U., and the international community, and obtaining U.N. state membership, as well as acceptance into the World Bank and the International Monetary Fund (84).

Legality of Claim

Eritrea had an excellent legal basis for its claim to independence. Eritrean separatists argued that theirs was a case of independence, not secession. The basis of their claim centered on having been historically denied their right to self-determination, first by Italian colonialism, and then through the U.N. imposed federation with Ethiopia, which illegally absorbed it soon after (Pavkovic, Radan 498). The claim was set in clearly defined borders demarcating contiguous territory, over which the separatists exercised effective control (Herbst 107). Moreover, the claim asserted the right of the Eritrean peoples, which were neither ethnically nor religiously homogenous. Ethiopia offered a weak counter to their assertion, claiming that pre-colonial unity had existed under a single empire, however even this only occurred from 1880 to 1885 (Pavkovic, Radan 43).

Costs, Resources, and External Actors

An array of external factors impacted the conflict or emerged as a result of it. In terms of external stability, the conflict had a major spillover effect, with an estimated 500,000–600,000 refugees fleeing to Sudan, Kenya and Djibouti (135). An additional 80,000–100,000 exiles and émigrés fled to Arab States (135). At roughly 100,000 strong in the U.S., the Ethiopian and Eritrean Diasporas supported their respective movements, but shared the goal of overthrowing the Ethiopian state (Paquin 148). Neighboring and regional states backed the Eritrean independence movements, particularly Libya, Egypt, Sudan and Somalia. Syria, Iraq, Kuwait, Saudi Arabica, China, South Yemen and Cuba also provided varying degrees of support (Rothschild, Olunsorola 217). The U.S. and the Russia came to be involved supporting the push for independence as means for Russia to demobilize an unaffordable military presence and for the U.S. to bring about stability. Eritrea lacks resources in the general view, having no significant discovered reserves of

oil or minerals; however it does possess access to the sea, a potentially critical good for neighboring landlocked countries.

South Sudan: Overview

Through the right of a referendum on independence held in January 2011, as designated in the 2005 Comprehensive Peace Agreement that ended the Second Sudanese Civil War, South Sudan became an independent and sovereign state in July of 2011. Waged across nearly 60 years and two major civil wars, the conflict was one of the deadliest and most destabilizing wars in Africa (Pavkovic, Radan 506). The origin of the conflict lay with colonial consolidation, and subsequent neglect and maladministration by the government, due in part to the centralizing of power in the North as well as the later discovery of significant oil holdings in the South (505).

Origin of Secession

After Britain and Egypt conquered Sudan in 1898, Britain administered Northern Sudan and Southern Sudan separately for the next 40 years or so (Hannum 309). This differentiation was necessary as the populations differ significantly; the North is of Arab descent, Islamic, and speaks Arabic while the South is of Black descent, is Christian and Animist, and speaks indigenous languages and English (308–9). Prior to decolonization in the 1950s, South Sudan was merged with Sudan, however Egypt and the United Kingdom gave executive powers and the role of civil administration to the North. (Pavkovic, Radan 505).

South Sudan never had a chance of contesting the shape of Sudanese decolonization, as the consolidated borders were solidified in a conference in 1947 (Hannum 307). Though held in Juba, the capital of the South, the only Southern representatives were on the payroll of the British, who sided with the interests of the North and of Egypt, which were for unity (307). Following this conference, Southerners remained politically marginalized through decolonization and the formal independence of Sudan in 1956 (307). This structure converged with religious and cultural differences and a weak post-colonial state to create conditions for civil war (Pavkovic, Radan 505).

Review of Secession

The conflict began in 1955, the year prior to Sudanese independence, and was drive by political frustration and marginalization of the South (506). As the war

progressed, Southern rebels succeeded in forcing the withdrawal of government forces (506). These gains contributed to the launch of a military coup in 1958. The military government combated dissent with force, and sought to weaken Southern political power and impose Islam and Arabic on the South (Hannum 309). In 1964 these measures induced massive general strikes in the North in protest of the government, leading to its fall (309).

The following year, a transitional government was formed, which held a round table conference in March in an effort to bring an end to the conflict (310). Though being attended by the representatives of major factions in the North and South, the conference reached no conclusion, and was undermined by anti-Southern violence (309). Amidst these overtures of peace, Southern groups set up a provisional government with local councils, courts, schools and clinics in the areas under Southern control (311). In what has become a theme of Sudanese administrations, the government was once again overtaken by a military coup, with Numeiri taking power (311). The First Sudanese Civil War then ended in 1972, as the government and major rebel groups signed the Addis Ababa Accords (Pavkovic, Radan 505), which established regional self-government for the South, granting it control over economic and security matters (Hannum 312). Peace was also secured through alteration of the constitution to recognize English as a primary language and the right of the South to religious freedom (312).

The measures of the 1972 agreement were sufficient in bringing about peace, with minor residual instances of conflict. However plans to construct a canal and divert Southern water, potentially causing the loss of arable land for 2.6 million people, created growing tension. The discovery of oil reserves in Southern and border territory in 1979 by Chevron exacerbated the situation, as the state built refineries in the North and sought to redistrict the oil sectors out of Southern control (Hannum 313). Reacting to these tensions in 1981, the state removed the regional president of the South, dissolved its assembly, performed wide scale arrests of dissenters and appointed a military government (313). These antagonistic tactics extended into 1983, when the government imposed the 'September Laws', institutionalizing an 'Islamization' of the state, as well as re-dividing the area of Southern provinces (313).

In response to these policies and actions, the Sudan People's Liberation Movement and its armed wing, the Sudan People's Liberation Army were formed with the intent of reforming Sudan, and the Second Sudanese Civil War began when Southern soldiers mutinied against orders to transfer to the North. As the civil war continued, the government was wracked by two military coups, one in 1985, which led to failed peace talks, and another in 1989 in the wake of a ceasefire agreement and a debilitating famine (314). The 1990s witnessed the spread of the conflict to other areas of Sudan as other groups took up arms against the government, and the efforts of regional powers, specifically the Intergovernmental Authority on Development, to bring about peace through talks (Pavkovic, Radan 506).

 By 1998, the military position of the North and its ability to sustain the conflict was rapidly eroding (Natsios 162). The situation of the state worsened considerably in 2001 as infighting between Southern rebels ended and they aligned in opposition to the state (Pavkovic, Radan 596). These circumstances brought about the signing of the Machakos Protocol by the SPLM and the state government (506). Through the Protocol, parties involved agreed that South Sudan had a right to self-determination within the structure of the Sudanese state. Additionally, the division of state and religion, and the practical implications of this division were agreed upon (505).

The Machakos Protocol was the first of a series of agreements that formed the 2005 Comprehensive Peace Agreement, which ended the Second Sudanese Civil War and established the date for a referendum among Southern Sudanese on whether to secede or not (506). Internal and external factors led to the participation, agreement and general adherence to the C.P.A. in contrast to past agreements. Internally, the civil war had had a debilitating economic cost, consuming more than half of the oil revenues of the state (Natsios 163). Politically, the state believed that the potential secession could be mitigated and that unity could be made attractive to the Southerners (163). Given that throughout the majority of the conflict rebels sought to reform the state, nor secede, this was not implausible. However, Khartoum did not consider how past experiences could undermine unity arguments, or how to communicate such a message. Externally, United States intervention in Iraq and Afghanistan made maintenance and improvement of

relations with Washington a high priority for Khartoum, which required an end to the civil war (Natsios 171).

The C.P.A. formally established the autonomous Government of South Sudan in Juba and set up the interim Government of National Unity in Khartoum. The key aspects were that it allowed the military wing of the SPLM to continue to exist and receive external assistance to train and modernize, it outlined the sharing of oil revenue between Khartoum and Juba, and most notably it set January 2011 as the deadline for the referendum on Southern independence (171). The C.P.A. differed from past agreements due the level of involvement by external actors, specifically the United States. Though Khartoum commonly discarded agreements in the past, the U.S. offered the government incentives to allow for peace and the referendum to take place (169). Beyond this external pressure, the provisions of the agreement that allowed the Southern military to train, expand and modernize made cost for disruption high and the plausibility of significant bargaining chips or territorial gains low (171). These factors and significant international attention enabled the referendum to be successful held in January of 2011, and with 99% of Southerners voting for independence, and South Sudan to achieve sovereign statehood in July later that year (BBC).

Legality of Claim

South Sudan had a clear and powerful claim to independence, although the aim of Southern actors for the majority of the conflict was reform, not secession. In the framework of *Uti Possidetis*, South Sudan had been administered separately from Sudan under colonialism. Beyond this, South Sudan had been granted the right to secede through the 2005 C.P.A., to which Sudan had agreed. By repeatedly violating past agreements that had measures short of independence, Khartoum demonstrated the intractability of the conflict without a more permanent solution.

Costs, Resources, and External Actors

External actors and factor were significantly involved in determining the course of the two civil wars and South Sudan's eventual acquisition of independence. Similar to the conflict in Eritrea, the conflict in Sudan generated hundreds of thousands of refugees abroad and internally (UNHCR). Beyond the already discussed role of the United States, Ethiopia, Israel, Kenya and Uganda all provided direct and or indirect assistance to the

Southerners (Hannum 326). The South also had support from non-state actors, primarily non-governmental organizations directly tied to the institutions or constituents of Christian churches (Natsios 172).[14] Though supporting the Southerners represented one way for other states to indirectly oppose Khartoum, another major potential incentive is the significant oil reserves found in the Southern region, roughly estimated at 5.4 billion barrels of oil (Trivett). An expanding Diaspora of Southern Sudanese existed throughout the conflict, supporting the effort through lobbying efforts, economic support and ultimately participation in the global referendum.

Katanga: Overview

On July 11th 1960, Katanga, a province in the Democratic Republic of Congo, declared independence and called upon states to recognize it (Rothschild, Olunsorola 201).[15] Though ethnic and political differences existed between Katanga and the D.R.C., the conflict was primarily centered on control over the significant mineral resources and potential wealth of the province (Pavkovic, Radan 122). The attempt at secession was relatively brief, being brought to an end by the intervention of U.N. forces at the request of the D.R.C. without recognition by any other state (122).

Origin of Secession

The origin of the conflict rests with the presence of the mining corporation 'Union Miniere du Haut-Katanga' (Rothschild, Olunsorola 201). Offering mining opportunities and services paralleling those of the central government, the area had a white settler population exceeding 30,000 in 1960 (201). The concept of 'Katanga' as being distinct from the rest of the D.R.C. emerged in the 1950s, when immigration into the area induced fear of displacement and discrimination by the inhabitants. This caused the different demographics of the region to align (201). The interests of the corporation and Belgium soon came to match those of the white settlers as the dependent condition of the country and future profits of corporations was challenged by local political elites (202). Moise Tshombe integrated these sentiments into his political party, 'Confederation des Associations Tribales du Katanga' (CONAKAT) (202). The party dominated regional

---

[14]These included a wide array of Christian non-state actors, ranging from the evangelical 'Samaritan Purse,' conservative Christians, the Catholic Church, black American churches and moderate protestant churches.
[15]I encountered conflicting dates as to the secession of Katanga, with some listing the 17th of July rather than the 11th.

elections, but did poorly in national elections, securing only two minor posts in the central government (202). This led the party to perceive itself as having been politically excluded, and caused it to consider secession as an option (202). CONAKAT initially planned to secede two days prior to the declaration of independence by the Democratic Republic of Congo; however Belgian officials persuaded them to delay their efforts (202). Plans resumed however, when Congolese soldiers mutinied against European officers and targeted white settlers with violence, destabilizing the central government in the process as it struggled to subdue the violence (202).

Review of Secession

Taking advantage of the opening provided by the mutiny and the presence of 6,000 Belgian soldiers within its borders, Katanga declared itself as a sovereign and independent state on July 11th and 'asked all to recognize in us the right of every people to self-determination' (202). Katanga immediately asked Belgium for technical, financial and military aid, specifically in the interest of training a Kantangan military force (202).

The immediate prospects of Katanga were mixed, as the secession effort had the resources of the settlers as well as those of the mining corporation. Belgium did not recognize Katanga but did provide assistance and had its troops in the territory disarm soldiers loyal to the central government (202). Its initial internal prospects severely contrasted with its international presence. African states opposed the secession, viewing it as a plot to undermine the independence of the D.R.C. and enable Europeans to further exploit the resources of the area (202). States throughout Asia expressed similar sentiments (203). Beyond Belgium, other European states did not offer recognition, as opposition by African and Asian states as well as the timing of the secession, being in the midst of decolonization, caused the costs of such an action to far outweigh the benefits (203).

The U.N. also came to oppose the secession, fearing that if successful, Katanga might set a precedent of both rich regions seceding from poor states, and enabling challenging of the territorial integrity of post-colonial states (Pavkovic, Radan 122). In reaction, the U.N. Security Council affirmed the territorial integrity of the D.R.C., asserting that the territory was being decolonized as a unified whole (123).

Patrice Lumumba, Prime Minister of the D.R.C. requested U.N. intervention in order to end the separatist bid and resolve the situation (Rothschild, Olunsorola 203). The Security Council agreed, however it primarily sought to deescalate the situation, not remove the secession (123). Having called upon Belgium to withdraw its troops previously, the Security Council passed Resolution 146, which authorized the deployment of a U.N. force, with the objectives of maintaining order, protecting human lives, and forcing the withdrawal of Belgian troops and foreign mercenaries from the Area (203). The U.N. Force was then deployed throughout non-separatist territory. U.N. troops refrained from engaging Katangan forces out of concern for casualties in what some members of the international community believed was a domestic issue outside the jurisdiction of the U.N (203).

In response to this, Lumumba ordered the D.R.C. Army to retake Katanga (203). The deployment of the Army failed to resolve the situation, as the campaign quickly degenerated into indiscriminate violence and drew into a stalemate when the President removed Lumumba, and threw the government into chaos. Initially this offered Katanga a chance to consolidate, however it soon faced a separatist movement within its own borders that escalated into a second conflict, bearing brutal violations of human rights (201).

These circumstances drove an alignment of positions among regional and international actors in favor of ending the conflict (204). However, concern regarding the potential precedent of such an intervention stymied consensus as to what measures could be taken, with Asian and African states supporting the use of direct force and European states largely favoring the use of negotiation (204). Though no clear harmony was achieved, the Security Council passed Resolution 169, which aimed to maintain the integrity of the D.R.C. and bring an end to the conflict. The Resolution circumvented concerns of setting a precedent by authorizing troops on the ground to essentially act as they saw fit to achieve their objectives, and disarm and remove all foreign forces by any means necessary (204). U.N. troops proceeded to crush the secession, and obtained legitimacy afterwards when the Government of Katanga ended the bid for secession on January 13th 1963, while in exile in Northern Rhodesia (204).

Legality of Claim

Katanga had no legal basis for its secession. It had not been administered separately from the rest of the D.R.C., nor had it been politically marginalized or antagonized by the central government. Its population did not identify as distinct from the D.R.C. except for a few years leading up to the secession, a point which was itself undermined by a secession from Katanga in the north. The weak legal basis of its claim is reflected in the near uniform opposition it met with from the international community. Lastly, it declared its independence after the D.R.C. had done so and been recognized by other states. As a result, the Katangan declaration of independence violated the territorial integrity and sovereignty of the D.R.C.

Costs, Resources, and External Actors

Local, regional, and international actors opposed the secession of Katanga, denying it recognition or support, except for Belgium in the latter case. Significant mineral resources, in the form of diamonds and precious metals, played a significant role in the push for secession; however their presence failed to either sustain the separatists or induce support for the secession. The conflict did not undermine the external stability of the area, as the costs of the conflict, economically and on a human level remained within the country. Finally, though the subject of sympathy from Europeans, Katanga lacked any type of support beyond that provided by Belgium.

Biafra: Overview

On May 30th, 1967, the Republic of Biafra declared its independence from Nigeria. Issues of identity primarily drove the push for independence, although the possession of significant oil reserves in the territory declaring independence played a role in both the preliminary analysis of separatists and the conflict itself. In contrast to Katanga, Biafra achieved recognition, though it was only to be a handful of states. Beyond recognition, Biafra received limited support from regional and international actors, yet the majority of rhetorical, logistical and material support was for the central government. After 3 years of conflict and famine, Biafran forces surrendered to the Nigerian High Command on January 14th, 1970 (Pavkovic, Radan 96). Waged primarily through traditional warfare and amidst starvation and disease, the cost of the conflict rounded out at an estimated 1,000,000 deaths (96).

Origin of Secession

The origin of the secession rests with the self-reinforcing inter-ethnic conflict between elites for control of the state that then devolved to the ground level where groups of individuals began targeting each other along ethnic lines. Nigeria transitioned into independence as a federal state, with three federal units, each corresponding with an ethnic identity majority, with one outlier (98). Over half of the population of Nigeria could fall into one of these three categories, the interplay of which dominated politics and civil services, making utilization of ethnicity necessary. The Hausa, numbering around 15 million, were predominantly found in the North of the country. The Yoruba, numbering 10 million, lived in the West and Midwest of the country (98). Lastly, and of significance for this study, the Ibo, numbering ten million, resided in the East and Midwest. This placement does not encompass the entirety of individuals who may be perceived or self-identify as belonging to these groups, instead existing spread across the country.

On January 15th 1966, a group of Ibo army majors staged a coup, killing political and military leaders of other ethnicities in the process (98). The coup was thwarted however by Major General Ironsi, an Ibo, who imprisoned the soldiers (98). Ironsi followed this by abolishing the existing divisions between regions of the federation, and integrating regional and federal services (98). In May of the same year, large-scale riots occurred in the North, leading to the massacre of several hundred Ibo (99). Then on the 28th of July, Northern military members launched what they deemed as a countercoup against Ironsi, deposing him and purging eastern officers (99). Facing a splintering central government, regional leaders convened to stabilize the situation, however they failed to reach agreement on a means of resolving the situation. This failure, and the widening violence against easterners as a result of the countercoup, led the eastern government to look towards secession (99).

In the wake of the failed conference, the Northern Army killed between 7,000 and 10,000 Ibos in the North. This prompted a mass exodus of Ibos and other Easterners to the East, estimated to be around 1.5 million. The situation was exacerbated in October when several hundred more Ibo were killed and hundreds of thousands were expelled (Pavkovic, Radan 99). Though the majority of Eastern military and government officials were already convinced of the merits of secession, the declaration of a state of emergency

by the central government and a redrawing of province borders persuaded those who had remained opposed to the idea (100).

Review of Secession

On May 30th, 1967, the Republic of Biafra declared its independence from Nigeria, launching a propaganda campaign centered on ethnic differences between Biafra and Nigeria, beyond calling upon foreign states to provide recognition and support (101). Advocates for Biafra and the Biafran government's Swiss public relations firm waged their campaign by framing the situation through rhetoric, photographs, and footage that echoed the Holocaust and made an obvious association (Baum). The propaganda eventually gained the attention of the foreign media, who quickly made it a major focus for people throughout the world, bringing in unprecedented humanitarian assistance (101). Despite remaining chaos among the central government, state forces occupied the Midwest. This offensive quickly stalled and was overrun by the Biafran Army. In what some historians view as the fatal error of the Biafran secession, the Biafran Army failed to make use of this victory, opting to wait for reinforcements rather than push for the capital region and potential force the central government to consent to the secession. As the central government stabilized itself, it recaptured the Midwest, and came to encircle Biafra and occupy all of its coastal areas by 1968 (101). A stalemate arose due to external aid mitigating the impact of the siege, however improved armament of the state forces and food and medicine shortages eventually led to the collapse of Biafran forces, and occupation of the territory by state forces. Then on January 9th, 1970, Biafran forces surrendered, bringing an end to the conflict (101).

Legality of Claim

From a strict legal perspective, Biafra had no legal basis to secede. It was not administered separately under colonialism, nor did domestic Nigerian law enable such an act. Biafra was ethnically and religiously distinct from the rest of the country, being primarily Ibo and Christian. However, pockets of different groups existed throughout the country, and were often integrated into local communities. The violence that precipitated the secession was significant, as too was that which occurred over the course of the conflict. The value of these factors as evidence of the necessity for a separate state was

undermined by lack of state directed violence, and the unity of other provinces and ethnic groups with the state.

Costs, Resources, and External Actors

On an external level, the O.A.U. opposed the secession, repeatedly trying to bring about peace and calling upon 'the secessionist leaders to cooperate with the Federal authorities for the purpose of restoring peace and unity to Nigeria' (Baum). Though formally opposed, the issue split the O.A.U (Pavkovic, Radan 101). Tanzania, Gabon, Ivory Coast, and Zambia extended recognition of Biafra, acknowledging the need for unity and territorial integrity, but arguing that these goals could not be achieved without consideration for the conditions necessary. Beyond the O.A.U., Haiti and South Africa also recognized Biafra (101).

The state received the majority of material support, being armed and supplied by the U.K., the U.S.S.R. and the U.S., varying from light arms to planes and heavy artillery. Biafra received support indirectly from Gabon, Israel, South Africa, China and France, mainly in the form of light arms (101). On the non-state front, Biafra received support from the International Red Cross, and joint church aid in the form of food and medicine (101). Although the conflict was highly intense, it was relatively short and external instability was limited (Baum).

Somaliland: Overview

On the 18th of May 1991, the Republic of Somaliland declared its independence, repealing the Union Act of 1960, and returning to its original status and borders prior to its federation with Somalia (Paquin 156). After resolving an initial period of internal conflict that lasted until 1997, Somaliland has progressively met and exceeded the qualifications of statehood: possessing an operational government, stable and effective institutions, a functional economy and having held both a referendum on the constitution and multi-party democratic elections (152). However, despite these indicators and a strong legal claim to statehood, it only possesses de facto recognition from other states (Pavkovic, Radan 503).

Origin of Secession

From 1897 until decolonization, British Somaliland existed as a British colony, distinct from Somalia (Paquin 152). On the 26th of June 1960, Somaliland declared its

independence and was recognized by the U.S. and 35 other states (152). Shortly after obtaining independent statehood, Somaliland passed the Union Act of 1960 and joined with Italian Somalia to form a federal republic, which was relatively homogenous ethnically, linguistically and religiously (153). The nature of this union shifted though when the president of Somalia unified the two states by means of a unilateral decree, without basis in a treaty or the consent of Somaliland. This amalgamation was then reinforced through a tyranny of the majority, when a constitution affirming the existence of a single state was put to referendum and supported by Italian Somalis (153).

Review of Secession

Events in Mogadishu altered the political landscape considerably in 1969, when Siad Barre launched a military coup against the state, seizing power, suspending the constitution, dissolving the legislature, and nationalizing private industry (153). Barre initially founded a military regime with the backing of the U.S.S.R.; however, his goal of uniting Somalis within one greater Somali state and consequent invasion of Ethiopia caused the U.S.S.R. to withdraw support, a gap that the U.S. cautiously filled (154).

In the 1980s, lingering dissent and further political marginalization brought about the formation of the Somali National Movement (SNM), which initially sought to overthrow Barre and restore the initial federal state arrangement (154). In 1988, operating out of Ethiopia with the support of the Addis Ababa, the SNM pushed Barre's forces out of important cities, Hargeisa and Burao. The state military regime responded through a campaign of aerial bombardment, focusing heavily on urban areas (Paquin 154). By the close of the 1980s, deaths estimated to have resulted from the conflict exceeded 50,000 (154).

Somaliland was not alone in its conflict against the regime as other groups, notably the United Somali Congress and the Somali Patriotic Front, also waged war to overthrow Barre (155). The weight of these conflicts resulted in a shift in the nature of the secessionist conflict over the course of the 1990s, as the central government lost administrative functionality and military capacity in 1990 and then dissolved shortly after in 1991 (155). Despite having brought about the fall of the Barre regime, the SNM was alienated from efforts to reform a federal state by infighting among other anti-Barre

factions and a sense of exclusion due to the unilateral formation of a new government by the United Somali Congress (Pavkovic, Radan 501).

Following internal reconciliation, the Republic of Somaliland declared its independence on the 18th of May 1991, repealing the Union Act of 1960 that originally incorporated it into a federation with Somalia (501). Citing its previous status as a separate, independent and sovereign state and delineating its borders as those it originally held, Somaliland asserted itself as a case of dissolution, not secession (Paquin 155). However, as no government existing in Mogadishu was considered competent, no consent could be acquired for the dissolution, a condition that led the O.A.U. to oppose recognition (156–7).

Conflict embroiled Somaliland once more however, as loyalists to the state launched attacks, and internal dissent arose due to issues of political representation and efforts to demobilize soldiers (Paquin 160). By 1997, internal stability and relative peace was reacquired, and external stability was finally established, hemming the generation of refugees, a major issue in the past with over 300,000 Somalilanders being displaced to Ethiopia alone in the 1980s (Pavkovic, Radan 501–2). From this point onward, Somaliland has increasingly not only fulfilled the basic requirements of statehood but also demonstrated its functional capacity as a state and government. Beyond improving its economy, multi-party democracy was instated, with a referendum on the constitution that identified Somaliland as separate and independent of Somalia passing (Paquin 161).

Somalia remains unstable and without a competent and capable central government, and in the two decades since its declaration of independence Somaliland remains without de jure recognition by any state (Pavkovic, Radan 502). However, de facto recognition has been achieved with Ethiopia, Eritrea, Djibouti, the U.S. and some other states (Paquin 161).

Legality of Claim

Somaliland holds a relatively strong legal basis for its claim to independence. It argues that its independence is not secessionist in nature but rather an act of dissolution. Though the African Union has remained opposed to this as Somalia has not consented to this dissolution, the forced unification of the two separately recognized states and violation of the legal basis of the federation in doing so supports the argument of

Somaliland. That a significant amount of the international community and major powers, such as the United States, initially recognized Somaliland when it first obtained independence undermines to a degree the notion that the re-extension of recognition would set a precedent for secession. On the issue of territory, the proposed territory of Somaliland is that which it originally was recognized as possessing and this claim does not infringe on the recognized territory of any other state except Somalia.

Costs, Resources, and External Actors

Though Somaliland is relatively isolated from the violence found throughout much of Somalia, the conflict stage left at least 50,000 dead with an estimated 180,000 civilians displaced within Somaliland and 300,000 displaced within neighboring states (Pavkovic, Radan 502). However, the shift in the nature of the conflict, marked by the loss of capacity by the central government to oppose the independence has led to mitigated continued internal and external costs and impact on stability.

Significant resource reserves have not been found to exist in Somaliland as of this time, though exploration rights from the 1980s to parts of its territory remain in the hands of oil companies (Paquin 168). Similar to Eritrea though, Somaliland does have significant coastal access, an important physical feature in the consideration of regional actors and political dynamic (Pavkovic, Radan 503). In part due to this coastal access, Ethiopia has consistently provided support for Somaliland. Egypt, with the backing of Saudi Arabia and other Arab states, has opposed the secession, to stymy Ethiopia (503). Kenya has supported the Somali state due to its investment in the Transitional Government of Somalia (503). Beyond the regional level, South Africa has provided support to Somaliland, and the UK has been involved in pushing for support of recognition (503). On the note of support, Somaliland does possess a Diaspora, however it is relatively small, unorganized and split between different countries (Paquin 160).

Differentiating the Exceptions

Examination of the cases of Eritrea, South Sudan, Biafra, Katanga and Somaliland enable the identification of local, regional and international factors that may explain their differing outcomes and offer insight beyond the specifics of each case. Factors drawn from the case studies are distilled to their basic form in the interest of being transferable

or able to find commonality among variations of details. Factors will be identified at the local, regional and international levels. Each will be discussed in terms of its potential merit for determining the outcome of a secession, and examples from the case studies will be drawn in support of this. These factors will come to form an analytical framework.

Local Level Factors

Separatist movements do not possess the guarantee of survival that states hold (Jackson 3). Separatist movements must not only avoid destruction by the original state and non-state factions, but must also maintain a sufficient level of functionality to continue their bid for statehood.

The first local factor is whether or not the host state is capable of opposing the secession. Where it cannot, the state may be more broadly dysfunctional, and be considered legally incompetent, and incapable of consenting to a secession. In all of the case studies, the host state of the separatist movement is at least initially capable of opposing the secession, or the groups that shift to separatist agendas (it would need to be capable enough to antagonize such groups sufficiently to drive them to secede in the first place). Pierre Englebert's thesis on the deficit of separatism in Sub-Saharan Africa supports this, as he asserts that elites and marginalized groups will generally seek to take control of the state rather than seek to secede, given the difficulty of secession and the relatively valuable nature of control in Africa over extant state institutions (Englebert, Hummel 405).

The second significant local level factor is whether or not the state is able to exercise effective control over the separatist territory. If the host state is able to reassert effective control over the territory in question, this tends to indicate the destruction or crippling of the separatist movement. If not signifying destruction on the battlefield, this may show a denial of resources, and affect the analysis of other actors. As Katanga and Biafra display, when forces representing or acting on behalf of the state took possession of the territory, the claim was withdrawn and the bid ended. However, the Eritrean case shows that if the host state captures the territory through especially destructive tactics, but

fails to gain complete control, the backlash may be greater than the initial opposition. Thus, reestablishing effective control requires time and a relative degree of stability.

As the state and central government have access to state resources and are able to receive logistical and military support from other states and organizations, it tends to be better equipped than the separatists. Yet, as the case of Eritrea demonstrates, this support may backfire if the host state troops are poorly trained or disciplined, and are captured by or defect to the separatist. The capacity of a secession to survive and defeat the forces of the central government relies upon support networks, such as a Diaspora, through which separatists are able to acquire supplies and assistance. For example, significant food aid was provided by religious organizations to separatists during the conflicts in Nigeria and Sudan (Natsios 172). Thus, whether or not a secession possesses a support network is important in determining its longevity and fate.

These factors then tie into the fourth and fifth significant elements, the intensity of the conflict, and its duration. A high intensity conflict incentivizes internal and external actors to directly or indirectly intervene to end the secession, due to the debilitating economic and human cost, as well as the effect the conflict has on the stability of the government and the rest of a state. Protracted conflicts tend to have residual effects that may induce subsequent cycles of violence. The duration of a conflict commonly relates to its intensity (though if a state is unable or unwilling to suppress the separatists then the two may not be related). In both Katanga and Biafra the conflict was relatively short when compared to that of South Sudan and Eritrea.

Regional Level Factors

The effect of a conflict on its surrounding region represents another significant factor in the outcome of a secession. The conflicts in Katanga, Biafra, and Somaliland all have been relatively contained within the area of the host state. In contrast, the Eritrean and South Sudanese cases had large impacts on the stability of the region via generations of refugees, the use of neighboring territory as bases, the necessity of border defense, the loss of trade and economic growth, and other spillover effects.

Regional actors have served key roles in the outcome of a secession through provision of material and political support to either the state or the separatist movement.

Who regional actors support depends on their relations with the host state, and how the outcome of the secession could affect their interests.

International Level Factors

Ultimately, the outcome of a secession is determined at the international level, as the present nature of the international system is such that once recognized, and a member state of the United Nations, the juridical sovereignty of a state cannot be removed. While factors at a local level may determine the longevity of a secessionist movement, and regional factors mitigate or reinforce their effect, without recognition as a sovereign state, the separatists remain arguably unsuccessful. As discussed in section 5-4, the extension of recognition to a separatist entity has potential domestic and international political and legal repercussions. Through the case studies, three international factors appear to have an influential role in the outcome of a secession.

The most fundamental factor is whether or not the separatist entity fulfills the basic requirements of statehood, as without these no recognition could be extended. Whether or not a secession has a legal claim to independence is also a significant factor. Different variations of a legal claim to secession appear in the cases of South Sudan, Eritrea, and Somaliland. Finally, other states and non-state actors may be driven by benefits that successful secession would bring about; thus whether or not the separatist region possesses resources is a potentially significant factor.

Applying the Framework: Trends and Findings

When viewed together, we are left with the following list, with the answers being yes or no unless indicated otherwise:

Domestic Factors

1) Host State Capacity   (Host Capacity)
2) Exercise of Effective Control by State Over Secessionist Territory
                                                                (Effective Control)
3) Presence of a Support Network for Separatists   (Support Network)
4) Intensity of Civil War   (Conflict Intensity)
    a. High   b. Low   c. None
5) Duration of Civil War   (Conflict Duration)
    a. Long   b. Short   c. None

Regional Factors

    1) Conflict generated regional external instability   (External Instability)

    2) Regional Actors Support of Secession   (Regional Support)

        a. Yes   b. No   c. Mixed

International Factors

    1) Does Secession Fulfill the Basic Requirements of Statehood? (Statehood Req.)

    2) Does Secession have Legal Claim?   (Legal Claim)

    3) Does Separatist Territory Possess Resources?   (Resources)

        By constructing this list into an investigative framework, they can be applied to all of the case studies:

Determinants of Secession Outcome

Factors	Eritrea	South Sudan	Somaliland	Katanga	Biafra
Host Capacity	Yes	Yes	No	Yes	Yes
Effective Control	No	No	No	Yes	Yes
Support Network	Yes	Yes	Yes	Yes	Yes
Conflict Intensity	High	High	None	High	High
Conflict Duration	Long	Long	None	Short	Short
External Instability	Yes	Yes	No	No	No
Regional Support	Yes	Yes	Mixed	No	Mixed
Statehood Requirements	Yes	Yes	Yes	Yes	Yes
Legal Claim	Yes	Yes	Yes	No	No
Resources	Yes*	Yes	Yes*	Yes	Yes

* Indicates non-traditional resources, specifically coastal access and the potential existence of material resources.

Research Findings

On the domestic level, where the host state was able to exercise effective control over the separatist territory, the secession failed. In all cases except Somaliland, the host state was capable of responding to the secession. It should be noted that when the SNM pushed for reform of the state, the Somali central government was able to react. The lack of state capacity may have enabled Somaliland to continue to exist, however it may also explain its lack of success, as the central government is also regarded as legally incompetent. The presence of a support network, though varying in form and level of strength, does not appear to act as any particular indicator of the outcome of a secessionist movement.

Cross-analysis of the intensity and duration of a secessionist conflict however, reveals a correlation between instances of long, high intensity conflict and successful cases of secession in Eritrea and South Sudan. Katanga and Biafra both generated high intensity conflicts but only for short periods of time, before being eliminated. Somaliland represents an exception, as Somalia has lacked the capacity to oppose it for the duration of its declared independence. On the regional level, extending from the nature of the civil war, whether or not a secessionist conflict generates external instability also seems to be a determinant. Though the separatist conflict in all of the cases had an impact on the citizens of the host state and the population residing in the separatist territory, Eritrea and South Sudan by nature of their duration and intensity caused significant regional instability due to spillover conflict, refugees, and other detrimental effects. Somalia itself may still generate regional instability; however its internal struggles are distinct from that caused by Somaliland.

Regional support does not prove to be of any significance in and of itself. In all of the cases the O.A.U. and the A.U. has maintained opposition to the secession, only altering its official line after recognition is achieved or after the host state has consented to the secession. In both of the exceptions to the norm, no regional actor has extended recognition without an internal process allowing for a legal secession. Where regional support has been aligned in support of the secession, tracking the support given is a difficult task. However, insofar as it indicates the lack of support for the host state by its neighbors, regional support will reinforce domestic factors.

Whether or not a secessionist group is able to fulfill the requirements of statehood does not offer much insight. Given the highly politicized nature of recognition in relation to the requirements of statehood, this is to be expected. That being said, the exceptional qualifications of Somaliland appear to be mitigated by other factors, potentially the lack of regional or international costs for maintaining the status quo. The presence of a legal claim does appear to relate to the outcome of a secession, however Somaliland stands out here, having existed as a recognized sovereign state before being illegally absorbed into Somalia. The disparity between the strong legal claim of Somaliland and its present lack of statehood also seem to point to idea that there is insufficient incentive to make the cost of non-recognition exceed the cost of recognition. Both Biafra and Katanga lacked legal claims, which undermined their efforts to obtain recognition.

Surprisingly, the presence and nature of resources are not, in and of themselves a strong predictive or determining element. Both Eritrea and South Sudan possess resources (coastal access and oil reserves respectively). However, the presence of similar coastal access in Somaliland, an important incentive for nearby landlocked countries, undermines non-material resources. On the material side, both Katanga and Biafra enclosed major mineral and oil reserves respectively, yet both bids proved unsuccessful. The difference between Katanga, Biafra, and South Sudan may rest with their ability to survive. In other words, the extension of recognition by states that would benefit was withheld due to the possibility of failure. Coastal access is also more dependent on the current political dynamic between regional actors, whereas minerals and oil require stability and sovereignty to be extracted.

A Defense of the Research Method

The methodology used to identify these factors and in turn formulate a hypothesis, is open to criticism—particularly concerns of confirmation bias, or circular reasoning. However, this thesis seeks to address a deficit present in current literature on the subject, and identify factors that potentially differentiate cases of secession, not to explicitly prove one above the others. Undoubtedly, the creation of a theory involves a degree of proving the potential validity of the idea, but it is not the same as to prove it superior to others, an area where such confirmation bias could undermine the integrity of the findings (van Evera 68, 71). These factors are drawn from across the five case studies

and then reapplied across all of them, satisfying the need to test the antecedent conditions involved (73).

Cases of secession may be rare but they have important implications and effects on the areas where they occur. Research on the subject is critical for future policy evaluation of secessions or even internal conflict. A better understanding of which factors are important, and how they interact and affect the actors involved is necessary to accurately judging a situation, and constructing solutions.

Conclusion

The principles of sovereignty and territorial integrity delimit the boundaries of the international state system and are mechanisms for maintaining stability. Secession enables an asymmetrical challenge of these tenets: it pits the stability of the international system against the stability of individual states. The international system acts in the interest of its own survival, which is apparent in the evolution of international law and its capacity to address novel or exceptional threats. Reflecting this reality, it follows that the intensity, duration, and external impact of a secessionist conflict are the primary determinants of the outcome of a secession. The legal viability of a secession should serve as a central factor. However, as comparison of Eritrea and South Sudan with Somaliland showed, the strength of a legal claim is dependent on the cost of continued conflict outweighing the domestic and international consequences of secession in the perspective of external state actors. Beyond their immediate cost, and in contrast to the cases of Katanga and Biafra, Eritrea and South Sudan demonstrated both intractability and a high probability of continued conflict at a similar or greater cost.

The opposition of the international community to the practice of secession except in the most extreme conditions may discourage potential secessionists and ensure a general measure of stability, but it has debilitated the development of consensus on the conditions of acceptable secession, and thus minor stability is broadly achieved at the cost of major conflict and selective instability. The reluctance of the international community to embrace secession is understandable; the relations with existing states and their allies are more valuable than those with a non-state actor. Efforts to determine the

conditions of secession provoke fears of a slippery slope. This is significant for those countries with separatists of their own or those that are more dependent on their juridical statehood than their own power to uphold their sovereignty. Yet, as the two exceptions have shown, the international system may be forced to come to terms with the necessity of a secession when the cost is sufficient and the consent of the host state is given, so long as it is judged to be legally competent.

As it stands, the necessity of a high-cost conflict and consent of the host state reflect the operating mentality of the international system and the complexity of the issue at hand. Waiting for these conditions to occur is not a viable means of ending or preventing instability and conflict. This is because the way a secession ends affects the future of participants on all sides (Tir 716–719). As the conflict continues and worsens, human and economic costs rise and grievances distill throughout the population, and the future of the region is threatened.

Concern does not rest simply with the actions of the general population, but their vulnerability to war rhetoric by their leaders. According to Jaroslav Tir, the leaders of the new state and the host state both have incentives to return to war (716–719). The leaders of the host state may seek to regain lost territory in order to regain control over natural resources, historical purposes, and/or to shore up domestic support (715). The leaders of the new state may seek to gain even more territory, as they may have only gained some of the territory they had claimed (719). The restart of conflict post-secession is a serious problem, as each side is legally able to receive logistical and material aid, and direct support from other states (Shaw 1042–3).

Interdependent economic structures, an expanding regime on the existence of human rights, and the potential threat of non-state actors through acts of terrorism and other means of non-traditional warfare necessitate a vigilant role for the international community in order to maintain stability. Failure to address cases of secession until the separatist movement displays its ability to survive enables states to temporarily circumvent potential infringements on the sovereignty of others. Yet, the ignorance of the international community to the legal basis of the secession, specifically Somaliland, undermines the delicate effort to strike a necessary balance.

The present lack of procedure and conditions for secession may discourage some actors by nature of the difficulty of the process, but can encourage others, as the myth of creating homogenous states persists, and the right to self-determination continues to be considered a generalized right to secede (Johnson 143–148). Determining the requisite conditions for a secession is a necessary step if secession continues to persist in the future, and until the right to self-determination no longer appears to be a generalized right to secede, it will.

Granting statehood to Somaliland would engrain basic acceptable conditions for secession, and show a differentiation between secessions with a legal basis and those without. This initial act would offer a stepping-stone for the system to reconcile the matter before it proves more debilitating, and discourage the more opportunity-seeking separatists.

Given the value of sovereignty and the potential for exclusion in any domestic power arrangement, separatist movements will persist in their efforts. That being said, it is in the interest of the international system to develop a functional framework for acceptable secession, without which intractable highly damaging civil wars that bring about the consent of the host state will remain the path to statehood, setting up the conditions for state failure.

Works Cited

Barth, Frederik. "Segmentary Opposition and the Theory of Games: A Study of Pathan Organization." *The Journal of the Royal Anthropological Institute of Great Britain and* Ireland vol. 89 no. 1 (1959): 5–21. Print.

Bates, Robert. *State Versus Ethnic Claims: African Policy Dilemmas*, ed. Donald Rothchild and Victor Olunsorola. Boulder: Westview Press, 1982. Print.

Baum, Phil. "The Tragedy of Biafra." *Jewish American Congress*. Web. http://www.biafraland.com/1968_Biafra%20_Report_By_American_Jewish_Congress.htm.

BBC. "Q&A: South Sudan Independence." July 4, 2011. Web. http://www.bbc.co.uk/news/world-africa-12111730

Bookman, Milica Z. *The Economics of Secession*. New York: St. Martin's Press, 1992. Print.

Cali, Basak. *International Law for International Relations*. New York: Oxford University Press, 2010. Print.

Carter, Johnson. "Partitioning to Peace: Sovereignty, Demography, Ethnic Civil Wars." *International Security* 32, no. 4 (2008): 140–170. Print.

Englebert, Pierre. *Africa's Sovereign Sorrow: The Politics of Poverty*. Youtube. April 4th 2007; Middlebury, VT: Middlebury College Guest Lecture.

Englebert, Pierre and Rebecca Hummel. "Let's Stick Together: Understanding Africa's Secessionist Deficit." *African Affairs* vol 104, Issue 416 (2005): 399–427. Print.

Evera, Stephen Van. *Guide to Methods for Students of Political Science*. Ithaca: Cornell University Press, 1997. Print.

Hannum, Hurst. *Autonomy, Sovereignty, and Self-Determination: The Accommodation of Conflicting Rights*. Philadelphia: University of Pennsylvania Press, 1990. Print.

---. "The Specter of Secession: Responding to Claims for Ethnic Self-Determination." *Foreign Affairs* March/April 1998 (1998): 13–18. Print.

Hannum, Hurst and Eileen F. Babbitt, ed. *Negotiating Self-Determination*. Lanham: Lexington Books, 2006. Print.

Herbst, Jeffrey. *States and Power in Africa*. Princeton: Princeton University Press, 2000. Print.

Jackson, Robert H. *Quasi-States: Sovereignty, International Relations and the Third World*. Cambridge: Cambridge University Press, 1993. Print.

Johnson, Carter. "Partitioning to Peace: Sovereignty, Demography & Ethnic Civil Wars." *International Security* vol 32. No. 4 (Spring 2008): 140–170. Print.

Krasner, Stephen D. "Structural Causes and Regime Consequences: Regimes as Intervening Variables." *International Organization* vol. 36, no. 2 (Spring 1982): 185–205.Web. http://www.jstor.org/stable/2706520.

Gleditsch, Nils P., Peter Wallensteen, Mikael Eriksson, Margareta Sollenberg, and Havard Strand. "Armed Conflict 1946–2001: A New Dataset." *Journal of Peace Research* 39, no. 5 (2002): 615–37. Print.

Green, Elliot D. "On the Endogeneity of Ethnic Secessionist Groups." PhD diss., London School of Economics, 2005.

Montevideo Convention on Rights and Duties of States. *Convention on Rights and Duties of States*. Washington, DC: International Treaty, 1934. Web. http://avalon.law.yale.edu/20th_century/intam03.asp.

Nabudere, D.W. "The Political Economy of Conflict and War in the Great Lakes Region." *Institute for Justice and Reconciliation* 1 (2004): 1–60. Print.

Natsios, Andrew S. *Sudan, South Sudan, and Darfur: What Everyone Needs to Know*. New York: Oxford University Press, 2012. Print.

Okhonmina, Stephen. "States Without Borders: Westphalia Territoriality under Threat." *Journal of Social Science* 24 no. 3 (2010): 177–182. Print.

Organization of African Unity. *Resolutions Adopted by the First Ordinary Session of the Assembly of Heads of State and Government held in Cairo, UAR, from 17 to 21 July 1964*. Addis Ababa: Secretariat, 1964. Print.

Palus, Nancy. "Deal Between Mali Tauregs, Islamists Breaks Apart." *Voice of America*, June 1, 2012. Web. http://www.voanews.com/content/deal-between-mali-tuaregs-islamists-breaks-apart/1146062.html.

Paquin, Jonathan. *A Stability-Seeking Power: U.S. Foreign Policy and Secessionist Conflicts*. Montreal: McGill-Queen's University Press, 2010. Print.

Pavkovic, Aleksandar, and Peter Radan, ed. *On the Way to Statehood: Secession and Globalisation*. Burlington: Ashgate Publishing Company, 2008. Print.

Pavkovic, Aleksandar, and Peter Radan. *Creating New States: Theory and Practice of Secession*. Burlington: Ashgate Publishing Company, 2007. Print.

Posner, Daniel N. *Institutions and Ethnic Politics in Africa*. New York: Cambridge University Press, 2005. Print.

Ratner, Steven R. "Drawing a Better Line: Uti Possidetis and the Borders of New States." *American Society of International Law* vol. 90 no. 4 (1996): 590–624. Print.

Rothchild, Donald, and Victor Olunsorola, ed. *State Versus Ethnic Claims: African Policy Dilemmas*. Boulder: Westview Press, 1982. Print.

Ross, Michael. "What Do We Know about Natural Resources and Civil War?" *Journal of Peace Research* vol. 41, no. 3 (May 2004): 337–356. Print.

Shaw, Malcolm N. *International Law*. 5th ed. Cambridge, UK: Cambridge University Press, 2003. Print.

Southall, Aidan. "State Formation in Africa." *Annual Review of Anthropology* vol. 3 (1974): 153–165. Print.

Spears, Ian S. *Civil War in African States: The Search for Security*. Boulder: First Forum Press, 2010. Print.

Tir, Jaroslav. "Keeping the Peace after Secession: Territorial Conflicts between Rump & Secessionist States." *The Journal of Conflict Resolution* 49, no. 5 (2005): 713–741. Print.

Trivett, Vincent. "Oil-Rich South Sudan has Hours to Choose Between North Sudan, China and The U.S." *Business Insider*, July 8, 2011. Web. http://www.businessinsider.com/southern-sudan-independence-2011-7.

United Nations. *U.N. Security Council Resolution* 1265, 1999. New York: U.N., 1999.

—. *U.N. Security Council Resolution 1973*, 2011. New York: U.N., 2011.

United Nations High Commissioner for Refugees. "Sudan," 2012. http://www.unhcr.org/pages/49e483b76.html.

Scottish Economic Development in the face of English Hegemony: Trade Imbalances, Banking, and the Union of 1707

Thomas Reubens

The economic fortunes of Scotland on the eve of the Union of 1707 were highly influenced by Scotland's relationship with England. Armed conflicts, most notably the English Civil War and the Glorious Revolution had devastating effects on the Scottish population and economy. Farmers suffered greatly from the pillaging of soldiers while severe losses of manpower could potentially produce economic woes for an agrarian society like Scotland. The fields of farmers who were off fighting or dead would go untended, resulting in food shortages and, in the most extreme circumstances, famines. Meanwhile Scotland was lagging behind in the acquisition and development of colonial possessions.[1] The eventful years before the Union saw Scotland directly confronting major challenges of economic development, including chronic deflation and English hegemony.

Disregarding the 1603 Union of the Crowns, the English Navigation Acts of 1660 and 1663 labeled Scotland a foreign country, thereby excluding Scottish ships from trade with the English colonies. However, trade records from between 1660 and 1707 show that the attempts to bar the Scots from the plantation trade were not always successful. In fact, trade between Scotland and the colonies generally increased throughout the second half of the 17th century despite the Navigation Acts. To England's frustration, this legislation was often disregarded by many Scottish and Colonial merchants because both parties stood to benefit from this trade.[2] After the Union, the Navigation Acts were expanded to protect *British* shipping, rather than only *English* shipping. This post-Union expansion of the Navigation Acts was to eventually prove beneficial to Scotland.[3]

---

[1] Francis Russell Hart. *The Disaster of Darien: The Story of the Scots Settlement and the Causes of its Failure 1699–1701*. (Cambridge, MA: Riverside Press, 1929), 11–2.
[2] Henry Hamilton. *Economic History of Scotland in the Eighteenth Century.* (London: Oxford University Press, 1963), 250.
[3] Hamilton, 255.

Nonetheless, the trade between the colonies and Scotland was not automatically profitable for Scotland as a whole. Wealthy Scottish landowners had fallen into the habit of purchasing luxury goods produced overseas; this led to a scarcity of financial capital, or deflation in Scotland as gold and silver were exported and foreign goods were imported. This trade deficit and the resulting shortage of capital were becoming increasingly troublesome for the economy of Scotland, causing the Parliament and economically minded Scots to seek a remedy for these problems. Far from being unique to Scotland, this economic trend reflected a common challenge among contemporary European countries at the dawn of the Industrial Revolution. In 1681, the Scottish Parliament passed legislation designed to improve the manufacturing sector of the economy and reduce the trade deficit. *A Proclamation for regulating and encouraging Trade and Manufactories in the Kingdom of Scotland* preceded this legislation and succinctly described the economic situation of Scotland in 1681:

> …by the undue balance of trade, occasioned chiefly by the import of many unnecessary and superfluous commodities, consumed upon vanity and luxury, a great part of the stock of the money of the kingdom was exported, and the improvement of the native Export and Manufacture of the Kingdom neglected.[4]

The subsequent legislation called for ship building materials to be duty free, a tax reduction on new stocks for manufactures in order to increase investments, and the removal of excise & customs taxes on dyes for the textile industry. Furthermore, the "Act for Encouraging Trade and Manufactures" sought to attract foreign investment and skills by offering legal protection to foreign entrepreneurs and citizenship to foreign Protestants who established manufactures in the country.[5] The aim of this was to increase investments in the local textile industry and to import foreign innovations in order to make Scottish manufactures more competitive.

However, there were still substantial economic obstacles ahead of a modernizing Scotland. Although foreign linen industries could threaten to out-compete the nascent Scottish industry, new protective tariffs had the adverse effect of causing foreign countries to close their markets to the Scots. At a time when the linen industry was the biggest source of employment

---

[4] *A Proclamation for regulating and encouraging Trade and Manufactories in the Kingdom of Scotland.* (Edinburgh: Printed by the Heir of Andrew Anderson, 11 April 1681), 4.

[5] *Proclamation for regulating and encouraging Trade and Manufactories,* 6–9.

after agriculture and a third of Scotland's exports were cloth, such tariff wars could be painful.[6] England was the biggest foreign market for the Scottish linen industry, so a tariff war could easily hurt the linen industry, but conversely the tariffs protected the Scottish linen industry from English competition. From 1681 and continuing beyond the Union of 1707, Scotland was consistently trying to improve this crucial industry by importing new techniques from the continent, particularly from the Dutch.[7] Scotland's other industries, notably coal, iron, fishing, and agriculture faced developmental challenges similar to those faced by the linen industry.

There were three identifiable factors that contributed to the deflation facing Scotland. The first, mentioned previously, was the trade deficit under which Scotland had been operating. The remedy sought for the trade deficit was to limit imports and develop the local economy, particularly the textile industry, so that exports would increase. This action by the Scottish Parliament in 1681 led to a backlash from certain foreign markets, such as the English market for linen.[8] The chosen solution to fix the trade deficit and therefore to indirectly fix the deflation problem caused a new problem: a deficiency of markets for Scottish goods. In a terrible twist of irony, the solution conceived by Scotland for opening new markets, the Darien Company, more formally known as the Company of Scotland Trading to Africa and the Indies actually aggravated the deflation problem by tying up £400,000 in shares. The Darien Company became the second contributing factor to the deflation problem faced by Scotland.[9]

The Bank of Scotland's first governor, John Holland estimated the number of "adventurers" in the Darien Scheme to be 1,500.[10] The amounts contributed ranged from the minimum of £100 to the maximum of £3,000 and came from titled men and untitled men from all parts of the country.[11] The transfer of money from all over Scotland in exchange for stock in the Darien Company aggravated deflation because stock is not nearly as liquid as money and

[6] Alastair J. Durie. *The Scottish Linen Industry in the Eighteenth Century.* (Edinburgh: John Donald Publishers LTD, 1979), 13.
Hamilton, 131.
[7] Durie, 14.
[8] Durie, 8.
[9] Hamilton, 291, 295.
[10] John Holland. *A Short Discourse on the Present Temper of the Nation with respect to the Indian and African Company; and of the Bank of Scotland. Also, of Mr. Paterson's pretended Fund of Credit.* (Edinburgh: Printed by John Reid, 1696), 10.
[11] *A Perfect list of the several Persons Resident in Scotland who have subscribed as adventurers in the joint-stock of the Company of Scotland Trading to Africa and the Indies together with the respective sums which they have severally subscribed in the books of the said company, amounting in the whole to the sum of £400,000 sterling.* Edinburgh: printed and sold by the Heirs and successors of Andrew Anderson, 1696.

therefore does not circulate like money. Although the Darien Company did spend some of that money back into circulation, it did not spend it all in Scotland.

The Company of Scotland was formed with the purpose of usurping the East India Company's trade monopoly. It was originally going to be funded by both English and Scottish investors; £300,000 from each nation for a total of £600,000, until the English Parliament had the company relocate to Edinburgh and the burden of funding the endeavor (now only £400,000) was shifted entirely to Scotland.[12] This move by the English was consistent with what Andrew Fletcher of Saltoun referred to as the "affair of Hamburgh," in which a minister of King William III in Holland persuaded the Dutch against purchasing stock in the Darien Company.[13] One of the early warning signs that may shed light on England's decision to back out of the new trading company was a decline in the East India Company's stock values when the new company formed.[14] Independent of the exact reasons for England's treatment of the Company of Scotland, the result was the same. The financial burden placed on Scotland for the formation of the company was increased but perhaps this was a contributing factor to the strong opinion in favor of the Darien Company in Scotland, since it made the company more exclusively Scottish.

The blame for the failure of the Company of Scotland was attributed to King William and the English Parliament by many Scots.[15] Their grievances against the English were not unfounded. William was in a predicament after the Glorious Revolution; he was the King of two independent parliaments who both had economic goals that could become contradictory. Furthermore, the ability of both the Scottish and the English Parliaments to pursue their own policies was increased because of compromises such as the abolition of the Lords of the Articles, a royal institution that had restricted the Scottish Parliament before the Glorious Revolution.[16] The English solution to the unwanted trading company was to refuse it aid until it collapsed.[17] This was despite a clause in the Darien Company's charter on behalf of William III stating that "His Majestie's officers Civil and Military, and all others, are hereby required in their respective

---

[12] Hart, 20, 183.
George Pratt Insh. *The Darien Scheme* (London: Staples Press Limited, 1947), 4, 10–2.
[13] Andrew Fletcher of Saltoun. *Selected Political Writings and Speeches of Andrew Fletcher of Saltoun.* Edited by David Daiches. (Edinburgh: Scottish Academic Press, 1979), 32–3.
[14] Insh, *The Darien Scheme*, 10.
[15] Hart, 162.
[16] Insh, *The Darien Scheme,* 5.
[17] George Pratt Insh. *The Scottish Jacobite Movement* (Edinburgh: Moray Press, 1952), 76–7.

stations, to be Aiding and Assisting the said company.[18]" The reason given by William for neglecting the trading company was that it would "disturb the general peace of Christendom,[19]" but he was also known to have complained that "I have been ill-served in Scotland; but I hope some remedies may be found to prevent the inconveniences which may arise from this Act.[20]" It is true that the Darien Company could have "disturbed the general peace of Christendom," because the location for the company's new colony—the Gulf of Darien—was in territory previously claimed by Catholic Spain on the Isthmus of Panama. However, William's reasoning here was quite hypocritical; William himself disturbed the peace of Christendom when he usurped the throne from James II and started a war. Additionally, a letter from Caledonia [the colony on the Gulf of Darien] to Boston dated 18 February 1698, referred to an attempt made by the Scots to contact the President of Spanish Panama for a peace bargain.[21] As evidenced by the Spanish attack on the colony and its ultimate surrender to the Spanish, this professed peace deal came to nothing.[22]

As previously mentioned, one of the primary reasons for the English attitude towards the Darien Company was a backlash from the East India Company and its patrons, although there were other reasons such as Captain Long's "Letter from Jamaica, concerning the Scots settling in Caledonia, 1699." Captain Long had expressed fears to the English Parliament about the impact that the settlement at Caledonia would have on the Jamaican colony. In particular, Long feared that a prosperous Scottish colony at Darien, especially if it were the result of a gold rush would lead to a labor shortage in Jamaica as men left to try their fortunes at Darien.[23] Lieutenant Governor of Jamaica William Beeston reiterated this point.[24]

However, the failure of the Darien Company cannot be attributed entirely to the English. When the legislation was passed to create the Company of Scotland Trading to Africa and the Indies, there was no reference to the Gulf of Darien or a plan to establish a colony in Spanish territory, so perhaps the founders of the company had not planned that far ahead. The company's

---

[18] "Act of Parliament constituting the Company of Scotland, Trading to Africa and the Indies." Edinburgh, 26 June 1695 in Hart, Francis Russell. *The Disaster of Darien: The Story of the Scots Settlement and the Causes of its Failure 1699–1701*. (Cambridge, MA: Riverside Press, 1929), 190.

[19] Insh, *The Darien Scheme,* 20.

[20] Hart, 34.

[21] Hart, 238.

[22] Insh, *The Darien Scheme*, 19–20.

[23] Hart, 219–22.

[24] Hart, 152.

charter forbade the inhabitation of an area by force; this however created a gray area because when the Scots settled at Darien, their immediate neighbors were not the Spaniards who had claimed the territory, but the Native Americans with whom the Scottish colonists had successfully established peace.[25] Nonetheless, a backlash from Spanish colonial authorities was far from inconceivable. Fortunately the company's charter did contain provisions permitting the colonists at Darien to make treaties and to defend themselves, their settlements and their trade.[26] The Company of Scotland was also burdened with other problems, although it can be argued that English aid could have helped resolve many of them. Provision shortages and illnesses, in addition to the immense length of time that it took to reinforce the expedition from Scotland, all proved to have severe consequences for the adventurers.[27]

The news of the Darien Company's failure was not passively received in Scotland. Various writers began to circulate accounts of its failure and the reaction from William III was negative, as indicated by his edict against petitioning grievances for Darien and bounties for the capture of some pamphlet writers.[28] In 1700, George Ridpath published "Scotland's Grievances, Relating to Darien, &co. Humbly offered to the consideration of the Parliament" which began with a reference to an anecdote of William Wallace and Robert the Bruce with very nationalist overtones.[29] The evocation of Scottish heroes from its medieval war for independence in the 14th century is just another indicator of the attitude that at least some Scots held at this time. Ridpath's publication defined and explained how the English Parliament, acting in concert with King William III was treacherous to the Scots. Their treachery alienated the relationship between the Scottish subjects and their King by causing the King to act against the interest of his subjects in Scotland.[30] Furthermore, the English Parliament did not consult the Council of Scotland and King William never consulted the Scottish Parliament before acting against the Darien Company.[31] One of Ridpath's chief protests against this "treachery" was that even if William's actions had been for the greater good of his realm,

---

[25] Hart, 45, 165, 224–5.
[26] Hart, 186.
[27] Insh, *The Darien Scheme*, 19.
[28] George Ridpath. "Scotland's Grievances, Relating to Darien, &c. Humbly offered to the consideration of the Parliament." (Edinburgh, 13 February 1700), 12.
[29] Ridpath, preface.
[30] Ridpath, 1, 6.
[31] Ridpath, 5, 12.

It will by no means acquit the pernicious counselors of treachery towards us, since the least they could have advised in this case was, that we should have had notice of such Proclamations before-hand, that we might have been upon our guard, and have done what we could to have prevented our colonies being frightened or starved from Darien.[32]

Unfortunately Scottish grievances manifested themselves in more than just ink. In reaction to the East India Company's seizure of the Scottish trading ship *Annandale*, the English trading vessel *Worcester* was apprehended by a Scottish crowd and three of its crew, including one Captain Green were hanged on the shores of Leith on the Firth of Forth.[33] If we can trust these events as indicators of a strong public sentiment in favor of the Company of Scotland, which seems reasonable, then we can acknowledge a sort of *economic* nationalism occurring in Scotland at this time. This economic nationalism was centered on the Company of Scotland and its goal of establishing foreign markets for Scotland's industries, in order to improve the fortunes of the nation. Subsequently, the failure and eventual dissolution of the Company of Scotland in 1707 was an undeniable blow to this economic nationalism.

Finally, the third factor contributing to the deflation problem was the reliance on precious metals to serve the function of money. The supply of financial capital at this time was vastly limited by the amount of precious metals in the country. The only traditional methods for ameliorating this predicament were to either fix the trade deficit, debase the currency, or to gain access to a new gold or silver mine. These solutions were generally making poor progress, so the Scottish Parliament with the aid of English financier John Holland, passed legislation to establish the Bank of Scotland on 17 July 1695. The shares for the new banking company, which was to have a twenty-one year monopoly in Scotland, were sold in London and Edinburgh.[34] The founding of the Bank of Scotland in 1695 followed very quickly after the establishment of the Bank of England in 1694 and the general principles of both were "to be employed in the trade of lending and borrowing money upon interest and in no other.[35]"

The new banking innovations—a national bank issuing notes of credit—were regulated by the government so that people would have a greater trust in the banks & their notes while also

---

[32] Ridpath, 5.
[33] Insh, *The Darien Scheme,* 4.
[34] Hamilton, 294.
[35] Hamilton, 294.

being protected from the "oppressive extortion" of goldsmiths and other private lenders.[36] The Bank of England accelerated this process by replacing the old tally stick system with their notes and ascribing a common seal to their notes in order to improve their circulation.[37] The development of a central bank, regulated by the government and initially funded via the sale of stocks to private investors who would then make money from the interest paid on bank loans, played a critical role in solving the deflation problem by circulating paper notes representing gold or silver alongside gold and silver.

> The more Credit the Bank has, and the more Money is lodged in it, the more it will lessen interest, for want of Occasions to improve it; and *those who Lodge their Money in the Bank have it as much at their disposal as if it were in the hands of the Goldsmiths, or in their own Cash-Chest, and there is a greater Value than the Money which is deposited in the Bank, that circulates by their Credit as much as if it were stirring in Specie*: And the Bank Bills serve already for Returns and Exchange to and from the remotest parts of the Kingdom, and will in a little time do the like in Foreign Parts, which will lessen the exporting of Bullion for the paying and maintaining our Armies abroad during this War; and if the Bulk of the Money of the Nation which has been Lodged with the Goldsmiths, had been deposited in the Bank 4 or 5 years past, it had prevented its being so Scandalously Clipped…[38]

Eventually the amount of money circulating in note form exceeded the actual amount of precious metals held at the banks. Due to its relationship with Parliament, the Bank of England, whose example was closely followed by the Bank of Scotland, was also advertised as being a more secure place to keep one's savings than in the vaults of independent goldsmiths or an unrestrained corporation.[39] This, alongside the 1707 recoinage, helped to increase deposits at the bank, thereby increasing the amount of money available to be lent out and subsequently decreasing the interest charged on loans by making them less risky for the bank since more borrowers meant more interest being paid back to the bank. This banking innovation solved the

---

[36] Michael Godfrey. "A short account of the Bank of England," (London: John Whitlock, 1695), 8.
[37] John Broughton. "Remarks upon the Bank of England, with regard more especially to our trade and government. Occasion'd by the present discourse concerning the intended prolongation of the Bank. Humbly address'd to the honourable House of Commons. By a merchant of London, and a true lover of our constitution." (London: A. Baldwin, 1705), 8–9.
[38] Godfrey. "A short account of the Bank of England," 2–3.
[39] Godfrey. "A short account of the Bank of England," 4–5.

deflation problem via artificially inflating the supply of gold & silver in the economy by increasing the amount of credit in circulation in the form of paper notes.[40] The Bank of Scotland was given permission to raise £120,000 Scots from £10,000 sterling by way of issuing credit.[41]

Contemporary London merchant John Broughton lamented these new innovations for a handful of reasons, including the obvious potential for corruption because bank members were simultaneously allowed to sit in Parliament.[42] He was suspicious of the strong efforts to prolong the existence of the Bank of England since it was originally formed under the conditions of wartime necessity.[43] Broughton was also strongly critical of the ability of the bank to lend money to the government at interest due to the immense power that comes with being the creditor of a country's government.[44] Adding to the new bank's profitability was its ability to also lend money and issue notes to the rest of England at interest.[45] Broughton feared "that all these consequences are chargeable upon that boundless Power which the Bank has of extending a Credit so current as this is, which, in its Nature, will always be increasing, until it grows too great for all opposition.[46]" However, despite these drawbacks, the Bank of England and the Bank of Scotland[47] still served their primary purpose of remedying deflation:

> The Power to extend their Credit, and upon so good a Foundation as the security of an Act of Parliament, is perhaps a more considerable Article of their Profit than even so great an Interest. They had a sufficient Prospect (and Time has made it good) of raising their Credit to a Par with Money; and wherever such Credit obtains, it affords all the real Advantages of so much Money. And a Credit thus Establish'd, and rais'd to Par with Money, is capable of being increas'd to an immense Value; considering the great Occasions for it, and Conveniences of it in Trade, which, its known, cannot be carry'd on to a due Extent in England, without *a far greater Sum in Credit, than there is in Specie in the whole Nation*: And therefore a Credit vastly extended, must bring in Vast Profits to them that are thus Credited.[48]

---

[40] Hamilton, 294–7, Broughton, 13–5.

[41] Hamilton, 295.

[42] Broughton, 10.

[43] Broughton, 11.

[44] Broughton, 15, 28–32.

[45] Broughton, 13.

[46] Broughton, 36.

[47] Note: Only the Bank of England lent to the crown whereas both could lend money to private individuals.

[48] Broughton, 13–4.

Not every contemporary of Broughton shared his sense of alarm and skepticism with the power being given to the Bank of England and, to a lesser degree, the Bank of Scotland. *A brief account of the intended Bank of England* by William Paterson discussed some different consequences of the central bank.

> …when this War [the Glorious Revolution & subsequent war with France] begun, the Credit of the Nation was low, and the Wits on both sides, found no better nor honester way to supply the Necessities of the Government, than by enhauncing the Price and Interest of Money; the effect of which was, that the Government was obliged to pay from double to treble, or higher interest: The Disease [debt] growing daily worse, Men were tempted to draw their Effects from Trade and Improvements, and found the best and securest Gain, in making Merchandise of the Government and Nation.[49]

The belief espoused here by Paterson was that the growing debt would lead to economic improvements and advancements, which is in accordance with the consistently growing national debt of Great Britain during the 18th century. Despite the economic growth of the 18th century, the national debt and consequently the taxes of Great Britain steadily rose.[50]

Other consequences of the new banking institution revealed themselves during the panics surrounding various Jacobite Risings. Jacobite scares usually led to bank runs, such as that seen in 1715, but the banks did not have enough coinage to convert everybody's paper notes. This situation led to the establishment of interest-bearing accounts. When banks went insolvent because a Jacobite scare caused a bank run, those left with bank notes were promised interest. Eventually this practice came to be used during stable banking periods as well.

Historian Henry Hamilton wisely remarked that capital and skill were the two ingredients holding back most sectors of the Scottish economy on the eve of the Union, but between new banking innovations initiated in the 1690s and the gradual development of domestic industries, economic growth in Scotland was indeed a reality before the Union and, less surprisingly, after the Union.[51] Yet domestic growth cannot always occur in a bubble and in this instance it was subjected to many factors. The available supply of financial capital, the availability of markets, competition for burgeoning markets and the threat of war, present in this case by the looming

---

[49] William Paterson. "A brief account of the intended Bank of England," (London: Randal Taylor, 1694), 3–4.

[50] John Brewer, *The Sinews of Power: War, Money and the English State, 1688–1783.* (London: Century Hutchinson, 1988), 30, 39, 96–7.

[51] Hamilton, 296.

succession crisis of Queen Anne's reign, combined to create a fragile atmosphere for the continuation of economic growth. Although it cannot be known how Scotland would have developed economically if the Union never happened, it is clear that the initiative for economic improvement happened domestically and was underway before the Union. Nonetheless, the Union was still a catalyst for economic growth because new markets became available for Scottish goods, the Union bolstered the prospects of stability and peace by resolving the succession crisis, and payment of the "equivalent" by the English in return for the dissolution of the Darien Company provided a helpful boost of financial capital to the Scottish economy.[52] The growth of economic nationalism in Scotland during the 1690s may have taken a blow when the Darien Company failed, but Scottish efforts to improve economically did not. Ultimately the Union proved beneficial to Scotland, but the impetus for Scottish modernization came from within Scotland.

---

[52] George Pryde, ed. *The Treaty of Union of Scotland and England 1707*. (London & Edinburgh: Thomas Nelson and sons LTD, 1950), 93–4.

Works Cited

Brewer, John. *The Sinews of Power: War, Money and the English State, 1688–1783.* London: Century Hutchinson, 1988.

Broughton, John. "Remarks upon the Bank of England, with regard more especially to our trade and government. Occasion'd by the present discourse concerning the intended prolongation of the Bank. Humbly address'd to the honourable House of Commons. By a merchant of London, and a true lover of our constitution." London: A. Baldwin, 1705, in *Bank of England: selected tracts 1694–1804*. Farnborough: Gregg Press, 1968.

Daiches, David. *The Paradox of Scottish Culture*. London: Oxford University Press, 1964.

Durie, Alastair J. *The Scottish Linen Industry in the Eighteenth Century.* Edinburgh: John Donald Publishers LTD, 1979.

Fletcher, Andrew, of Saltoun. *Selected Political Writings and Speeches of Andrew Fletcher of Saltoun.* Edited by David Daiches. Edinburgh: Scottish Academic Press, 1979.

Hamilton, Henry. *Economic History of Scotland in the Eighteenth Century.* London: Oxford University Press, 1963.

Hart, Francis Russell. *The Disaster of Darien: The Story of the Scots Settlement and the Causes of its Failure 1699–1701*. Cambridge, MA: Riverside Press, 1929.

Holland, John. *A Short Discourse on the Present Temper of the Nation with Respect to the Indian and African Company, and of the Bank of Scotland Also of Mr. Paterson's Pretended Fund of Credit.* Edinburgh: Printed by John Reid and sold by Mrs. Beiglie, 1696.

Godfrey, Michael. "A short account of the Bank of England," London: John Whitlock, 1695, in *Bank of England: Selected Tracts 1694–1804.* Farnborough: Gregg Press, 1968.

Insh, George Pratt. *The Darien Scheme.* London: Staples Press Limited, 1947.

Insh, George Pratt. *The Scottish Jacobite Movement.* Edinburgh: Moray Press, 1952.

Pagan, William. *The Birthplace and Parentage of William Paterson Founder of the Bank of England, and Projector of the Darien Scheme*. Edinburgh: William P. Nimmo, 1865.

Paterson, William. "A brief account of the intended Bank of England," London: Randal Taylor, 1694, in *Bank of England: Selected Tracts 1694–1804.* Farnborough: Gregg Press, 1968.

A Perfect list of the several Persons Resident in Scotland who have subscribed as adventurers in the joint-stock of the Company of Scotland Trading to Africa and the Indies together with the respective sums which they have severally subscribed in the books of the said company, amounting in the whole to the sum of £400,000 sterling. Edinburgh: printed and sold by the Heirs and successors of Andrew Anderson, 1696.

A Proclamation for regulating and encouraging Trade and Manufactories in the Kingdom of Scotland. Edinburgh: Printed by the Heir of Andrew Anderson, 11 April 1681.

Pryde, George, ed. *The Treaty of Union of Scotland and England 1707*. London & Edinburgh: Thomas Nelson and Sons LTD, 1950.

Ridpath, George. "Scotland's Grievances, Relating to Darien, &c. Humbly offered to the consideration of the Parliament." Edinburgh, 13 February 1700.

The Cultivation of Uncertainty: An Atheist's Experiences With Agnosticism

Kim Hall

"The Cultivation of Uncertainty: An Atheist's Experiences With Agnosticism." That wasn't supposed to be the title of my presentation today. My intended title was going to be something like "The role of VMAT2 neurotransmitter proteins in self-transcendental experiences in epileptic patients." This would have been a seemingly more appropriate topic for a biology student interested in genetics, as well as an inquisitive but hardcore atheist (one who does not believe in a god or gods) anxious to understand a strange phenomenon called faith. With this skeptical, secular mindset and my scientific background, the questions I was interested in for this capstone project were about the material and objective bases for religious belief. I wanted to figure out what was going on in what I thought of as the real world—the material, physical, observable world—that made some people believe that there was something else. Reading *The God Gene: How Faith Is Hardwired Into Our Genes*, by Dean Hamer, inspired my original research topic. Hamer hypothesized that the levels of a particular neurotransmitter transport protein in the brain could affect feelings of self-transcendence, a mysterious feeling that formed the base of spiritual beliefs. I decided to vet this hypothesis by comparing it to the other primary literature on the subject, and then discuss the philosophical implications of a material basis (or lack thereof) for belief in God. I found a willing professor who liked talking about both biology and philosophy and happily started my project by sifting through psychology, sociology, and neurology articles, trying to get a good understanding of what the scientific community meant when it discussed things like "spirituality," "faith," "self-transcendence," and "belief."

Rather quickly this approach went to pieces. The papers I was reading could not settle on a consistent definition of any of these terms. Spirituality was defined in terms of self-transcendence and vice versa. Qualities like peace of mind and mental well-being were found to correlate with a spiritual mindset, but then people with greater spiritual tendencies were found to have increased mental well-being. The tests and scales scientists used to measure traits like belief in people were widely varied and none were given much credibility. I found this somewhat

frustrating—I had set out with what I thought was a clear question to ask and my research was not returning any answers. The scientific method places a lot of emphasis on finding a good question and phrasing it correctly in order to obtain a valid answer, so I thought the solution to my problem would be to come up with a new, better question. I returned to *The God Gene* looking for other terms to investigate and other hypotheses to evaluate, only to find that none of the ideas I was interested in pursuing went anywhere in the primary literature. In fact, I found more poor reviews of the science in Hamer's book than anything else.

For a long time, my project stagnated as I tried to find an appropriate question that would yield a clear answer. I was loath to give up the pursuit of faith and equally reluctant to abandon the possibility of a biological explanation for it, but my research indicated more and more firmly that I could not examine both. The scientific community simply did not have a clear idea of what spirituality was, much less a reliable way to observe or measure it. Dr. Everett endured a lot of increasingly frantic scenes where I related my difficulties in finding a question to ask of this confused and uncertain community. Each time, he suggested not rebelling against this uncertainty, but stepping back, observing the hazy terms and cloudy uncertainties for what they were, and cultivating my own uncertainty. Letting go of the simple question, simple answer approach was a challenge for me, and it took me a long time to accept the idea that there was a phenomenon surrounded by many very complicated questions without simple answers from the realm of science. With encouragement, however, I managed to get some distance from this mindset, and toward the end of the fall semester I had concluded that my original project was just not going to work out.

I knew I was still interested in understanding spirituality and religious faith, but my trusted and familiar approach to problem-solving had failed me. At Dr. Everett's suggestion I decided to try something radically different, and started simply talking to everyone who would listen about spirituality and their experiences with it. I was still trying to get a handle on definitions at this point, still hoping that with enough clear terms I could ask a good question with an answer to be found in the library (but you already know how this turned out). The atheists I spoke with talked about physical sensations experienced in the presence of natural beauty, but the Catholic students I talked to were much less specific when describing their faith. Many were taken aback by my questions about feelings of self-transcendence and physical reactions, instead comparing their belief to something like breathing, unconscious but vital. I felt

I was on to something after these discussions—the faithful do have something different going on! there is something to investigate!—but I still was not sure what it was.

Then another professor pointed me toward a book by Huston Smith entitled *Why Religion Matters: The Fate of the Human Spirit in an Age of Disbelief.* Smith, a prominent scholar of the world's religions, made a case I had never before encountered for the reasons to have spiritual faith. He said that the mindset of the scientific community had warped from a tradition of explaining natural phenomena to a cult-like worship of the infallibility of science, a cult he termed "scientism." By worshipping science and attempting to use it to explain all questions, Smith claimed, the practitioners of scientism were flattening the world and arrogantly robbing it of any higher morality, purpose, or meaning. Science had a very important place in the quest to understand the world, but it was dangerous to make it the sole method of inquiry. Smith then advocated the retention of religious faith as a way to understand the world in terms that may not be as clear and articulate as those in science, but nonetheless true and valid.

This was hard for me to accept or even evaluate fairly. The idea of coming to a valid conclusion without using the tool set I had carefully acquired throughout my education in the sciences made me want to throw my copy of Smith's book against the wall more than once. But his argument was persuasive. His description of the cult-like devotion to science held by so many ontological naturalists sounded familiar, as I recalled my angst over not finding answers in the scientific literature and my crumbling certainty that an answer to the right question was there. While I remained skeptical that religious faith was a valid method for describing and comprehending the world, I had to accept that I was subscribing to the narrow and dogmatic principles of scientism myself, and that I had to step away from this devotion if I wanted to think critically about spiritual belief.

Abandoning scientism was uncomfortable. Without certainty in science's ability to explain everything in the world, I was grappling with questions that were new to me but familiar to any epistemology seminar—how do we know anything? How do we know we know it? What does it mean to "know?" How much can we know about anything? Specifically to my project, how do the spiritually faithful know what they do if they don't seem to be using the methods I use to know things? This was a frightening question for several reasons—had I been going about learning things the wrong way during my entire education? Was there something out there to understand that I wasn't able to see, much less comprehend and appreciate? How was I going to

make a project out of a question like this? But with patient encouragement from my mentor and from those I spoke with about spirituality, my fears subsided. I remembered that my uncertainty was something to be nurtured, cultivated, if my questions were to be articulated. Gradually, I recalled the original questions that prompted my research into the science of spirituality. I was a curious atheist who wanted to understand faith, who had started off by trying to understand it in the only way I knew. Dr. Everett and others persuaded me that this was a topic worth reading and writing about. The questions I settled on addressing from a confused atheist point of view were these: what is spiritual faith, why do some people have it and some lack it, how do we know who is right, and which kind of person am I?

Since the definition of faith was not to be found in scientific literature and not to be articulated by the student believers I spoke with, I started reexamining writings by people who sounded like they had a good, certain definition of faith and were not afraid of relating it. Richard Dawkins' *The God Delusion* and Christopher Hitchens' *God Is Not Great: How Religion Poisons Everything* are anything but reluctant to expound upon the wicked characteristics of faith. These "new atheist" works frame religious belief as a kind of rejection of logical processes, driven by a sickness of the mind or some primitive artifact of evolution from our distant pasts. Anything that leads people to believe in phenomena without "sufficient evidence" must be some kind of perversion and evidence of gullibility, of an inferior mind, or of dangerous, evil lunacy. Hitchens asserts the difference between this insanity and the more rational, nobler atheistic mindset thus: "Our belief is not a belief. Our principles are not a faith. We do not rely solely upon science and reason, because these are necessary rather than sufficient factors, but we distrust anything that contradicts science or outrages reason. We may differ on many things, but what we respect is free inquiry, open-mindedness, and the pursuit of ideas for their own sake."

These words, in light of Huston Smith's allegations of scientism, seem to me to no longer uphold principles of free inquiry or open-mindedness. The tactic of the new atheists is not one that encourages open and free discussion; rather, it ridicules and demonizes those it purports to be engaging in dialogue. Dawkins, Hitchens, and their ilk single-mindedly attack religion and defend the scientific worldview, without considering the possibility that the faithful have a valid reason for being so, and without considering the idea that their method of understanding the universe might not be the only reasonable method—or, indeed, might have some evils of its own.

As I reread the new atheists' works, I realized that this scientism and the men who promote it sound almost indistinguishable from fundamentalist religious sects and their members—our way is the one true way, and the others who dissent are sinners, demons, and the damned.

Feeling disgusted with how long I had counted myself among such a dogmatic group of atheists, I switched to reading books by more philosophers and scholars of religion, something it is pointedly obvious Dawkins and Hitchens do not do. This was where I finally found some claims about religious faith that made sense to me. Unlike the new atheists, scholars like John F. Haught and Terry Eagleton seemed to me to be much more willing and able to uphold principles of open-mindedness. They were well acquainted with the arguments and grievances of the new atheists and were in perfect agreement with them on some points, such as past and present wrongs committed by people of faith and in the name of faith, and the need for a rational basis for holding beliefs. Also unlike the new atheists, these scholars argued that these points in no way decreased the validity of the reasons for having religious belief, and they made the (what seemed to me at the time) extraordinary claim that faith is something that must be arrived at with reason and logic. People of faith are not encouraged to disavow their rational principles; in fact, these believers encourage the use of reason to give meaning to belief. Eagleton uses an analogy of falling in love—if I say you love someone, I have to back it up with reasons why if the love is to be meaningful. Yet another person could agree with all the reasons—the pretty color of my beloved's eyes, for example—and not feel the same love. That love, greater than the sum of the reasons or evidence, is like religious faith. I had a hard time buying this but all of these authors pointed out how belief in scientific claims, or any kind of claim, works much the same way. I believe in evolution, I have many good and rational reasons to believe in evolution, for example, but the combination of those reasons alone is insufficient to justify my belief. There is a commitment and trust in the methods of science that I have no objectively measurable evidence for, and this trust is what makes my belief a conviction.

The allegation is made by the new atheists that people who believe are altered and deficient in some way. They do not possess all the facts or knowledge that the skeptics do, or their capacities for logic are diminished, or (as I at one point suspected) they are physiologically different from skeptics. I have come to believe that this is neither a rational conclusion nor a fair one. What evidence do people have that believers are sick or stupid beyond their faith? When a scholar like Terry Eagleton makes an elegant analogy for his beliefs that even an undergraduate

atheist like me can comprehend, what reason do I have to think he is a lunatic? It now seems to me not only hasty but ridiculous to write off an individual as irrational simply because they happen to believe something I do not when all other evidence points toward their rationality, a conclusion I am deeply embarrassed not to have applied to the debate between religion and atheism before.

If the proponents of religion and the proponents of atheism both seem to be rational, why do they end up at different conclusions? Good science is supposed to be replicable; identical procedures are supposed to yield identical results. A believer and an atheist may indeed be using the same sorts of logical processes and identical types of evidence to arrive at conclusions, but as Theodore Everett points out in "The Rationality of Science and the Rationality of Faith," the content of that evidence is somewhat different. The faithful and the skeptical alike have trust in people whose convictions they rely upon, like parents, teachers, experts, or other persons of authority. When surrounded by trusted people who have faith, people who are trusted for good reasons (like a history of being right about other things they assert, or for being more likely to be right about a given subject), it would seemingly be irrational to suddenly buck the trend and start believing something different. The religious, the spiritual, the skeptics, and the scientists do exactly the same thing. They have a lot of good evidence for what they believe and they draw rational conclusions from it, but most of it is not directly observed and objectively measured data. For believers, the evidence may be based on the authority of a parent or priest who is trusted to be more knowledgeable. For scientists, it is based on faith in other scientists' observations (theoretically replicable, but who has time or funding to check all the observations?), which are based on trust in other scientists. Freethinking, a rejection of trusted authority, may not always be the most rational choice as the new atheists claim it is.

So how do we know who is right?

I do not know. We all have our different convictions for reasons both rational and ineffable. We are mostly reasonable people, mostly not deficient or perverted lunatics. Beyond a trust I have built up in a certain set of authority figures, a trust that has eroded as my research has progressed, I do not seem to have a sufficient body of evidence to decide the matter. I have no idea if anyone does.

Faced with the rational behavior of believers, the dogmatic principles of the new atheists, the failure of my familiar scientific approach to resolve a question, and my inability to pick a

side in the debate between religion and atheism, I now feel I can no longer define myself as an atheist. Letting go of this label and belief set has been harder by far than coming up with a topic for this project. Often I feel like I am rejecting a skepticism I worked hard to develop and defend for years against its detractors, and I worry that I myself might be succumbing to the beautiful delusions the new atheists scoff at. Then again, I fear that I am failing to understand the truth about how the world really works, failing to comprehend what lies beyond the physical phenomena scientists like me try to measure. I do not know if I will ever resolve these doubts and questions. I do not know if it is even possible for anyone to resolve them. I cannot even call myself a strict agnostic, one who steadfastly maintains that the answers to these questions and the nature of the spiritual cannot possibly be known. I have cultivated my uncertainty to the point where I know a few more things that I do not understand, and they are all I can present to you today.

I owe many thanks to my mentor, Dr. Theodore Everett, for his encouragement and patience, to Drs. Herzman and Nicodemi for their feedback and support, and to everyone who shared with me their beliefs and convictions.

The Economic Implications of Working-Class Authoritarianism: A Comparative Perspective

Steve Fisher

ABSTRACT

Seymour Martin Lipset first proposed his theory of working-class authoritarianism in his 1959 paper "Democracy and Working-Class Authoritarianism." His thesis held that individuals of lower social and educational classes were more prone to sympathize with authoritarian and illiberal regimes, be they to the political left or right. This paper will discuss and then explore Lipset's theory in a modern context. In doing so, it will argue that working-class authoritarianism still exists in liberal democracies today. To do so, case studies of several states and their elections will be analyzed and compared. These examples will show that while the phenomenon can still be seen today, the manner in which it appears has changed following the fall of communism. This paper will conclude its discussion of Lipset's theory by arguing that the authoritarian tendencies of the classes Lipset indicated now lean exclusively to the political right. Furthermore, the economic implications of this theory in a modern context will be illuminated. Finally, the argument will be made that working-class authoritarianism has the effect of preventing individuals affected by it from voting for the liberal political policies that could aid them in their economic struggle.

❖

Since the beginnings of democracy, citizens' voting habits have been a great deal of interest to political scientists around the world. In 1959, Seymour Martin Lipset made his contribution with an article titled "Democracy and Working-Class Authoritarianism." In this paper he argues that individuals of lower social strata are more likely to support authoritarian regimes than their counterparts of middle or upper socioeconomic class. Furthermore, Lipset correlated this to both past fascist regimes and also to communism in the (then) present and the future. As mentioned, the majority of his arguments referenced (either directly or otherwise) historical instances of fascist governments shortly after their demise. Furthermore, he made note of since-fallen communist systems as the opposite extreme. With neither of these types of

illiberal regimes surviving in the sense Lipset knew them to be at the time of his writing, the validity of his thesis is worthy of re-examination in a modern context.

Lipset's use of data almost exclusively from a time in history directly succeeding the fall of authoritarian regimes to the political right and the rise of them to the left raises the possibility that given the mitigation of instantiated political extremes to both the left and right, working-class authoritarianism has been dealt a crippling blow. In addition to the above theory, this paper evaluates the economic implications of working-class authoritarianism in order to see what effects the concept has on the ability of people to vote toward economic policies that are favorable to them and their low economic standing. Examining both Lipset's work and modern-day case studies, it is apparent that working-class authoritarianism is, in fact, still a relevant phenomenon in liberal democracies throughout the world and that it is directly derogatory to the economic well-being of the lower-class individuals affected by it.

In discussing the core tenants of working-class authoritarianism, Lipset claims, "A variety of evidence from many countries suggests that low status and low education predispose individuals to favor extremist, intolerant, and transvaluational forms of political and religious behavior" (Lipset 482). Taking this assertion a step further, Lipset asserts that these past trends can be used to "anticipate wide-spread support by lower-class individuals and groups for extremist movements" (483). Prior to discussing the present-day applications of Lipset's argument, it is necessary to first discuss and analyze the argument itself.

Lipset bolsters his thesis with statistics from several world democracies. In particular, he cites a study completed in post–World War II Germany in which 3,000 German men were asked what their preferred governmental system would be. These men were then given the choice between a single-party system, a multiparty system, and a system absent of political parties (Lipset 485). This test was designed to determine the relative strength of authoritarian and illiberal opinions within a particular occupational group. An answer of "one party" would indicate a preference for an authoritarian style of governance, given the monopoly of national politics this party would attain. On the opposite side, an answer of "several parties" would imply a preference towards liberal government. These results would then be compared by occupational group, revealing relative feelings of authoritarianism or liberalism.

The results, as reproduced in Figure 1, revealed that men who worked as unskilled and semi-skilled workers were significantly more likely to support the idea of a single-party

governmental structure than their counterparts who identified as artisans or civil servants, among others. Lipset further explains that in addition to this 1953 study, "comparable results were obtained in 1958 when a similar question was asked of national or regional samples in Austria, Japan, Brazil, Canada, Mexico, the Netherlands, Belgium, Italy, and France" (485).

A second study cited by Lipset was conducted on almost 5,000 Americans in 1955. The study asked participants a series of questions regarding the rights of free speech for different groups including communists, critics of religion, and several others.[1] Based upon their responses to these prompts, the participants were grouped into three categories of relative "tolerance." These categories were considered to be "more tolerant," "less tolerant," and a middle ground referred to in the study as "in-between" (Lipset 486). These results were then arranged by the occupation of the test subject. Lipset examined the concluding statistics of the study and found that farmers and manual laborers had a 20 percent and 30 percent chance of being members of the "more tolerant" group, respectively. This data was compared to the 66 percent of subjects labeled as being professional or semi-professional workers who tested as being "more tolerant" (Lipset 1959, 486). The results of this study are reproduced in this paper as Table 2.

Lipset also indicates that other factors can explain the increased levels of authoritarianism found in lower-class individuals throughout the world. These include religion, personal upbringing, and education. Of education, Lipset argues that an increase in levels of education has a direct correlation to an increase in liberal attitudes. As evidence for this, Lipset cites a study finding that in both the United States and Germany the likelihood of an individual to support a system with multiple parties increased as the individual's education increased (Lipset 489).

In making this argument, Lipset states something that few today would refute, claiming that a lack of "… formal education, itself closely correlated with social and economic status is also highly correlated with undemocratic attitudes" (Lipset 489).[2] Lipset is positing the notion that education and income levels combine to form a sort of infinite regress in society. To expand on this, further consider the following: If one is educated then they are more likely to lead a life

---

[1] These groups were chosen because they would be seen as heretical to the "average" American's predispositions at the time.

[2] Lipset is using the term "undemocratic" in the way that I use the terms "illiberal" and/or "authoritarian," that is, to mean a system which lacks civil rights and personal freedom, not necessarily a system which fails to hold elections. I submit that the two are not mutually exclusive, a "liberal" system with personal freedoms and civil rights may exist absent of democratic elections and, conversely, a system can hold elections and thus be democratic without the aforementioned traits of a "liberal" system. For more discussion of the distinction see Chapter 2 of Fareed Zakaria's *The Future of Freedom: Illiberal Democracy at Home and Abroad*.

on a higher economic level than an equivalent individual with no education. This will all but certainly lead to an increased opportunity for the former's children to be educated, thus continuing this cycle of education. If the opposite approach is taken, however, and an individual with no education and thus, a lower economic class is assumed, it is easy to see the difficulty in financing an education for either themselves or their children, perpetuating a lineage of the uneducated lower class. In this way it is difficult in all but the most exceptional cases, to make a firm distinction between economic class and education level. This is certainly part of the reason for Lipset's conclusion that "both inferior education and low occupational position are highly intercorrelated, both are part of the complex making up low-status, and are associated with a lack of tolerance" (Lipset 490).

The first modern case study to be evaluated for congruence with working-class authoritarianism is that of the United Kingdom. The British system of liberal democracy is parliamentary in nature, a factor that leads to a system in which third parties have the potential to gain representation within Parliament. As a result of this feature the government can be more volatile than in more strict two-party systems like that of the United States.[3] With that being said, the majority of British politics do remain dominated by a center-right and center-left party.

In the United Kingdom general elections data from 2005, one sees at first glance little indication of working-class authoritarianism. Indeed none of the three top vote-getting parties, the Labour Party, the Conservative Party, and the Liberal Democrats, can be characterized as authoritarian or radical in any regard. They are, in fact, best described as center-left, center-right, and centrist respectively, with all demonstrating strong commitments to liberalism. Furthermore, statistics support the notion that the economic lower class is voting for the left-leaning party. This aligns directly with what one would expect to be in the economic interests of that group. In the 2005 election, the Labour Party earned 48 percent of the vote from the lowest class studied, and 40 percent from the second lowest. This can be compared to the center-right Conservative Party's percentages of 25 percent and 33 percent, respectively, or the politically moderate Liberal Democrats with 18 percent and 19 percent (Tetteh 31).

This data strongly indicates that working-class authoritarianism does not have a strong hold on the voting in the United Kingdom. Accordingly it can be concluded that, on a national

---

[3] A more detailed comparison of the impact of presidential and parliamentary democracy will be given later in this paper.

level, it does not affect the economic concerns of the lower classes, which still vote in large numbers for the center-left Labour Party. However, simply because working-class authoritarianism in the UK is not perceivable on this level does not mean it is completely extinct. On a smaller scale evidence does suggest that authoritarian opinions still maintain a hold on the British lower classes.

To find proof of this one has to only look as far as the radical right British National Party (BNP). Recently this party garnered over 500,000 votes, making it the most popular extreme right-wing party in British electoral history with interest in the party "growing sharply" (Ford). While these votes were not enough to win a seat in Parliament, they were enough to show an alarming demographic trend. As Robert Ford explains, "The BNP does best among older, white working class men with few educational qualifications living in the North of England" (Ford). Statistics further support the notion that a label of "working class" is the best indicator of support for the BNP, with a 70 percent chance of a BNP voter fitting this description (Ford). This is a group that generally favors the center-left Labour Party.[4] What is perhaps more telling is that it is a perfect match to the portrait painted by Lipset's thesis. He had posited that individuals of poor economic and educational standing would be more tempted to support an authoritarian party or governmental system than a more wealthy or educated individual, all other factors being equal. This fact, combined with the meteoric rise of the BNP, indicates that though small in size, working-class authoritarianism seems to be on the rise in the UK.

One factor from Lipset's thesis that is not present is a sizeable contingent from the authoritarian left as well as right. Lipset does not specifically argue that working-class authoritarianism will favor either fascism or other illiberalism on the right or communism on the left, only that it will be an illiberal regime. This is not so in the study of Britain, where lower-class individuals who have shown themselves as sympathetic towards authoritarianism seem to gravitate to one end of the political spectrum. While the radical right BNP was able to win 1.9 percent of total votes in the 2010 election, the Communist Party of Britain (certainly a left-wing authoritarian group) was only able to take 0.38 percent of the vote in the general election ("Communist Party of Britain…"). Further damaging to the notion that working-class

---

[4] Ford's label of "working class" is very much akin to the DE and C2 labels used by Tetteh. Tetteh's terms come from the social scale used by the National Readership Survey (NRS) to indicate unskilled individuals (D) and those at the very edges of economic sustenance (E) and skilled manual labors (C2), respectively. Both authors are referring to members of the lowest income-earning economic group.

authoritarianism could easily cleave to either the political left or right are the membership numbers of both parties. While the BNP reports a membership of 10,256 in 2010, the Communist Party of Britain can only claim 931, or less than 10 percent of the BNP's membership ("Communist Party of Britain…"; "British National Party…"). Clearly the authoritarian working class of Britain is gravitating toward the political right rather than the left.

The second state to be examined is Australia. Australia, like the United Kingdom, features a liberal, parliamentary system of democracy. As one would then perhaps expect, voting trends and behavior in Australia seem to mirror those of the UK in every way discussed above. Like the UK, there is a strong majority of lower-class individuals voting for the center-left party, while their upper-class counterparts tend to vote center-right. Simon Jackman, who has analyzed Australian election results from 1987–2001, indicates that in the case of Australia, this center-left party favored by the poor is known as the Australian Labor Party (ALP) and has won the majority vote of the self-defined "working-class" during this period. As for the center-right Liberal Party of Australia (LP), the same study shows that it has been favored by the economically wealthy across the same period of time (Jackman 16).

One of the findings of this Australian case study is in regards to illiberal parties. As with the mainstream liberal ALP and LP, these illiberal parties also follow trends established by the UK case study. Election results from 2010 indicate that while working-class authoritarianism does not have a crippling hold on Australian politics, it does exist. One of the foremost authoritarian parties in Australia is known as One Nation. This party, based out of the Australian state of Queensland, is characterized by its radical right-wing political stances. Damon Alexander's work "Modelling the Socioeconomic and Demographic Characteristics of One Nation's Vote" examines the voting base of the party. This study concludes that earners in excess of $1500 are "decidedly less likely to support One Nation than those earning less than $699…" (Alexander 9). Lipset's thesis is once again confirmed by these results; the majority of support for an illiberal political party in a democracy is coming from lower-income classes.

While Alexander's work illuminates the voting trends for One Nation, on the political far right, it does not discuss One Nation in comparison to far-left authoritarian parties. Results from the most recent Australian Federal elections (held in 2010) indicate that Australian working-class authoritarianism also leans disproportionately towards the radical right. One Nation, the illiberal right-wing example party discussed above, gathered 27,184 popular votes for 0.2 percent overall.

This is significantly greater than the figures reported for the illiberal left-wing Communist
Alliance, which received 656 votes. So small was the voter support for the Communist Alliance
that they received a statistical zero percent of the popular vote in this election ("Party Totals").
Interestingly, as in the case of the BNP, One Nation enjoys much of its support from one
geographic region, in this case the state in which it is headquartered in, Queensland (Alexander
2).

The final country to provide a case study is the United States. Also a liberal democracy,
the United States differs from the previous case studies in the means by which its democratic
actions are carried out. The United States relies upon a presidential electoral system rather than a
parliamentary system. This has resulted in the emergence of two strong parties with little to no
practical chance of a viable third party emerging and earning seats in Congress in the foreseeable
future, as they frequently do in the Parliaments of the UK and Australia. In the United States, the
two main political parties are the center-left Democratic Party and the center-right Republican
Party.

Despite the aforementioned difference, statistics indicate a high degree of similarity
between the US and the previous countries discussed. *New York Times* exit poll data confirms
that the majority of individuals with family income equal to or below \$49,999 have voted
Democrat since as far back as 1992 (the first year in which this data was compiled). In contrast,
those who make \$100,000 or over per family have statistically voted Republican, with the
exception of the 2008 election in which results were split 49 percent to each party ("National
Exit Polls Table"). This data confirms that, as in the previous case studies, members of lower
social strata tend to vote to the left while wealthier members of society vote for right leaning
parties.

What is, however, different about the United States is the lack of a "dominant" right-wing
authoritarian party. Whereas in the UK and Australia a lower-class individual sympathizing with
authoritarianism can find and vote for right-wing illiberal parties like the BNP and One Nation,
the political structure of the United States is arranged in such a way that no true, viable, third-
party exists, let alone one characterized by right-wing illiberalism.[5] Two of the most illiberal
right-wing parties in US politics, the United States Nazi Party and the American Fascist

---

[5] The term "viable" is used in this context to indicate a third-party with a realistic capability of winning a seat in
Congress as many third parties in Parliamentary governments like the UK and Australia can. This idea is expanded
upon later in this paper.

Movement, do not publicize membership statistics. Furthermore, neither party indicates its intention to sponsor their own candidate to run for national office (though the American Fascist Movement has seemingly endorsed Ron Paul for President), a stark contrast to the national ambitions of the BNP and One Nation ("American Fascist Movement…"). As far as the illiberal left in the United States, it appears to be just as irrelevant as elsewhere in the world. The Communist Party USA, like the US Nazi Party and American Fascist Movement, has not entered its own candidate in any recent elections, preferring instead to endorse President Barack Obama in both 2008 and 2012 (Webb).

Despite the lack of instantiated working-class authoritarianism with an independently functioning illiberal right-wing party, lower-class individuals can still be seen voting against trends for right-wing parties in the United States. Similar to the UK and Australia, this also occurs in certain geographic locations; in this case, the region in question is the US South. Exit polls from the 2004 Presidential election indicate that the Republican Party was favored in the South by a margin of 58 percent to 42 percent, the largest discrepancy between parties of any US region ("US President/National/Exit Poll 2004"). This trend continued in 2008 when the South was the only region to vote Republican with a greater frequency than Democrat ("Election Center 2008: Exit Polls"). This pattern can be traced back as far as 1980, with the Southern vote either favoring Republican or being split with Democrats in every election ("National Exit Polls Table"). Perhaps coincidently, the American South is also statistically the poorest region of the United States, a title that recent data indicates it is putting distance on other regions for (Velasco). This evidence of geopolitical cleavages, in which the lower class from one particular region votes in a manner apart from the same class elsewhere, in indicative of a working-class authoritarian sentiment within United States politics.

When taking the totality of the case studies together, it becomes clear that working-class authoritarianism is in fact, still relevant in modern-day democracies. Though these authoritarian sympathies seem to have shifted along the political spectrum since the time Lipset introduced the theory, it nonetheless holds true: certain groups of people of a relatively low socioeconomic class are indeed more prone to sympathize with authoritarian regimes. Another finding that is perhaps more surprising is that working-class authoritarianism seems to have a greater practical effect in parliamentary democracies. This is due to the higher propensity for third parties to earn seats in parliament in either the United Kingdom or Australia than for one to win a seat in Congress in

the United States. This results in a scenario in which a system at face value may seem to be more representative and "democratic," is actually much more prone to allowing power to authoritarian parties, thus creating an overall more illiberal regime. In the case of the United States, the incredibly rigid two-party structure that exists forces would-be radical right or left voters closer to the political center, robbing authoritarian parties of potential votes. This notion explains the lack of a viable radical right third-party in the United States, as was apparent in the UK and Australia, despite the presence of the seemingly prerequisite geographically concentrated lower class present in all three of the countries.

Given the modern prevalence of working-class authoritarianism, it is valuable to discuss its economic implications upon those affected by it. The relationship between the liberalism or illiberalism of a regime and the economic ramifications of that distinction were of particular interest to political scientist Barrington Moore. In his 1966 book *Social Origins of Dictatorship and Democracy: Lord and Peasant in the Making of the Modern World*, Moore explores the relationship between economics and liberalism in governments around the world. One of Moore's conclusions holds that liberal democracies are largely the result of lower class individuals gaining more wealth and eventually becoming part of an expanding middle class. This newly created middle class will then go about pressing for political rights. This process leads, according to Moore, inevitably to an increase in liberalism, culminating in the establishment of a liberal democracy (Moore 1966). In making this claim, Moore discovered a positive correlation between liberalism and economic growth, resulting in quality-of-life gains for the lower and middle classes.

On the opposite end of the political spectrum, Moore argues that illiberal and authoritarian regimes stifle the economic growth of the lower classes, seemingly with a goal of preventing them from creating a middle class. To Moore, a right-leaning authoritarian regime results from changes from the economic top, as opposed to the middle, as would happen in a liberal political shift (Moore 1966). This upper class then imposes authoritarian measures to solidify its control over the country in question. Unsurprisingly, this is done at the expense of the classes below it. In this manner Moore makes another positive correlation, this time between illiberal, authoritarian regimes and lower social, economic, and political status for those of the lower class. This results in an overall lower standard of living for members of this lower class.

With Moore demonstrating that favorable economic circumstances (more specifically, the creation of a large middle class) implies an increase in political liberalism, what does this mean for the economic implications of modern working-class authoritarianism? Simple *modus ponens* indicates that if beneficial economic circumstances to the lower and middle classes of society imply an increase in liberalism, then an illiberal system would lead to harmful economic circumstances. This asserts that working-class authoritarianism, by leading people to favor authoritarian regimes, is also leading them to support a system that is damaging to their own economic best interests.

The theory of working-class authoritarianism presented in this work is very much a product of its time. Lipset's work was written shortly after the conclusion of World War II, in which three of the belligerent fascist states (Germany, Japan, and Italy) were vanquished and during an era when common opinion presented communism as the foremost threat to liberalism throughout the world. This thesis can best be summarized in Lipset's own words: "Both evidence and theory suggest, however, that the lower strata are relatively more authoritarian, that (again, other things being equal) the will be more attracted towards an extremist movement than toward a moderate and democratic one, and that, once recruited, they will not be alienated by its lack of democracy, while more educated or sophisticated supporters will tend to drop away" (Lipset 484). While Lipset's theory was, at the time, focused on discussing the communist appeal to democratic lower classes it is still relevant today, despite the fall of communism. Through the course of the over half-century since Lipset's theory was published, the lower socioeconomic classes he indicated have undergone a shift from sympathizing with the illiberal left to the illiberal right.

The trend explained and discussed by this theory bears a poor omen for the economic status of those affected by it. Time has shown that authoritarian regimes, be they politically left or right, seldom benefit the lower classes. In this there is the tragic irony of working-class authoritarianism: many of those who would benefit most from the liberal economic policies of centrist political policies are among those who challenge them most vehemently. Should this trend continue to progress and expand, as data suggests it has and will, those of the lower socioeconomic classes in liberal democracies around the world, particularly those based upon a parliamentary structure, risk doing drastic damage to their economic self-interest.

Table 1

Responses of Different German Occupational Groups to Preferred Party System in Percentages (Males Only), 1953					
Occupational Group	Several Parties	One Party	No Party	No Opinion	Number
Civil Servants	88	6	3	3	111
Upper White-Collar	77	13	2	8	58
Free Professionals	69	13	8	10	38
Skilled Workers	65	22	5	8	277
Artisans	64	16	9	11	124
Lower White-Collar	62	19	7	12	221
Businessmen (Small)	60	15	12	13	156
Farmers	56	22	6	16	241
Semi-Skilled Workers	49	28	7	16	301
Unskilled Workers	40	27	11	22	172

Source: Lipset, Seymour Martin. "Democracy and Working-Class Authoritarianism." *American Sociological Review*. 24. no. 4 (1959): 482–501.

Table 2

Proportion of Male Respondents Who Are "More Tolerant" With Respect to Civil Liberties Issues		
Occupational Group	Percentage	Number
Professional and Semi-Professional	66	159
Proprietors, Managers, and Officials	51	223
Clerical and Sales	49	200
Manual Workers	30	685
Farmers or Farm Workers	20	202

Source: Lipset, Seymour Martin. "Democracy and Working-Class Authoritarianism." *American Sociological Review*. 24. no. 4 (1959): 482–501.

Works Cited

Alexander, Damon. "Modelling the Socioeconomic and Demographic Characteristics of One Nation's Vote" (paper presented at 2007 APSA Conference, Monash University, September 24–26, 2007).

"American Fascist Movement (The Third Way)." The American Fascist Movement. www.americanfascistmovement.com.

"British National Party: Report and Accounts for the Year Ended 31 December 2010." UK Electoral Commission. http://www.electoralcommission.org.uk

"Communist Party of Britain: Report and Accounts for the Year Ended 31 December 2010." UK Electoral Commission. http://www.electoralcommission.org.uk.

"Election Center 2008: Exit Polls." CNN News. www.cnn.com.

Ford, Robert. The University of Manchester Institute for Social Change, "Voting for Extremists: Demographic, Attitudinal and Contextual Predictors of Support for the British National Party." http://www.humanities.manchester.ac.uk.

Jackman, Simon. "Political Parties and Electoral Behavior." *The Cambridge Handbook of the Social Sciences in Australia* (2002).

Lipset, Seymour Martin. "Democracy and Working-Class Authoritarianism." *American Sociological Review*. 24. no. 4 (1959): 482–501.

"National Exit Polls Table." *The New York Times*. elections.nytimes.com.

"Party Totals." ABC Elections. www.abc.net.au.

Tetteh, Edmund. UK Parliament. "Election Statistics: UK 1918–2007." 2008. House of Commons Library.

"US President/National/Exit Poll 2004." CNN News. www.cnn.com.

Velasco, Schuyler. "South Bears the Brunt of America's Rising Poverty Rate." *The Christian Science Monitor*, September 13, 2011. www.csmonitor.com.

Webb, Sam. "Does It Matter Which Party Wins?" Catholic Family News. www.cfnews.org.

Zakaria, Fareed. *The Future of Freedom: Illiberal Democracy at Home and Abroad*. New York: W. W. Norton, 2007.

The Internet and the Sociopolitical Development of Nation-States

Ben Jay

ABSTRACT

Over the past two decades, the Internet has revolutionized the spread of information across the world. Much like the printing press of the Renaissance, the Internet has enabled access to a wealth of ideas and facilitated infinitely greater communication for millions of people. As Benedict Anderson has argued, Gutenberg's press and the ideas it helped to spread played a major role in establishing the many national identities of Europe, and eventually today's modern nation-states, out of the ashes of the Roman Catholic–dominated Middle Ages. Twenty-one years since the launch of the World Wide Web, the Internet is already starting to have a similarly dramatic effect. In the West, the Internet has made social activity and research far easier than ever before possible. Elsewhere in the world, the effect is even greater. In China, the Internet is playing a major role in rapid commercialization and industrialization, while social networks are giving citizens an increasingly prominent voice against an authoritarian regime, undermining the constant pressure of strict government censorship and propaganda. In India, rapidly expanding mobile networks are connecting hundreds of millions in remote areas to each other and to government officials, ensuring the effectiveness of the bureaucracy and the endurance of the world's largest democracy.

Printing Press: Primary Weapon of the Revolutionary Arsenal

Anderson explores the immediate impact of Gutenberg's press and its subsequent effects on the role of language in European society and power structures in Chapters Two and Three of his 1983 book *Imagined Communities: Reflections on the Spread of Nationalism*, entitled "Cultural Roots" and "The Origins of National Consciousness," respectively. Johannes Gutenberg first completed his eponymous bible during the 1450s, and between then and the close of the fifteenth century, around 20,000,000 volumes were printed throughout Europe, with presses operating in over 110 towns throughout Europe by 1480, particularly in what are now Italy and Germany (Febvre and Martin 186). During the subsequent sixteenth century, an estimated 150,000,000 to 200,000,000 more volumes were published throughout Europe (Febvre

and Martin 262). Needless to say, over this 150-year period, print media became a ubiquitous aspect of European society, with publishers establishing a market early on, and eventually expanding it to include the virtual entirety of Europe's increasingly literate populace.

Initially, this market only included those who could read Latin, then the primary language of written text in Europe, as it was (and still is) the official language of the Roman Catholic Church. But as this small market composed entirely of polyglots rapidly became saturated, publishers looked to broaden their market, producing content throughout Europe in local vernaculars for the monoglot majority (Anderson 38). The effect of this on society was twofold. First, the widespread distribution of printed media in a variety of languages led to the standardization and consolidation of printed language. Numerous similar dialects across Europe became consolidated into standard printed versions for the sake of convenience and versatility. For example, content printed in High German was distributed across an area populated by people that spoke a large number of adequately similar "lesser" dialects, such as the northwestern Platt Deutsch (Anderson 45). Other neighboring vernaculars, such as Bohemian Czech, were ultimately not quite similar enough to High German to be included under its umbrella by publishers, leading to the development of a distinct printed Czech language (Anderson 45). Similar examples could be found elsewhere in Europe, including in Britain, where dialects from the midlands and north of England fell under the umbrella of King's English from the south, whereas Welsh and Gaelic remained distinct. As these language barriers became more established, the idea that Europe was divided along these lines into a few larger nations, as opposed to countless smaller regions, gradually gained widespread support, leading commoners to eventually identify with certain nationalities over others.

Second, the increased variety of content and the greater number of languages in which that content was available facilitated the free spread of information to the masses in a way never before possible. Until the rise of print media, all information of any consequence, terrestrial or celestial, was tightly controlled by the Catholic Church, which controlled its dissemination through its own channels. All church documents and bibles were written entirely in archaic Latin, making them inaccessible to virtually everyone outside the clergy, allowing the Church to suppress all interpretations of sacred texts other than its own (Anderson 39). As such, when Martin Luther posted his Ninety-Five Theses to the church door in Wittenberg in 1517, it was one of the very first documents of any religious significance to be published in any commonly

understood language, in this case German. Luther's German version of his Theses (the copy on the church door was in Latin) was widely distributed throughout Germany, and according to accounts from the time, they had "within 15 days [been] seen in every part of the country" (Febvre and Martin 289–90). The Reformation ultimately owed its success to Luther's star power as a writer. Between 1518 and 1525, Luther's works accounted for one-third of all books sold in German. Between 1522 and 1546, Luther published 430 editions of Biblical translations (Febvre and Martin 291–95). As such, Luther's commercial success and prolificacy made him into history's first bestselling author.

In short, the print medium gradually weathered away at the domineering influence of the Church, first by ending its monopoly on the distribution of information, and then by spreading information contrary to Church dogma. Eventually, national identities were forged at the expense of religious ones. This led to the consolidation of political power around secular authorities, further undermining the Church's influence in political matters. From there on, the print medium only grew in influence, continuing to spread revolutionary ideas worldwide. Without the printing press, William of Orange's *Declaration of the Hague* and *The Declaration of the Rights of Man* would never have been distributed, the Glorious and French Revolutions, respectively, might have never occurred, and popular sovereignty arguably would not have developed in Britain, France, or anywhere else.

World Wide Web: Printing Press of the Digital Age

At this point we can fast forward several hundred years to 1991, where Tim Berners-Lee, an English computer scientist was working to develop the World Wide Web on the campus of the European Organization for Nuclear Research (CERN), located on the border between France and Switzerland (Kottke.org 2012). Computer-based information networks, now commonly referred to as the "Internet," have existed in various forms since the 1960s, originally commissioned by the United States government for commercial and military interests. Berners-Lee's World Wide Web was intended to be a distribution medium for hypertext documents, or web pages. When users open their web browsers and log onto the web, they're greeted with a web page, which usually contain clickable text links, or hypertext, and if clicked upon will bring the user to another web page with more text links. The World Wide Web, as envisioned by Berners-Lee, is an interconnected network, or "web," of such pages, theoretically containing the collective knowledge of all of humanity, in a format that is readily accessible to anyone

anywhere, and relatively lightweight (accessible on a computer or a mobile device, as opposed to a physical library containing thousands or millions of cumbersome printed volumes).

On August 6, 1991, Berners-Lee posted the first web page, a basic explanation of the World Wide Web project and instructions on how to use a web browser and set up a web server, at the address info.cern.ch (Tim Berners-Lee 1991). On April 30, 1993, CERN announced that the basic foundations of the World Wide Web would be released into the public domain on a royalty-free basis, ensuring the free and widespread growth and development of the Web by both corporate and individual actors (CERN 2003).

In a sense, Tim Berners-Lee is like a modern Johannes Gutenberg, and the World Wide Web is his modern-day printing press. Similarly, Gutenberg used his press almost exclusively to print bibles, while Berners-Lee's initial web pages were purely informational, providing information mostly about CERN and the World Wide Web project. Eventually, as time passed and other innovators rose to prominence, the printing press reached its full potential, ultimately being used constantly to publish countless books and periodicals worldwide. Likewise, CERN's release of the World Wide Web into the public domain opened it up to public innovation in a similar manner. As a result, the Internet of today facilitates the spread of information for research purposes, as it was originally intended, as well as social activity between individuals, as evidenced by the rise of social networking websites like Facebook and Twitter.

The Web in the Western World: So Ubiquitous, It's a Human Right!

In the Western world, the ability of users to access the Internet is largely unrestricted and exploited to near full potential. Individuals can and regularly do use the Web to communicate with each other (social networks have made it infinitely easier to send messages, images, videos, etc., to friends anywhere), keep up with the current events of the day (many newspapers and press organizations post their content online in some form, in addition to any print or broadcast editions), and perform scholarly research (much of this paper was researched using online source material). Similarly, corporations use this technology to provide important product information to existing customers and gauge interest among potential future consumers, while elected officials and government agencies do the same for their constituents. Internet access has become so ubiquitous over the past two decades that a 2011 United Nations Human Rights Council report declared uninterrupted Internet access to be a human right, calling "upon all States to ensure that Internet access is maintained at all times, including during times of political unrest"

(United Nations General Assembly 21). In another example, the government of Finland has taken this logic one step further, declaring in 2010 that broadband Internet access is a legal right for all citizens, and promising to provide connections nationwide by 2015 (Finland makes broadband a "legal right," 2010).

The Web in China: Opening Pandora's Box in a Socialist Vacuum

While citizens of the industrialized states of Western Europe and the Anglosphere now enjoy abundant and relatively unrestricted access, the Internet's current potential is rapidly being realized elsewhere in the world. As two of the world's fastest growing economic and political powers, China and India are at the forefront of this online boom.

Currently, China has one of the world's fastest growing online presences. According to the China Information Network Information Center's (CNNIC) January 2012 Statistical Report on Internet Development in China, 513,000,000 people in China, or 38.3% of the total population, were online, an increase of 55,800,000, or 12%, from the year prior (CNNIC 4). By comparison, in 2004, there were 204,000,000 users in the United States in 2004, or about 75% of the population then (Nielsen/NetRatings 1). Additionally, 73.4% of Chinese Internet users access the web on desktop computers, while 69.3% access the web on mobile devices (CNNIC 2012 4). Of all home computer users, 392,000,000, or 98.9%, have broadband connections (CNNIC 2012 4), meaning that high-speed access is well within the reach of China's growing middle class. Furthermore, 136,000,000, or 26.5% of all Chinese Internet users are located in rural areas, an increase of 11,130,000, or 8.9% over the year before (CNNIC 2012 4), suggesting slower, but still significant growth outside China's bustling metropolises.

Politically, China is an authoritarian one-party state under the control of the Communist Party of China (CCP), and has been since 1949. As such, the government has consistently sought to carefully measure the Internet's level of penetration in China, so as to promote economic growth opportunities while also attempting to limit any potential for online antigovernment dissent. In 1993, the "Golden Bridge" project was introduced, aimed at modernizing China's telecommunications infrastructure (Lewis 5), which included a number of key Internet regulations first implemented in 1994, followed by the Internet's commercial launch in China in 1995 (Lewis 3).

Three main regulations govern China's Internet. The first is the Temporary Regulation for the Management of Computer Information Network International Connection, passed in 1996

and updated in 1997 (Qiu 10). The Temporary Regulation prohibits individual users from "establishing a direct international connection by themselves," requires all Internet service provider (ISPs) connections go through one of four networks, ChinaNet, GBNet, CERNet or CSTNet (Qiu 10), and mandates that all ISPs obtain government licenses (11). The second is the Ordinance for Security Protection of Computer Information Systems, which was issued by the State Council in 1994 to give the Ministry of Public Security the power to "supervise, inspect and guide" nationwide Internet security and prosecute any violations (Qiu 11). This regulation led to the introduction of a subsequent regulation in 1997, the Computer Information Network and Internet Security, Protection and Management Regulation, which lists specific illegal online activity (Federation of American Scientists, New PRC Internet Regulation, 1998), including any use of the Internet intended to "overthrow the government or the socialist system" (Section 5, §2), disrupt national unity (Section 5, §3), "inciting hatred or discrimination among [minority] nationalities" (Section 5, §4), promoting superstition or vice (Section 5, §6), inciting terrorism, criminal activity or slander (Section 5, §7), or spreading any vaguely defined "falsehoods" or "rumors" (Section 5, §5). The third is State Council Order No. 292: Measures for Managing Internet Information Services, issued in September 2000, which set the first restrictions aimed directly at content providers, or providers of "Internet information services" (IIS) (China Culture, 2003, Measures, 2000). Specifically, IIS providers must be licensed and registered with the proper provincial authorities (Articles 7 and 8), are responsible for maintaining the legality of their content (Article 11), must record all user access data for at least 60 days and make those records available to any "relevant state authorities" upon request (Article 14), and further reinforce and apply the regulations set forth by the 1997 Management Regulation to IIS providers (Article 15).

As such, these regulations and ordinances ultimately give the Chinese government the technical and legal capability to monitor and censor Internet content and activity within its borders, define broadly what content and activity is illegal, and conduct fully unwarranted investigations related to any illegal activity and content. In short, the Chinese government theoretically has the ability to absolutely control the Internet within its borders to further its own political ends, entirely at the expense of the ideals of freedom and openness that the World Wide Web project has strived to promote. In China today, access to many websites is blocked sporadically or permanently, including social networks such as Facebook and Twitter,

informational websites such as Wikipedia, and foreign news outlets such as *The New York Times* and BBC News. Internet police can and do monitor the activity of individual users, and often remove subversive blog and forum posts at will.

Foreign Internet companies are similarly bound to state regulations in order to legally hold operations in China. Google attracted controversy in 2006 after launching google.cn, the Chinese local version of its search engine, which censored search results in accordance with Chinese law (Grossman 2006). For example, when a user does an image search for "Tiananmen" in google.com, the results mostly include pictures of the iconic tank man at the 1989 Tiananmen Square democracy protests, but when a user does the same search at google.cn, the results include mostly official pictures of the square and the Gate of Heavenly Peace (Grossman 1). Google co-founder Larry Page responded to the controversy by admitting that filtering results compromises Google's mission and "Don't be evil" ethos to a degree, but "failing to offer Google search at all to a fifth of the world's population, however, does so far more severely" (Grossman 1). Google eventually stopped filtering its results in 2010, following cyberattacks that it traced to Chinese computers, possibly related to the government (China condemns decision by Google to lift censorship, 2010). Despite, or perhaps because of the blow to China's image caused by the fallout over Google, domestic Internet companies have managed to thrive in China, regardless of their greater compliance with government censorship standards. China's largest search engine, Baidu, still holds the bulk of China's search market (Jin 2011), despite Google's presence, while social networks Renren and Weibo are popular local alternatives to their banned foreign counterparts, Facebook and Twitter, respectively, despite heavy government monitoring.

Despite all of this, controversial user-generated content has appeared in China, and government censors have continually had to play catch-up to stop them. Numerous organizational websites and individual blogs promoting sensitive political topics, including democracy, the Falun Gong movement, Taiwanese independence, and anticorruption, have drawn the continued ire of the state. Paris-based nongovernmental organization Reporters Sans Frontières has referred to China as "the world's biggest prison for netizens," and London-based Amnesty International has noted that China "has the largest recorded number of imprisoned journalists and cyber-dissidents in the world" (Yu 2012). However, the government has not been

able to achieve the gargantuan task of wiping out all online dissent, and subversive content continues to emerge regularly.

In addition to controversial political content, the government has also attempted to crack down on what it perceives to be "sexually suggestive" content (Federation of American Scientists, New PRC Internet Regulation, Section 5, §6, 1998). As such, Internet pornography is banned in China in largely the same way that subversive political content is. Li Li, a Guangzhou based blogger writing under the pseudonym Muzi Mei, caused considerable controversy in 2003 when she started posting extensive and detailed accounts of her sex life. Li's blog became extremely popular in China, a country where public sexuality is seen as highly taboo, and Li became a national celebrity despite government attempts to censor her (Sex and the Single Chinese 1), so much so that she was the most searched entry on search engine Sohu.com, ahead of runner up Mao Zedong (Yardley 2003).

The case of Muzi Mei brings up two important issues related to the liberalization of media. First, her blog's success demonstrates the ability of new media to effect social change in addition to political change. Today's young Chinese are considerably more open about their sexuality then their parents, who were largely educated under the hard line suppression of the Cultural Revolution. According to sociologist Li Yinhe, 70% of Beijing residents admitted to having had premarital sex in 2005, compared to 15.5% in 1989 (Sex and the Single Chinese 2005, 1). While China's budding sexual revolution has had some growing pains, as evidenced by rising HIV infection rates and increased abortions among single mothers, increasingly widespread sexual knowledge will, in all likelihood, make Chinese society more able to effectively deal with sexual issues as a whole in the long run. Second, her blog also raises questions about the role of obscene and otherwise intellectually non-stimulating content in a free society, as the blog is still composed entirely of accounts of Li having sex, including a 25 minute long audio podcast of her orgasm noises that crashed its host server when 50,000 people attempted to download it simultaneously (Sex and the Single Chinese 1). Like the romance novels and Playboy magazines of the print medium, spam and porn have a ubiquitous presence on the Internet, with porn sites accounting for an estimated 30% of all online data transfer (Anthony 2012). While this content may be a nuisance for much of the online public, its existence does highlight the fact that in open societies, content producers are allowed the freedom to express whatever they please and content consumers can consume whatever they

please. Furthermore, producers and consumers have the responsibility to make that choice wisely, and if they didn't have that choice, their content would ultimately be considerably less effective in effecting any real social or political change. It's not ideal, but it's arguably better than having highly controlled content from limited sources. To quote Li, "I express my freedom through sex, it's my life, and I can do what I want" (Sex and the Single Chinese 1).

Currently, the Chinese government's approach to the Internet suggests a desire to maintain a strong and stable political order and a fairly conservative social order, while also promoting the country's sustained economic growth. In short, the government is attempting to clear a path of least resistance for China's future political and social advancement as it grows economically. Despite the strong historical links between sociopolitical liberalization and capitalist growth, it remains to be seen whether China's current model of Internet censorship will be sustainable. While the so-called Great Firewall of China certainly blocks a considerable amount of content, it also hasn't been able to keep up with all of the content, and as such, exists to some degree to intimidate rather than comprehensively censor. What is certain is that China's ability (or inability) and willingness (or lack thereof) to liberalize will ultimately have a great effect on its digital future, and vice versa.

The Web in India: A Brilliant Tool for Liberal Growth, Under Threat

To the immediate southwest of China is another one of the world's rapidly growing Internet markets, and a vastly different one at that. In India, the world's largest democracy, there are an estimated 121,000,000 Internet users, or 10.2% of the total population (Asia Internet Usage Stats, 2012), a massive online presence, albeit one with much smaller penetration than China's or the United States'. Additionally, only 2% of rural Indians have Internet connections (Vaidyanathan 2012). The quality of India's connections lag behind China's as well. Of that 10.2% online, only about 11% of them, or 13,350,000 have access to broadband connections (Indian Telecom Services, 4), as the growth of high-speed networks is currently hampered by the physical limitations of existing network infrastructures. However, India's online presence is growing rapidly, as the number of Internet users in China rose 25% in 2011 alone, and 59% of users access the Internet exclusively through mobile devices (Vaidyanathan 2012), of which there were 791,380,000 subscribers in February 2011 (India Telecom). Unlike China, individual connections are minimally regulated, and numerous private companies dominate the Internet market alongside a few state owned enterprises.

The Indian government has aggressively sought to expand access in several ways. First, the government has proposed setting up a national fiber optic network (NOFN) to provide telecommunications links to every village in India by 2014, an estimated ₹20,000 *crore* ($3,807,200,000 US) investment. According to Bharat Sanchar Nigam Limited (BSNL) CMD R.K. Upadhyaya, the NOFM "would help in offering governing, banking and health and other basic services online up to the villages are rural areas" once created (Telecom Links in all Indian villages by 2014: BSNL, 2012). Second, the government and British tech company DataWind have recently collaborated to produce and sell the Aakash, an inexpensive tablet computer running Google's Android operating system aimed largely at students (Magder 2012). After an initial pilot run of 100,000 free units, the Aakash will be sold to students for $35 apiece, with a commercial version called UbiSlate 7+ retailing for $60 (Magder 2012). DataWind CEO Suneet Tuli has described the Aakash as not quite living up to Apple's ubiquitous iPad in terms of quality, but argued that "if [it's] one of your first tablet experiences, you'll be impressed," adding "The iPad customer isn't our customer—but you can buy about 15 of these for the price of one of those" (Magder 2012). Telecoms and Education Minister Kapil Sibal stated, "The rich have access to the digital world; the poor and ordinary have been excluded. Aakash will end that digital divide" (Magder 2012). Third, India is already a major information technology hub, as much of the population speaks English, giving it a major advantage in a world where the most consistent Internet access exists in North America and Western Europe, especially in the Anglosphere. Eventually, these efforts and many others will likely facilitate the expansion of Indian Internet access through free capitalist growth.

Despite India's strong democratic tradition, the government does engage in some censorship activity under the guise of security, censoring content that the government believes could incite violence, especially anything related to the ongoing dispute with Pakistan over Kashmir (ONI Country Profile 300). However, journalists are rarely detained, and are usually released quickly when they are (ONI Country Profile 300). Censorship and monitoring efforts have increased considerably since the 2008 bombings in Mumbai (Internet Enemies Report 7), whereas before that, such measures were rare. Furthermore, according to Freedom House's Freedom on the Net 2011 report, bloggers and moderators can face "libel suits and even criminal prosecution for comments posted by others on their websites" (Freedom on the Net 170). Additionally, communications interceptions can be conducted without prior judicial approval,

and all ISPs are required to allow authorities access to user data (Freedom on the Net 2011 171), effectively allowing unwarranted online surveillance.

While this illiberal surveillance is quite unfortunate, it is not nearly as comprehensive or invasive as China's online censorship, and has not had nearly as detrimental an effect on free expression. It's also worth noting that similarly illiberal laws and bills have emerged in older, more established democracies, including the United States and Canada. The USA PATRIOT Act of 2001, passed shortly after the attacks of September 11, 2001, vastly expanded government surveillance powers online, requiring ISPs to provide detailed individual usage and financial data for subpoenas (USA PATRIOT ACT: H.R. 3162, Public Law 107-56, Title II, Sec. 210), for instance. The Stop Online Piracy Act (SOPA) sought to inefficiently punish and block websites for hosting copyright infringing content posted by individual users (Stop Online Piracy Act: H.R. 3261, Title I, Sec. 102) before it was defeated in the U.S. House of Representatives, while Bill C-30, currently proposed in the Canadian House of Commons, would allow for "legal access," or unwarranted access by authorities of user data from ISPs and telephone providers, as well as a "back door" provision to allow easier communication interception at will (C-30, Summary). Even C-30's title, the "Protecting Children from Online Predators Act," serves to propagandize the bill, much like the Indian and U.S. governments have sought to propagandize the Mumbai bombings and the 9/11 attacks, respectively.

Despite all of this, however, India's efforts to expand quality Internet access nationwide generally serve to further promote liberal free expression. By extension, this would also promote further economic growth and the deepening of India's democracy. Unfortunately, as a democracy, India's government is susceptible to public paranoia, political opportunism and other undue influence that can threaten individual privacy rights and hold back a society commercially and creatively.

Conclusion: State and Society in the Digital Future

The rise of the Internet has had a tremendous effect in revolutionizing global communication over its two-decade existence, much like the printed volume it's rapidly replacing. The growth of both mediums has been driven largely by economic motivators, as well as political ones, as service providers, corporations and governments seek to cash in, much in the same way that publishers and early authors did during the fifteenth and sixteenth centuries, facilitating the rapid transfer of information and ideas on a level never before possible. In China,

the government has simultaneously sought to foster Internet growth for purposes of economic growth, while also seeking to strictly control content for political purposes. Whether China can maintain this balance in the long term remains to be seen, but what is certain is that the Internet has had a dramatic overall effect on Chinese government and society so far, and has served to somewhat liberalize society. In India, the government has sought to foster Internet growth for purposes related to both economic development and the expansion of democracy. While much of India is still not connected, and there have been some hiccups along the way in the area of user privacy, India's future as a major and open digital hub remains bright. Therefore, it can be assumed that the continued growth of the Internet worldwide, both in industrialized Western states and in the developing world, will only serve to further promote political, economic and social growth and change well into the future, much like printed volumes did in centuries past, all the way back to Gutenberg.

Works Cited

Anderson, Benedict. *Imagined Communities: Reflections on the Origin and Spread of Nationalism.* 1983. London: New Left Books, 2006. Print.

Anthony, Sebastian. "Just how big are porn sites?" *Extremetech.com*. Extreme Tech, 4 April 2012. Web. 25 April 2012. <http://www.extremetech.com/computing/123929-just-how-big-are-porn-sites?print>.

"Asia Internet Usage Stats Facebook and Population Statistics." *Internet World Stats.* Internet World Stats, 2012. Web. 2012 April 16. <http://www.internetworldstats.com/stats3.htm>

Canada. Parliament. House of Commons. "Protecting Children from Online Predators Act." Bill C-30, 41st Parliament, 1st Session, 2011–12, Ottawa. 14 February 2012. Reading.

CERN. *10 Years Public Domain*. CERN, 2003. Web. 15 April 2012. <http://tenyears-www.web.cern.ch/tenyears-www/Welcome.html>

"China condemns decision by Google to lift censorship." *BBC.co.uk.* British Broadcasting Corporation, 23 March 2010. Web. 2012 April 16. <http://news.bbc.co.uk/2/hi/asia-pacific/8582233.stm>

CNNIC. 2012. Statistical Report on Internet Development in China. <http://www1.cnnic.cn/uploadfiles/pdf/2012/2/27/112543.pdf> (2012, Apr. 16)

"Country Profile: India." *Opennet.net.* OpenNet Initiative, 2011. Web. 16 April 2012. <http://access.opennet.net/wp-content/uploads/2011/12/accesscontested-india.pdf>

"Country Report: India." *Freedomhouse.org.* Freedom House, 2011. Web. 16 April 2012. <http://www.freedomhouse.org/sites/default/files/inline_images/India_FOTN2011.pdf>

Febvre, Lucien, and Henri-Jean Martin. *The Coming of the Book: The Impact of Printing, 1460–1800.* 1958. 3rd ed. London: Verso, 1976. Print.

"Finland makes broadband a 'legal right'." *BBC.co.uk.* British Broadcasting Corporation, 1 July 2010. Web. 15 April 2012. <http://www.bbc.co.uk/news/10461048>

Grossman, Lev. "Google Under the Gun." *Time.com*. Time, 2006 Feb 5. Web. 16 April 2012. <http://www.time.com/time/magazine/article/0,9171,1156598,00.html>

Jin, Tony. Feb. 15, 2011. "China Economic Watch: Regional GDP, Inflation, Search Market, Rate Hike." *The China Perspective.com.* The China Perspective, 15 February 2011. Web. 16 April 2012. <http://www.thechinaperspective.com/articles/chinaeconomicwa-8181/>

Lewis, James A. "The Architecture of Control: Internet Surveillance in China." *CSIS.org.* Center for Strategic and International Studies, 2006. Web. 16 April 2012.

<http://csis.org/files/media/csis/pubs/0706_cn_surveillance_and_information_technology.pdf>

Magder, Jason. "Aakash will end 'digital divides'; Designed in Montreal, world's cheapest tablet computer priced at $60." *Montrealgazette.com.* The Gazette, 3 January 2012. Web. 16 April 2012. <http://www.montrealgazette.com/opinion/Aakash+will+digital+divide+Designed+Montreal+world+cheapest+tablet/6115731/story.html>

"Measures for Managing Internet Information Services." *China Culture.org* China Culture, 2003. Web. 2012 April 16. <http://www.chinaculture.org/gb/en_aboutchina/2003-09/24/content_23369.htm>

"New PRC Internet Regulation." *Fas.*org. Federation of American Scientists, 1998. Web. 15 April 2012. <http://www.fas.org/irp/world/china/netreg.htm >

Qiu, Jack Linchuan. "The Architecture of Control: Internet Surveillance in China. Center for Strategic and International Studies." *International Journal of Communications Law and Policy*. Vol. 4. Winter 1999/2000. Web. 16 April 2012. <https://docs.google.com/viewer?a=v&q=cache:KprFbiPUySUJ:citeseerx.ist.psu.edu/viewdoc/download%3Fdoi%3D10.1.1.106.532%26rep%3Drep1%26type%3Dpdf+VIRTUAL+CENSORSHIP+IN+CHINA:+KEEPING+THE+GATE+BETWEEN+THE+CYBERSPACES&hl=en&gl=us&pid=bl&srcid=ADGEESiZoAOv6FILaBNtRo6IIoANRshufv2Ksqc8LJ8GG-GSGCnEqJ1FeIesBIhWv4ecc7-2K17cYC0gB80kzYfoFOUDjUmfZ2qWbpqPLuWYlIHw2N-wRCtsK20Uz1gSJ5GP35lRCGe_&sig=AHIEtbSgWeUhXAJQgxeUNGvV1JFjXyqvFw>

Reporters Without Borders. "Internet Enemies Report." *Reporters Sans Frontières.* R2012, 12 March 2011. Web. 16 April 2012. <http://march12.rsf.org/i/Report_EnemiesoftheInternet_2012.pdf>

"Sex and the Single Chinese." *Time.com*. Time Magazine, 5 Dec 2005. Web. 16 April 2012. <http://www.time.com/time/magazine/article/0,9171,1137697-1,00.html>

Stop Online Piracy Act (U.S. H.R. 3261); House Judiciary Committee, 26 Oct 2011. Web.

Telecom India Online. India Telecom Stats—February 2011. <http://www.telecomindiaonline.com/india-telecom-growth-and-subscribers-2011.html> (2012, Apr. 16)

"Telecom links in all Indian villages by 2014: BSNL." *Timesofindia.indiatimes.com.* The Times of India, 25 March 2012. Web. 16 April 2012. <http://articles.timesofindia.indiatimes.com/2012-03-25/india/31236443_1_mobile-connections-telecom-links-bsnl-official>

Telecom Regulatory Authority of India. "Indian Telecom Services Performance Indicator Report for the Quarter Ending December 2011." *Trai.gov.* Telecom Regulatory Authority of India, 2012. Web. 16 April 2012. <http://www.trai.gov.in/WriteReadData/PressRealease/Document/quarterly%20press%20release.pdf>

"The web was invented in France, not Switzerland." *Kottke.org.* Kottke, 2012. Web. 2012 April 15. <http://kottke.org/12/01/the-web-was-invented-in-france-not-switzerland>

Tim Berners-Lee. <tbl@cernvax.cern.ch>, 1991, Aug. 6. alt.hypertext. 2012, Apr. 15. <WorldWideWeb: Summary>

U.N. General Assembly, Human Rights Council. "Report of the Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression, Frank La Rue, A/HRC/17/27." 16 May 2011. Report.

USA PATRIOT ACT (U.S. H.R. 3162, Public Law 107-56).

Vaidyanathan, Rajini. "Is 2012 the year for India's internet?" *BBC.co.uk.* British Broadcasting Corporation, 3 Jan 2012. Web. 16 April 2012. <http://www.bbc.co.uk/news/business-16354076>

"Virtual Censorship in China: Keeping The Gate Between Cyberspaces." *Nielsen-online.com.* Nielsen//Netratings, 2004. Web. 16 April 2012. <http://www.nielsen-online.com/pr/pr_040318.pdf>

Yardley, Jim. "Internet Sex Column Thrills, and Inflames, China." *Nytimes.com.* The New York Times, 30 November 2003. Web. 16 April 2012. <http://www.nytimes.com/2003/11/30/world/internet-sex-column-thrills-and-inflames-china.html?pagewanted=all&src=pm>

Yu, Miles. "Inside China." *Washingtontimes.com.* Washington Times, 8 February 2012. Web. 16 April 2012. <http://www.washingtontimes.com/news/2012/feb/8/inside-china-719761130/?page=all>

The Oppressive Institution of Victorian Sexual Morality as Represented By the Women of Wildean and Shauvian Drama

Emily Webb

Through their various portrayals of women in late nineteenth-century drama, Oscar Wilde and George Bernard Shaw demonstrate disparate approaches towards the dispute against Victorian sexual morality as informed by their respective views of contemporary social convention. In the wake of the social purity movement that strove to outlaw male lust, Wilde explores the sexual politics in the seduction of his female protagonists—specifically Mrs. Arbuthnot in *A Woman of No Importance* (1893) and Lady Windermere in *Lady Windermere's Fan* (1892)—in such a way that suggests the futility of an ethical hierarchy that condones the sexual transgressions of men at the expense of feminine purity. Influenced by his advocacy of women's rights as a major proponent of the Fabian socialist movement, Shaw further denounces this degradation of women under corrupt patriarchal authority through the development of his authoritative female characters, such as Vivie Warren in *Mrs. Warren's Profession* (1893) and Gloria Clandon in *You Never Can Tell* (1897), who claim dominance over men in both the domestic and public spheres. While Wilde merely exposes the double standard of sexual morality as his female protagonists risk public shame for their susceptibility to male vice, Shaw actively condemns oppressive notions of proper womanhood to promote sexual tolerance in Victorian society through various representations of the self-respecting woman.

Without compromising his reputation as a notoriously extravagant dandy whose penchant for hedonistic pleasures was tolerated in nineteenth-century high society, Wilde tested the hypocritical moral hierarchy that social purists strove to reform through his own transgressive behavior. Though Wilde refused to support this movement to outlaw the sexual crimes of men, British journalist and renowned social activist W. T. Stead observed how the criminal charges against the playwright for his acts of gross indecency

exposed the discriminatory moral values that are promoted in Victorian society (Foldy 129–130). When reflecting on the public conviction of Wilde in 1895 for his crimes of sodomy, Stead tries to rationalize Wilde's lewd homosexual exploits and argues that

> If the promptings of our animal nature are to be the only guide, the punishment of Oscar Wilde would savour of persecution, and he might fairly claim our sympathy as the champion of individualism against the tyranny of an intolerant majority…If Oscar Wilde, instead of indulging in dirty tricks of indecent familiarity with boys and men, had ruined the lives of half a dozen innocent simpletons of girls, or had broken up the home of his friend by corrupting his friend's wife, no one could have laid a finger upon him. The male is sacro-sanct: the female is fair game (Stead 491–2).

Although he regards Wilde's homosexual transgressions as abnormal conduct, Stead asserts that the social persecution of such deviant behavior ultimately sheds light on the accepted standards of Victorian morality that tolerate the equally lecherous crimes of heterosexual men. Indicating the oppressive demands for female purity within patriarchal society, Stead claims that the established hierarchy of sexual morality would typically overlook similar acts of sexual deviance that are committed by men and rather exploit women. As he inadvertently called into question this double standard of Victorian sexual morality through his criminal acts of homosexual lust, Wilde further evaluated the need for a universal code of ethics in contemporary society in the dramatic action of his societal comedies, which unfolds through the discourse of his morally dubious characters that upholds arbitrary decorum among the privileged class.

Translating this marginal evaluation of the hypocritical hierarchy of Victorian ethics into the sexual politics between the men and women of his modern plays, Wilde constructs brief moral conflicts for his female protagonists that merely assess the oppressive demands for female purity under depraved patriarchal authority. While examining the evolution of his righteous female characters, Gregory Mackie claims that Wilde portrays the dramatic action of his plays only to analyze the moral outlooks of these women and further states that, "The process of moral evaluation, however, is limited by its being confined to a field of binaries and dualistic terms, and character development in these comedies is ultimately subject to a relatively firm ethical (and verbal) typology" (Mackie 157). As the playwright assesses the moral position of his female protagonists, Mackie argues that Wilde limits the development of these characters

within a binary code of ethics in such a way that refrains from obstructing the hypocritical hierarchy of sexual morality. In exploring the perspective of a victim of male vice through the social struggles of the scorned Mrs. Arbuthnot in his play *A Woman of No Importance* (1893), Wilde restricts her bout of moral enlightenment to the dualistic expectations of female purity and male depravity as she endures public condemnation for her weakness in the face of seduction. When she confronts her depraved seducer, Lord Illingsworth, after experiencing years of public shame, Arbuthnot pleads that he leave their love child in her custody and bemoans that, "I have had twenty years of sorrow, and I have only had one thing to love me, only one thing to love. You have had a life of joy, and pleasure, and success. You have been quite happy, you have never thought of us" (Wilde 139). As she tries to inform Illingsworth about the repercussions of his lecherous actions that defamed her reputation, Mrs. Arbuthnot notes how the depraved dandy continues to thrive in high society despite his similar sexual transgressions. Establishing her status as a tragic character, Mrs. Arbuthnot acknowledges the paradoxical hierarchy of moral values that denounces her submission to male vice while condoning such acts of impurity among men as she laments her ostracism from society. Though she first mourns her scorned existence outside of high society, Mrs. Arbuthnot refuses her son's plea to marry her vindictive seducer for the sake of salvaging her reputation and insists that, "I am disgraced; he is not. That is all. It is the usual history of a man and a woman as it usually happens, as it always happens. And the ending is the ordinary ending. The woman suffers. The man goes free" (Wilde 161). While providing keen insight into the discriminatory code of ethics imposed upon men and women, Mrs. Arbuthnot seems to come to terms with her status as a "fallen woman" who exists outside of this paradoxical moral hierarchy when she rejects social expectations of convenient marriage to reclaim her purity. Despite the fact that her enlightened character chooses to preserve her dignity over her social reputation, Mrs. Arbuthnot ultimately evades taking an active stance against the degradation of women at the hands of patriarchal sexuality and allows acts of male vice to persist through the social construct of sexual morality. Thus, Wilde refrains from disturbing the rigid social structure, which perpetuates these discriminatory standards of feminine purity, through the fleeting moral conflict of his female protagonists and, in turn, only exposes the sexual double standard in Victorian society.

In order to maintain this arbitrary decorum within high society, Wilde further evaluates the paradoxical nature of contemporary moral values through the witty discourse of his male and female characters that provokes the vexing moral dilemma of his female protagonists. Analyzing the stylistic function of the playwright's dramatic language, Mackie deconstructs the epigrammatic dialogue shared between his male and female characters and asserts that

> The structure of the epigrams that appear in both dramatic dialogue and lists of maxims is a matter of balance and hierarchy, of setting contrasting terms such as goodness and badness, virtue and vice, truth and lies, in opposition to one another, while maintaining the structure of an ethical hierarchy that gives such oppositions their meaning (Mackie 156).

While juxtaposing the notions of good and bad through the epigrammatic language of his morally dubious characters, Mackie claims that Wilde establishes equilibrium among such disparate concepts in this lexical opposition in an attempt to affirm the presence of a hypocritical moral hierarchy in Victorian society. In his first play, *Lady Windermere's Fan* (1892), Wilde exemplifies this balanced contrast through the clever exchange between the morally upright Lady Windermere and the seemingly nefarious dandy, Lord Darlington, which prompts the moral conflict of his female protagonist and tests the paradoxical expectations of sexual purity. When Lady Windermere expresses her belief in a universal standard of moral conduct for a presumably hypothetical case of male infidelity, Lord Darlington contests the absolute division between purity and depravity in society and jokes that, "It is absurd to divide people into good and bad. People are either charming or tedious" (Wilde 43). As he mocks the institution of Victorian morality through their witty discourse, Lord Darlington undermines contemporary moral values through the dual disparity between ethical behavior and social bearing that is posited in these epigrams and suggests the futile existence of this moral binary. Once she finds herself caught in this discrepancy of proper moral conduct when she must choose between the forbidden love of Lord Darlington and her duty to her allegedly adulterous husband, Lady Windermere soon begins to question social demands for female purity as she laments, "And yet which is the worst, I wonder, to be at the mercy of a man who loves one, or the wife of a man who in one's own house dishonours one? What woman knows? What woman in the whole world?" (Wilde 73) While her righteous character

struggles to commit her own sexual transgressions against her husband for fear of social condemnation, Lady Windermere internalizes this lexical opposition between the polarized concepts of virtue and etiquette and further perpetuates Darlington's notion of the worthless ethical hierarchy that condones male depravity. Even though she moves to disrupt this moral double standard that often admonishes the sexual deviance of women rather than men, Lady Windermere refrains from challenging its oppressive structure in patriarchal society by reassessing her personal code of ethics once the play's dramatic action reaches its abrupt resolution (Wilde 88). Through the marginal development of his female protagonists' moral conflict as represented in the epigrammatic dialogue of his male and female characters, Wilde preserves the arbitrary decorum within Victorian high society that upholds this hypocritical hierarchy of sexual morality.

While Wilde merely tested the paradoxical construct of nineteenth-century Victorian morality through his own criminal acts of sexual deviance, George Bernard Shaw worked to overturn oppressive demands of ideal womanhood that derived from this binary code of ethics as a result of his social activism. As a major proponent of the Fabian Socialist movement, Shaw expounded his belief that the element of humanity molds both men and women into functioning members of society who make equal contributions to its development. While the playwright believed in the human potential of both men and women in improving society, Sally Peters states that Shaw urged for the establishment of equal rights among the sexes and explains that, "[Shaw] asserted that 'the denial of any fundamental rights' to a woman is really 'a violation of the soul'…He saw his work as that of guiding the Fabians toward a new society to benefit both men and women" (Peters 19). In his public advocacy of such feminist principles, Shaw denounced the lack of female involvement in the public sphere at the expense of social reform and, in turn, worked to subvert patriarchal edicts of ideal domesticity and sexual purity that were inflicted upon women. Demonstrating his adherence to Fabian Socialist doctrine, Shaw promoted the gradual reform of social demands for proper womanhood through his many works of modern drama, which scrutinized such discriminatory standards of Victorian morality in support of his campaign for sexual equality.

In the development of his provocative "discussion plays," Shaw openly condemns social conventions of ideal womanliness through the final resolution of their dramatic

action that spawns from the moral conflict of his defiant, authoritative female characters. As expressed in his essay "The Womanly Woman," Shaw argues that patriarchal expectations of female responsibility within the domestic and public spheres deprive women of their individual rights as equal members of society and maintains that, "Therefore Woman has to repudiate duty altogether. In that repudiation lies her freedom; for it is false to say that Woman is now directly the slave of Man; she is the immediate slave of duty" (Shaw 44). By urging Victorian women to renounce their sense of noble, self-sacrificing duty and assert their own independence under patriarchal authority, Shaw fosters the concept of an empowered, self-reliant woman who thrives in contemporary society, which he further illustrates through strong-willed female characters in his plays who thwart standards of proper womanhood. In attributing typical masculine traits to the determined character of Vivie Warren in his play *Mrs. Warren's Profession* (1893), Shaw constructs the image of his autonomous female protagonist as she establishes her position within the patriarchal public sphere in such a way that undermines expectations of her rightful duty to the household. After the influential Sir George Crofts offers her financial stability and high social standing through marriage, Vivie declines his proposal along with the extravagant leisures of privileged high society when she states, "I am much obliged to you for being so definite and business-like. I quite appreciate the offer: the money, the position, Lady Crofts, and so on. But I think I will say no, if you don't mind. I'd rather not" (Shaw 74). Though the insincere gentleman tries to appeal to her pragmatic personality with his forthright terms of a marriage of convenience, Vivie rejects this call to her feminine duty of submissive domesticity so that she may focus on advancing within the working world and achieve her own economic means among her male counterparts. When she later faces much scrutiny from Praed and Frank Gardener for her austere, unromantic character, Vivie dismisses their insistence on her necessary sentimentality and domesticated nature as a woman and asserts that, "You are welcome to any illusions you may have left on these subjects: I have none. If we three are to remain friends, I must be treated as a woman of business, permanently single [*to Frank*] and permanently unromantic [*to Praed*]" (Shaw 86). Resisting the male influences that intrude upon her space within the public sphere, Vivie discredits their expectations of her responsibilities as a pious, delicate woman and a doting wife by upholding her proclivity

for self-reliance in such a way that contests Victorian standards of love and marriage. Although she may not seem to triumph over patriarchal authority as she continues to fulfill her noble duty as a businesswoman under the dominant males of the working world, Vivie still asserts her independence after enduring the oppressive demands for proper womanhood that drove her mother to engage in sordid business so that she may achieve her own respectable means for economic success. Rather than limit the development of his female protagonists to the binary of Victorian morality as illustrated in Wilde's plays, Shaw fully explores their moral conflict to the point where their characters subvert the social construct of ideal womanliness and claim dominance within patriarchal society.

Despite the fact that he portrays several of his female protagonists in a conventional domestic role as they fulfill their supposedly feminine duties of self-sacrificing love and sexual purity, Shaw manages to permutate his conception of the independent, authoritative women through the dramatic action that unfolds within the domestic sphere to closely scrutinize and subvert such patriarchal standards of ideal femininity. When he considers the helpless circumstances of ideal Victorian wives within the household, Shaw upholds that these women must assume a matriarchal position over their domestic domain in order to reclaim their dignity and reasons that, "The self-respect she has lost as a wife she regains as a mother; in which capacity her use and importance to the community compare favourably with those of most men of business" (Shaw 40). By attaining authority within a household that thrives on their self-sacrificing duty, Shaw argues that women can undermine the oppressive standards of ideal domesticity that demean their natural womanhood and regain a sense of autonomy and self-worth through her dominance. In the development of his female protagonist Gloria Clandon in his comedic play *You Never Can Tell* (1897), Shaw illustrates this notion of female empowerment through domestic duty while her strong-minded character works to elicit her command over the institution of Victorian marriage. As she struggles to protect her strong convictions against the affection of the lustful bachelor, Mr. Valentine, Gloria laments her lack of moral strength after she kisses him and admits that, "I am one of those weak creatures born to be mastered by the first man whose eye is caught by them; and I must fulfil my destiny, I suppose" (Shaw 86). By conceding to her presumably

natural sentimentality as a woman overcome by passion, Gloria appears to revoke her self-reliant feminine identity and succumb to patriarchal expectations of passive womanhood as she reluctantly pursues this love affair with her male counterpart. Once she learns that Valentine earns meager wages as a dentist, Gloria aggressively embraces her suitor and takes it upon herself to propose marriage when she declares, "Now let us have no false delicacy. Tell my mother that we have agreed to marry one another" (Shaw 122). Knowing that the irresistible bachelor actually holds weak economic standing, Gloria moves to acquire control over Valentine through marriage so that she may reclaim her self-respect as a morally upright, autonomous wife within their household. Although she seems to compromise her self-reliant character in assuming the oppressive domestic duty imposed upon women, Gloria ultimately redeems the weakness of her own principles as she gains dominance over this financially unstable patriarch and asserts her independence. Thus, Shaw resolves the moral struggles of his authoritative albeit conflicted female protagonists who face oppressive demands for ideal womanliness in such a way that suggests the rightful independence of women through their matriarchal authority acquired through domestic duty.

Though both playwrights illustrate patriarchal demands for proper womanhood that are upheld within Victorian society, Oscar Wilde and George Bernard Shaw convey their divergent perspectives on the social construct of sexual morality through the different magnitudes of dramatic action that develop the female protagonists in their respective works. Considering his own immoral behavior as a sexually deviant dandy within high society during the late nineteenth-century, Wilde restrains the moral conflict of the righteous women in his plays—specifically Mrs. Arbuthnot of *A Woman of No Importance* (1893) and Lady Windermere in *Lady Windermere's Fan* (1892)—to maintain arbitrary decorum in the paradoxical hierarchy of ethics that condones male depravity and merely expose the double standard of Victorian sexual morality. Conversely, Shaw fully explores the social struggles of his strong-willed, autonomous female characters, such as Vivie Warren in *Mrs. Warren's Profession* (1893) and Gloria Clandon in *You Never Can Tell* (1897), to subvert discriminatory expectations of ideal womanliness and further encourage social reform for women's rights. While Wilde eludes a disruption of the established ethical hierarchy—which would overlook his

crimes of gross indecency in a heterosexual context—by confining the development of his female protagonists within the oppressive binary of sexual morality, Shaw actively thwarts patriarchal demands for feminine purity and noble domestic duty in developing various portrayals of authoritative women who gain dominance in the domestic and public spheres in order to promote equality among the sexes.

Works Cited

Foldy, Michael S. *The trials of Oscar Wilde: deviance, morality and late-Victorian society*. Wiltshire: Redwood Books, 1997. Print.

Mackie, Gregory. "The Function of Decorum at the Present Time: Manners, Moral Language, and Modernity in 'an Oscar Wilde Play'." *Modern Drama* 52.2 (2009): 145–167. Web. 10 November 2011.

Peters, Sally. "Shaw's life: a feminist in spite of himself." *The Cambridge Companion to George Bernard Shaw*. New York: Cambridge University Press, 1998. Web. 10 November 2011.

Shaw, George B. "Mrs. Warren's Profession." *Plays by George Bernard Shaw*. New York: Penguin Group, 2004. 30–98. Print.

Shaw, George B. "The Womanly Woman." *The Quintessence of Ibsenism*. London: Walter Scott, 1891. 31–45. Web. 10 November 2011.

Shaw, George B. *You Never Can Tell*. New York: Brentano's, 1905. Print.

Stead, W. T. "The Conviction of Oscar Wilde." *The Review of Reviews* June 1895: 491–492. Web. 19 November 2011.

Wilde, Oscar. "A Woman of No Importance." *The Best of Oscar Wilde.* Ed. Sylvan Barnet. New York: Penguin Group, 2004. 102–173. Print.

Wilde, Oscar. "Lady Windermere's Fan." *The Best of Oscar Wilde.* Ed. Sylvan Barnet. New York: Penguin Group, 2004. 38–99. Print.

The Other Side of the Struggle: Freeborn Abolitionists, Communication, Perseverance, and Brotherhood

Sarah Ahearn

On the eve of the Civil War, John Mercer Langston, born a free man in Virginia, addressed a tremendous crowd in Cleveland, Ohio. Langston, who went on to become one of the first African-Americans elected to public office in Ohio and the first African-American elected to Congress in Virginia, was an vivid political activist, lawyer, and abolitionist. In his lecture, entitled "The World's Anti-Slavery Movement; Its Heroes and Its Triumphs," Langston highlighted some of the most important contributions that Northerners had made to the abolitionist effort. Dubbing the people of the North the movement's "living orators," Langston stressed the importance of movements crafted at the hands of freeborn black men and women (Langston 62). "Our anti-slavery books… are not only read by all classes, in every section of this country, but are read and admired by the common people…The influence… is silent, yet potential, well-nigh omnipotent" (61–62). There existed a fire that was brought about by a people who chose to remain in our nation and to work within that nation's Constitution in order to change their world. This idea is best articulated in Langston's own words:

> Animated by the same spirit of liberty that nerved their fathers, who fought in the Revolutionary War and War of 1812, to free this land from British tyranny, they are inveterate and uncompromising enemies of oppression, and are willing to sacrifice all that they have, both life and prosperity, to secure its overthrow…To have brought the means of education and development to such a people, a people possessing such a spirit, is certainly one of the more desirable and valuable achievements of the American anti-slavery movement (63–4).

Langston was not alone in the fight for freedom. What has been sorely lacking in the study of the abolitionist movement of the North is the experience of the black men and women who had been born free, and the role that they took in freeing their race in the South. Those men and women lived in a time where it was a simple matter of chance that they themselves had not been condemned into a life of slavery; it was purely a matter of location that had spared them their lives. It is this particular circumstance that cultivated this group of people who were publicly connected to the abolition movement decades before Civil War erupted in America. The

experiences of freeborn blacks in the North help to fill in the holes of the history of abolition. Aside from being historically momentous, this movement is an intriguing story of perseverance and unity—one that is characteristic of the American drive for freedom that we today pride ourselves on. Their methods of protest set an important precedent for future civil rights struggles both within and beyond the black community.

There was not one singular opinion among northern blacks concerning slavery and the actions to be taken thereabout in the time period before the Civil War. In reality, there was significant stratification of beliefs. For example, there was a movement based on the idea that mass immigration to Liberia was the only way to establish a land of free blacks. That idea and its subscribers, however, did not change the course of history. The others, those who believed in remaining in America, the land of their birth, and working within the country to change unjust policies, did. David Walker, black abolitionist, wondered of those "colonizationists," "Do they think to drive us from our country and homes, after having enriched it with our blood and tears…?" (Walker 77). Walker, along with so many others, sought to change America rather than escape it. A monumental achievement of freeborn black abolition was development of a public celebration of black American history. Though the common plea was for a complete separation of North and South, many freeborn blacks felt the same sense of belonging to America as whites did. They asserted themselves as a pertinent part of their country, celebrated their own past, and looked confidently and boldly into the future. It was those individuals who rebelled against the unjust policies of the government by promoting violence and civil disobedience, unity through celebration, and pride in the black race. The celebration of the accomplishments of the abolitionist movement and of the overall role that blacks had played in American history, be it through literature, action, or speech, were thus vital to the motivation and success of the freeborn blacks of the North.

One of the means by which Northern blacks began to promote abolition was by communicating ideas across the nation through literature. In the mid-nineteenth century, a culture of reform in response to various social problems emerged, and began for the first time to use the print media to communicate their messages. Under this umbrella of social reform, an abundance of abolitionist newspapers came into existence. This fledgling social change facilitated the growth of the freeborn anti-slavery movement by providing an important outlet to spread awareness: the press. Revolutionary David Walker provides an example of an early utilization of print media and the effects that it can have on public opinion. *Walker's Appeal in*

Four Articles; Together with a Preamble, to the Coloured Citizens of the World, But in Particular, and Very Expressly, to Those of the United States of America was a pamphlet that was first published in 1829. Walker's pamphlet expresses his distaste with the institution of slavery and shocked many with its violent suggestions. *Walker's Appeal* brutally chastises not only slaveholders, but all white Americans. Walker boldly calls for the slaves to rise up against their masters, darkly predicting that white Americans "will yet curse the day that [they] were ever born" (Walker 82). As one of the earliest examples of black activism via the print media, Walker's pamphlet marked the beginning of a long period of change in the United States.[1]

The exact extent to which David Walker's pamphlet influenced public opinion is difficult to estimate. It was distributed across the nation, gaining success in some areas and becoming banned in others. What is apparent, however, is that Walker's words did not fall upon deaf ears. Reaching whites and blacks across the North and even into the Deep South, *Walker's Appeal* enraged and inspired an audience far more vast than Walker himself could have foreseen. Despite being chastised for its violent implications, Walker's opinions remained in the public discourse for quite some time, bringing the idea of abolition and violence further into the national spotlight. Credited with radicalizing the abolitionist movement, the repercussions of Walker's pamphlet are often connected with events such as the infamous Nat Turner rebellion of 1831. Word had spread throughout the South of Walker's ideas, creating a tense, fearful environment in slaveholding communities as the threat of insurrection loomed overhead. Though the two men never actually conversed, Walker's ideas reached and inspired Turner. The resonance of the written word shook America to its core as the only way of life many of its citizens knew began to change (Risley, 26–27; Harding 75–100).

The fear of a slave insurrection that ran rampant through the South was solidified on October 16, 1859, when John Brown led an army of men into Harpers Ferry in order to raid its military arsenal and to arm slaves, so that they could forcibly overthrow their masters. Fear spread throughout the nation following the bloody seizure, as newspapers such as the New York *Herald* reported an "Extensive Negro Conspiracy"[2] in the South. In regards to the action that the Richmond *Enquirer* dubbed one that had "advanced the cause of Disunion, more than any other

---

[1] John C. Inscoe, "David Walker, 1785–1830," Documenting the American South, http://docsouth.unc.edu/nc/walker/bio.html (accessed November 19, 2011).
[2] "Harper's Ferry," *New York Herald*, October 18, 1859. http://www.pbs.org/wgbh/aia/part4/4h1538b.html (accessed November 21, 2011).

event that has happened since the formation of the Government,"[3] Thomas Hamilton, journalist for the *Weekly Anglo-African,* published an article justifying the use of violence. Hamilton joined the chorus of voices who chastised the enslavement of human beings and the practice of regarding blacks as property: he justified the actions of Brown and his men by pointing out that humans are not property, and therefore any attempts to "restore stolen chattels to their proper owners—to restore the slave to himself" are morally justified (Hamilton 41–42). It is words such as these that exemplify the fledgling sense of camaraderie and brotherhood that drove the stories of these abolitionists. While most newspapers in the country were focusing on the terror and the fear that John Brown's raid on Harpers Ferry had brought upon the South, freeborn black men such as Hamilton spoke for the cause, for the abolitionists. From Walker to Turner to Brown, the movement began to swell, fueled by years of injustice that could only be combated by an unyielding unified force.

Violence was naturally not the only form of disobedience that drove abolitionists. There has been a longstanding tradition of civil disobedience in America, especially in the struggle for black civil rights, which saw its beginnings during the era of the abolitionists. Abolitionists who chose to be lawfully acknowledged carried the theme of civil disobedience from international slave discourses into America. In 1849, philosopher Henry David Thoreau published his infamous discourse on civil disobedience. Thoreau's reverberating declaration for "not…no government, but at once a better government," to be obtained by letting "every man make known what kind of government would command his respect," which will "be one step toward obtaining it," was embraced by the next wave of freeborn blacks (Thoreau 2).

When the Ohio Supreme Court decreed that mulattoes or any voter of African descent did not have the right to vote, the Ohio Anti-Slavery Society fought back. William J. Whipper was one of the many who attempted to vote in town elections on April 4, 1859. He recounted his actions in a letter to the editor of the Salem *Anti-Slavery Bugle,* which was then published on April 23. Drawing upon the state's legislature regarding voting, Whipper detailed his belief that the law accommodated his right to vote by not expressly prohibiting it. In a few passionate lines, Whipper summed up his determination to the cause of abolition and the establishment of basic human rights, as well as his disdain for the Supreme Court of Ohio:

---

[3] "The Harper's Ferry Invasion as Party Capital," *Richmond Enquirer*, October 25, 1859. http://history.furman.edu/editorials (accessed November 21, 2011).

> I guess they will conquer, but conquer who? not the Mexicans, nor the Brighamite Mormons. Well, then, who is it? A colored man in Charlestown, Portage County, Ohio, who stands armed with the statute of Ohio, and a determination to insist upon his rights, notwithstanding the established custom of past years. So come on with your forces, for being quiet is no part of my mission (15–17).

Whipper's passion for the abolitionist movement and drive to spread awareness and unity is evident in his choice to publish his story. His disobedience represents one instance of many, where freeborn blacks used the law to defend themselves against the tyranny of oppression. These men who believed in fighting for civil rights in the North as well as the South saw that the introduction of black voters would inevitably aid the cause of abolition. They understood that unity in the face of adversity would change their world. The struggle for abolition among freeborn blacks may not have encompassed one form of disobedience, or purely civil disobedience for that matter, but did draw upon the same moral framework. Every article that resembled Whipper's sent shock waves through the country. The movement was dynamic, growing with each action, threatening to irrevocably alter American society.

At the very core of the movement stood the all-encompassing moral elevation of the black race. The betterment and unity of the black race as a whole would lead ideally to the freedom of those enslaved in the South. In 1860, John S. Rock, a black lawyer and anti-slavery activist, spoke before a large audience in Boston, Massachusetts, addressing such ideals. William C. Nell, another freeborn black abolitionist, had organized the rally for the purpose of celebrating black involvement in American history, particularly recalling the memory of Crispus Attucks. Attucks's death in the Boston Massacre symbolized black involvement in revolutionary history to these men and women who gathered annually to honor his memory.

Speaking to a crowd much larger than the traditional town hall meetings encountered previously in American history, John S. Rock sought not only to commemorate Attucks, but also to connect that black revolutionary spirit to the present day. In lieu of the Dred Scott decision, Rock called for further activism and perseverance. He stated, "I believe in insurrections—and especially those of the pen and of the sword" (Rock 73). Rock believed John Brown to be something of a second coming of Crispus Attucks, as both were men who had started a revolution. It was the revolutionary spirit of Brown's raid on Harpers Ferry that represents Rock's "sword." Anti-slavery publications and written discourse, the "pen," symbolize the other leg of Rock's belief. Organizations and celebrations such as the one organized by Nell instilled unity and pride in the members of the black community who were among the most important

figures in the fight against slavery. It was morally just and sound that the colored men and women of America fight back against the oppression cast upon them by the slaveholders of the South and the government; it was only by presenting a united front that the black race would emerge prosperous (58–66).

Thomas Hamilton joined Rock and the many others in the call for black unity. In an 1860 publication of the *Weekly Anglo-African*, he demanded the betterment of the black race and for the education of the masses in the ways of politics and abolition. "We must rely on ourselves," he stated (Hamilton 73). Hamilton, as previously discussed, advocated for violence in the years prior. His later article is not reflective of his separation from radical ideas, but rather it is representative of the all-encompassing nature of black pride and solidarity. The push for the celebration of black memory throughout history and the unity of the black race in the face of adversity were two of the most common themes in the freeborn black abolitionist discourse. It was the theme that united the push for violence and civil disobedience and that mended the split in ideologies. Encompassing moral and political elevation, the unity of the black abolitionists and the spreading of their message to the masses though speech and print culture are strongly intertwined the with success of the abolitionist movement in America.

The passion and the diligence that drove the freeborn anti-slavery movement is one that is characteristic of the American spirit. The effects of freeborn black activism led to a widespread awareness of the condition of blacks in the South, thus ultimately fueling America's descent into Civil War. Their activism spread education among blacks in both the North and the South, uniting the race and fueling rebellion. The means by which these protests and stories of activism reached such a broad audience is through the conscious effort on the part of the abolitionists to record, publicize, and preach their message. John Mercer Langston said of the anti-slavery movement: "[I]t has brought the subject of slavery itself distinctly and prominently before the public mind" (Langston 59). The spreading of awareness of the struggles endured by both slaves and free men fighting for abolition was a large contributing factor in the dissolution of North and South. Without acting as they had, freeborn abolitionists may not have had such a tremendous effect on the fate of the United States.

The study of freeborn black activism in the United States opens the door to many other areas of research. The men and women who contributed to this cause were passionate and determined. They evoked various methods of protest that one can see reflected in the struggle for black civil rights throughout American history. Racial pride and solidarity have proven to be

successful and powerful tools in the face of adversity. They bridge the gap between violent rebellion and civil disobedience; they mend the cracks between those who fight and those who are oppressed. Their demands were met with bloody civil warfare and begun the long struggle for equality in America. Their collective story is largely untold, as they were simply one part of a larger machine. However, the personal and ideological movement that was freeborn black activism is deserving of recognition; it is imperative to understand the other side of the struggle.

Works Cited

Primary Sources

Anderson, Osborne P. *A Voice from Harper's Ferry with Incidents Prior and Subsequent to Its Capture by Captain Brown and His Men: A Narrative of Events at Harper's Ferry.* Boston: Printed for the Author, 1861. http://books.google.com/books (accessed November 13, 2011).

Boyd, Herb. *Autobiography of a People: Three Centuries of African American History Told by Those Who Lived It*. New York: Doubleday, 2000.

"Constitution of the New-England Anti-Slavery Society Together with Its By-Laws, and a List of Its Officers." Boston: Garrison and Knapp, 1832. OpenLibrary.org. (accessed October 28, 2011).

Langston, John Mercer. *Freedom and Citizenship; Selected Lectures and Addresses*. Miami, Fla: Mnemosyne Pub, 1969.

Newman, Richard S., Patrick Rael, and Phillip Lapsansky. *Pamphlets of Protest: AnAnthology of Early African-American Protest Literature, 1790–1860.* New York: Routledge, 2001.

Redmond, Charles Lenox. "An Anti-Slavery Discourse." Speech, Address to the   Massachusetts Anti-Slavery Society, Boston, Mass., July 10, 1857. TEXT file. http://www.blackpast.org (accessed October 28, 2011).

Ripley, C. Peter (ed.). *The Black Abolitionist Papers: Vol. V: The United States, 1859–1865.* The University of North Carolina Press, 1992.

Thoreau, Henry David. *Civil Disobedience and Other Essays.* Dover Publications: Mineola, N.Y., 1993.

Walker, David. *David Walker's Appeal, In Four Articles: Together with a Preamble to the Coloured Citizens of the World, but in Particular, and Very Expressly, to Those of the United States of America,* eds. Sean Wilentzhill and Wang, New York, 1995, http://www.psb.org (accessed November 14 2011).

Secondary Sources

Anderson, Michael John. *Anger and Detachment: A Study of Arden, Osborne, and Pinter*. London: Pitman, 1976.

Harding, Vincent. *There Is a River: The Black Struggle For Freedom in America*. New York: Mariner Books, 1993.

Inscoe, John C. "David Walker, 1785–1830." Documenting the American South. http://docsouth.unc.edu/nc/walker/bio.html (accessed November 19, 2011).

Leeman, Richard W., ed. *African-American Orators: a Bio-Critical Sourcebook*. Westport, Conn.: Greenwood Press, 1996.

Risley, Ford. *Abolition and the Press: The Moral Struggle Against Slavery*. Evanston, Ill: Northwestern University Press, 2008.

"World Association of Newspapers: A Brief History." World Association of Newspapers. http://www.wanpress.org/ (accessed October 30, 2011).

Too Complex a Topic: Rewriting White South African Identity in J. M. Coetzee's *Summertime*

Kate Harlin

The genre of life-writing in its current form has been evolving since Jean Jacques Rousseau's *Confessions* was published after his death in 1782. It is an elusive and amorphous category of texts that is somewhat difficult to describe and define adequately. It serves many varied purposes, especially in a post-colonial context; this includes not only an attempt to write the self, but in many instances, an attempt to write the historical, social, and cultural contexts in which the self is located. As a country, South Africa has had a unique struggle with the concept of cultural identity in the two decades following the end of the racist apartheid regime. The white minority, the historical colonizer in the country, has since lost power to the black majority and has had to struggle to find its place in an evolving society. In his memoir *Summertime,* one of J. M. Coetzee's characters suggests that Coetzee never confronted the problem of white identity in South Africa because "it might have seemed too complex a topic to be explored in a memoir […]—too complex or too close to the bone" (209). Ironically, in this paper I will explore how *Summertime* is a self-conscious attempt at rewriting white South African identity post-apartheid and how Coetzee eventually abandons not only the writing, but white South African identity as a whole.

J. M. Coetzee's *Summertime* is as much about the genre as it is about identity; therefore, its unique use of life-writing must be understood even before examining how it creates identity. Following *Boyhood* (1997) and *Youth* (2002), it is the final installment in an autobiographical trilogy. While the previous two texts adopt a third-person narrator utilizing present tense, they still remain well within the confines of autobiography or memoir. The structure creates a temporal distance between the narrator and his author/subject, which does not exist in traditional autobiographical writing; however, the reader is never confronted with contradictory information that calls into question the validity of the content. Even a reader familiar with Coetzee's life can accept that *Boyhood* and *Youth* are separate endeavors from Coetzee's fiction, and unmistakably

autobiographical. *Summertime,* though, presents the reader with an immediate problem. The first and last section are journal entries, written in Coetzee's now characteristic present tense third person, but they frame what is presented to the reader as a series of interviews of five individuals who knew an author named John Coetzee in South Africa in the 1970s. The interviews are being conducted by an academic who is referred to fleetingly as Mr. Vincent. The John Coetzee of *Summertime* is dead, and Mr. Vincent is writing his posthumous biography. Further complicating its relationship to the previous two memoirs and its validity as life-writing, Coetzee intentionally falsifies much of the account of his life in the 1970s. However, as several scholars have recognized, Coetzee is much less concerned with facts than with communicating the various complex and often contradictory elements that comprise a real person's identity. As a result, he puts his confidence in the power of the literary over the historical (read: factual) narratives surrounding a public figure's life.

Interestingly, though the present of the narrative of *Summertime* is 2008, nearly two decades after the end of the apartheid regime, and during the ongoing emergence of the "New South Africa" and a Rainbow National identity, the stories actually takes place between 1972–75, perhaps one of the most violent and oppressive periods of South African history. However, both author J. M. and persona John Coetzee emigrated permanently to Australia in 2002, and though it is not addressed in *Summertime*, the real J. M. Coetzee went so far as to become an Australian citizen in 2006. Coetzee and all of the characters of *Summertime* experienced apartheid and its eventual demise; therefore, by setting the main narrative during a period when he was truly living in South Africa, but retaining a contemporary vantage point, Coetzee manages to set a post-apartheid text during apartheid.

How can a critic differentiate between novel and memoir in this instance, when the fictionalized self is created as a representation of the actual self? In an interview with David Atwell, Coetzee explains that in his mind, "Writing is not free expression. There is a true sense in which writing is dialogic: a matter of awakening the countervoices in oneself and embarking upon speech with them. It is some measure of a writer's seriousness whether he does evoke/invoke those countervoices in himself." He uses the term "monologic ideal" to describe this flawed notion of authenticity in the supposed non-fiction genre. Matthew Cheney compares Coetzee's text to *A Million Little Pieces,* James Frey's infamously fabricated memoir of addiction; with the adoption of the third-person narrative, Coetzee effectively creates a

persona—one that retains his name and characteristics, but is essentially a literary character meant to represent, rather than be the self. Because of this, the casual reader does not demand the same "honesty" from *Summertime* as he or she would from a more conventional piece of life-writing. But for the literary critic, this still presents a difficulty. Whereas memoir and other traditional life-writing require consistency—the creation of a defined self with identifiable values, characteristics, ideology, etc.—*Summertime* is a collage of different voices, all of which have different and even conflicting views of the self being written, but all of whom, most importantly, are written by Coetzee to construct his identity, however externalized that identity may be. For Coetzee, this strategy offers the reader a greater degree of insight into the complex reality of his life. Facts become subordinate to ideas, and the internal dialogue (or perhaps more appropriately, "polylogue") creates a self-subject which embodies the instability of self-identity. Though perhaps *Summertime* cannot be fairly characterized as memoir, it is without question life-writing. What kind of life-writing, and how the genre informs the interpretation of the John Coetzee constructed in the pages of *Summertime* is the subject of much scholarly investigation—including mine.

This analysis of genre and form is especially integral to the interpretation of *Summertime*'s construction of a white South African identity because the meticulous formal manipulation of the text reflects the level of self-consciousness with which Coetzee approaches the project, and therefore the representation of his self. John Coetzee is a writer but he is also a white South African, and this aspect of identity is full of diffuse and sometimes paradoxical meanings, both in Coetzee's own life and writing, and in the South African social, cultural, and political landscape as a whole. J. U. Jacobs describes how Coetzee's relationship to his white identity is reflected in Coetzee's penchant for *chiasmus* throughout his works. Jacobs argues that not only does Coetzee employ this device regularly, but that such rhetorical behavior reflects how Coetzee interprets the world and his own identity. This interpretation is far too reductive, as it implies both an assumption of a fixed dichotomy between Afrikaner and English, and an orderly reconciliation of the two; however, what Jacobs does illuminate is the detachment Coetzee feels from both of these white identities. Where Jacobs identifies dualities created by Coetzee, suggesting a resulting identity that incorporates these oppositional forces, holding them together like a seam, it appears that Coetzee's relationship with his white South African identity is far less cohesive.

Summertime provides the reader with several perspectives on John's identity, and it is notable that the greatest amount of authority, at least in terms of page numbers, is divided between Julia Kis, the daughter of Hungarian Jewish immigrants with whom Coetzee has an affair, and his Afrikaans cousin Margot Jonker, with whom he grew up on his grandfather's farm. The three subsequent perspectives are considerably shorter and narrower in their capacity to contribute to John's identity: Martin, a former friend and colleague at the University of Cape Town literature department; Adriana Nascimento, a Brazilian refugee who is the mother of one of John's students whom he pursued persistently and ultimately without success, and Sophie DeNoel, another university colleague and woman with whom Coetzee had an affair.

Julia's interview with Mr. Vincent is the first of the five interviews, and so it is up to Julia to introduce John to the reader. As a Jew and an immigrant unhappily married to a South African businessman during the period of her affair with Coetzee, Julia ought to be an outsider; yet in the apartheid era of politicized identity, she is classified as white, and experiences the privilege of that status. She presents herself to Mr. Vincent as extraordinarily self-assured, empowered, and erotic, though she does hint that in her first marriage and during the affair with Coetzee, she has been emotionally immature and naïve. Her portion of the narrative is characterized not only by her intense sexuality, but also by John's hopeless lack of sexual aptitude. This serves as an interesting contrast to the other white South African men Julia describes, such as her husband and his business associates, whom she portrays as uniformly adulterous, greedy, and superficial. Referring to her husband's colleagues as "men who knew about systems, about which systems are sustainable and which are not" (27), and eventually condemning all white South African men as speaking "as if there were some conspiracy they all belonged to that was going to create a fake, trompe l'oeil future where no future had seemed possible before" (53). She suggests that it is her identity as an outsider, however insignificantly it may have affected her daily life, that allowed her the distance necessary to glean how her husband and his peers were not only reinforcing the oppressive system, but benefiting from its continuation, and plotting to abandon it when the inevitable chaos began. At least according to her own account, Julia was the only member of her social circle not willingly complicit in its continuation (54).

John also seems to understand South African identity in racial terms. Julia is the first of all four women in *Summertime* to comment upon John's physical awkwardness and

unremarkable ability as a lover. Later, Adriana will suspect that "perhaps this is how Dutch Protestants behave when they fall in love: prudently, long-windedly, without fire, without grace" (172). John himself attributes his clumsy ways to a fundamental Europeanness. When John meets Julia he is living with his father in a suburb of Capetown, and in spite of the disconnect which seems to exist between his mind and body, he is endlessly improving his father's house which has fallen into disrepair. He explains to Julia many times that he is doing the work himself not simply because it is cost-effective, but because he wishes to "overthrow the taboo on manual labor" for white South Africans (61). John's insistence on doing his own construction projects is mentioned several times throughout *Summertime,* yet the political implications of the act are never addressed. Rather, it appears that it is more for the whites' benefit than blacks', as if he believes that the physical labor will rectify in him what he views as the soft, emasculated quality brought about by employing generations of Africans to do all of the country's heavy lifting. Julia recognized John's condemnation of the indulgent white lifestyle in his message, confessing that "the paid labor of [her] black domestic set [her] free to have idle affairs with strange men" (61). Yet somehow this acknowledgement of apartheid inequality is not a demand for egalitarianism, but a lament that John, as a white man in Africa, has been deprived of the privilege of laying concrete. It is this aspect of white South African male identity that is most distasteful to Julia. She even criticizes their identification with the Jews:

> For instance, white South Africans in those days like to think of themselves as the Jews of Africa: […] cunning, unscrupulous, resilient, running close to the ground, hated and envied by the tribes they ruled over. All false. [They] were not tough, they were not even cunning. [They were] a tribe of babies looked after by slaves. (54)

Whereas John identifies as a victim, born into a system he has not created, on a continent where he does not belong, Julia sees him and his counterparts throughout the country as helplessly reliant on the status quo and incapable of surviving alone in Africa. Sophie echoes this problematic tendency John has, to understanding racial identities in terms of stereotypes. She explains that "he saw Africa through a romantic haze. He thought of Africans as embodied in a way that had been lost long ago in Europe" (231). The cerebral and awkward John feels out of place in Africa throughout his life, and it appears that he attributes this discomfort at least in part to his whiteness. To John, there is something intrinsically African about the black people of the country which he does not possess and cannot hope to attain.

While it is easy and tempting to censure John entirely for claiming to be victimized by apartheid while being its primary beneficiary and taking no action toward political change, it is vital to remember that the John interpreted by Julia is only one sixth of a character. Margot, John's Afrikaans cousin and childhood confidant, reveals to the reader a much more sympathetic man. Margot has at an earlier date recalled a holiday gathering on her grandfather's farm, Voelfontein, and Mr. Vincent has subsequently taken her story and constructed a narrative from it. The text in Margot's chapter of *Summertime* is Mr. Vincent's interpretation of her story, occasionally punctuated by Margot's protests at his creative liberties. It is significant to note that, despite the discomfort Margot occasionally expresses, the form and ethos of this section most closely reflect those of Coetzee's previous autobiographical works; Margot is also the only voice present in *Summertime* that has known John throughout his life. Therefore, it is quite likely that the John constructed by Margot is the most comparable to the self that Coetzee would construct, had he followed conventional life-writing methods. This John is experiencing the homelessness and unbelonging associated in post-colonial theory with members of a diaspora—and of course, John is such a person. Margot offers the reader a detail of John's childhood that is central to his identity: he is an Afrikaner by heritage, but grew up linguistically English. It is this fact that Julia is vaguely aware of but unable to appreciate as an outsider, which causes John to identify as a victim, rather than a perpetrator, of colonialism. On the family farm, he is met with disdain from Margot's sister and her family who claim that "from the heights of his *englese* education […] John looks down on the Coetzees" (90); but when he returns to the farm to visit with this family, he speaks Afrikaans, however difficult and unnatural the language has become to him as an adult. Margot, perhaps the most affectionate of John's friends throughout *Summertime,* even wonders at his identity: "Does he really consider himself an Afrikaner? She doesn't know many real Afrikaners who would consider him one of the tribe […] nowadays at the very least you need to vote National and attend church on Sundays" (95). In reality, Margot has more in common linguistically with the black farmhand Hendrik, who can speak Afrikaans with an informal, conversational ease. Ironically, John's Afrikaner name causes Adrianna Nascimento to argue that he is unqualified to tutor her young daughter in English, asking the girl, "Can't this school find a proper English teacher? I want you to learn proper English from a proper English person" (157). When John defends his qualifications to Adriana, he ends with the statement, "English is just one language among many" (161), emphasizing that English as well as Afrikaans

are foreign languages to him. Linguistically dispossessed, John confesses to Margot that he has been learning Hottentot, the language spoken by the nomadic Khoi tribes of Southwest Africa, a dead and nearly forgotten language of the region. Margot cannot understand his interest in learning it, as opposed to Xhosa or another language still widely spoken by the Africans of the region. The answer he offers—"I am interested in the things we have lost, not the things we have kept. [With Hottentot] you can speak to the dead" (104)—does not resonate with a woman who has grown up certain in her Afrikaner identity, among those who share her language and culture, and on land which has belonged to her family for generations. John, alternately, experiences none of those grounding certainties. As such, he identifies with the Hottentot people; they have not vanished simply by being slaughtered to extinction by colonizers, they have been absorbed by another ethno-linguistic group, and their identity has been corrupted a bit more each generation until it disappeared. John's anxiety is that he is an intermediate, that his culture has been so diluted that it is indiscernible.

It is important to remember that Coetzee warns against the monological ideal because as a life-writer, he desires license to contradict himself. Although John views his own identity as being in crisis, Sophie reveals that John awaits a Utopian day "when family histories would have become so tangled and intermixed that people would be ethnically indistinguishable" (233). Notwithstanding the alarming undertones of social engineering inherent in this aspiration, John's idealism is coupled with an ardent aversion to engaging with the politics of the moment. When Mr. Vincent suggests John has been apolitical, Sophie disagrees, characterizing him instead as "anti-political. He thought politics brought out the worst in people" (228). She explains his feeling that "politics is a symptom of our fallen state and expresses that fallen state. […] Even the politics of liberation" (229). In fact, for someone who would like to project an image of John that is opposed to politics, she is extraordinarily familiar with his political attitudes during their relationship, claiming with confidence that John had been opposed to the liberation struggle in South Africa—a particularly inflammatory suggestion to be made in 2008 about a beloved national figure. She does qualify the statement by claiming he had believed that "the struggle was just but the new South Africa toward which it strove was not Utopian enough for him" (230). However, Sophie's interview, the final voice before the bookend of John's journals, leaves the reader with a surprisingly naïve and potentially alarming image of John as a white South

African who is indifferent to the suffering of his fellow countrymen because their revolutionary goals do not reflect his own.

Summertime is an incontrovertibly South African text; however, it begins shortly after John's homecoming from the United States, under a cloud of suspicion. In fact, his distance from South Africa is one of the most fundamental pieces of his white South African identity. As Sophie observes with uncharacteristic astuteness, "[John] was happiest in the role of outsider. He was not a joiner" (239). It is worth noting that three of the five interview subjects were born outside of South Africa, and Martin has since left for England, leaving Margot as the only "true" South African. In fact, Martin describes the strained relationship he and John shared with their birthplace as white South Africans, noting how "our presence there was legal but illegitimate […] Whatever the opposite of native or rooted, that was what we felt ourselves to be" (209–10). This guilt and uprootedness, as described earlier, result from their membership within a diasporic community. Martin attributes their individual decisions to leave South Africa to this identification. Another potential explanation for Sophie's anti-political characterization of John is John's divestment from South Africa, which likely took place before they ever met. As Martin explains, "We were reluctant to invest too deeply in the country, since sooner or later our ties to it would have to be cut, our investment in it annulled" (211). It is this image of severing ties that triggers within John the greatest emotional response. While on holiday at Voelfontein, he reflects on how the vast expanses of landscape which surround the family farm "fill him with melancholy [and] spoil him for life" (97). Regardless of his difficult relationship with the culture of South Africa, the natural geography still conjures in him a response comparable to Wordsworth's sublime, and it is this aesthetic that binds him more than anything else to the land. John's inability to fully realize a South African identity, however, inspires in him the resolve which leads eventually to his permanent expatriation. It is abundantly evident that this is not a peculiar insecurity in John's psyche, but instead indicative of a shared South African existence. Margot, the quintessential South African of *Summertime,* is driven to wonder, "Why are we spending our lives in dreary toil if it was never meant that people should live here, if the whole project of humanizing the place was misconceived from the start?" (140). Still, Margot has spent her whole life on the same piece of land, speaking her Afrikaans language with her Afrikaans family. She is an Afrikaner first and a South African subsequently, and this secure identity prevents her from ever leaving. No amount of encouragement from John or misfortune in health

or finance will uproot Margot, whose own destiny is inextricably linked to the future of her adopted homeland. This resignation to homelessness in the face of a nation that will soon belong even less to the whites than it does in 1972 is why John's decision to leave is met with resentment from much of his family. "Not without reason, the Coetzees took it to mean he had disowned his country, his family, his very parents" (131); if the Coetzee clan is doomed to work their whole lives in the foreign soil on which they were born, the same should be expected of John. This is the essence of Coetzee's South African identity, the inheritor of crimes they did not commit, on land they did not conquer, amidst a war they did not begin. In his final journal entry, John is faced with whether to sacrifice his personal life to nurse his ailing father.

John's relationship to the old man is in many ways a reflection of his relationship with South Africa; his father is an Afrikaner who listens to German opera and belongs to a previous generation. Ultimately, *Summertime* ends with John's final address to his father: either he must dedicate himself to his ailing parent until he eventually dies, or explain to him, "I cannot face the prospect of ministering to you day and night. I am going to abandon you. Goodbye. One or the other; there is no third way" (266). Whiteness in Coetzee's South Africa is as much an inadequacy as it is a privilege. Ultimately both John and J. M. must decide whether they will keep up the charade of becoming a South African in the face of the new "Rainbow Nation." In the end, both John and J. M. are unable to reconcile their white South African identity, and since they cannot renounce their whiteness, they abandon their country.

In the end, Coetzee is unable to write for himself an acceptable identity as a white South African. As Sophie tells Vincent, "[John] believed our life stories [are] ours to construct as we wish within or even against the constraints imposed by the real world" (227). This fluidity of identity eventually gives Coetzee license to abandon the project of white identity construction in the New South Africa, and declare himself Australian in spite of his own ethnic and historical reality. Perhaps it is this failure, to fully understand and constitute himself in the world into which he was born, that accounts for the impotence and disembodiment which characterizes his persona. John, after all, is not a white South African at all but "a fiction of himself" (225).

Works Cited

Cheney, Matthew. "Intentional schizophrenia : J M Coetzee's autobiographical trilogy and the falling authority of the author." *The Quarterly Conversation.* The Quarterly Conversation, 7 Dec 2009. Web. http://quarterlyconversation.com/intentional-schizophrenia-j-m-coetzees-autobiographical-trilogy-and-the-falling-authority-of-the-author

Coetzee, J. M. *Summertime*. Thorndike, Me.: Center Point Pub., 2010. Print.

Farred, Grant. "Autobiography." *South Atlantic Quarterly* 110.4. (2011): 831–847. Print.

Jacobs, J U. "(N)either Afrikaner (n)or English : Cultural cross-over in J M Coetzee's 'Summertime'." *The English Academy Review* 28.1. (2011): 39–52. Print.

Neuman, Justin. "Unexpected cosmopolitans : media and diaspora in J M Coetzee's 'Summertime'." *Criticism* 53.1. (2011): 127–136. Print.

Parks, Tim. "The education of 'John Coetzee'." *The New York Review of Books.* The New York Review of Books, 11 Feb 2010. Web.

XRF Analysis of Maya Ceramics
at Tipu and San Pedro, Belize

Molly R. Carney

Whenever the Maya are mentioned, large sweeping pyramids and vast jungle cities often come to mind. Most people associate the Maya with the Classic period, which ranged from A.D. 300–900 and is often considered the height of Maya civilization. The Maya "collapse" took place around A.D. 800–900 and constituted a decline and abandonment of Maya centers in the southern lowlands. However, the Postclassic period marked a time of florescence in the northern lowlands. Tipu, located near the Belize-Guatemala border in western Belize, and San Pedro, located on Ambergris Caye off the coast of Belize, are both Maya sites with Postclassic period Maya occupation. The Postclassic period ranged from A.D. 900–1521 and marks the time after the Classic period collapse up until Spanish contact. Long thought of as a period of cultural decline, the Postclassic Maya period is now believed to be one of exchange, commercialism, and extensive trade (Aimers personal communication 2011). There is evidence of this extensive trade along the Caribbean coasts and rivers, and inland at sites such as Mayapan and Chichen Itza.

According to Rice (2012) a study of a collection of archaeological pottery generally proceeds in three steps: classification, analysis, and finally interpretation and explanation. For the purpose of this study, we are attempting to chemically classify examples of two closely related types—Payil Red and Palmul Incised from the Red Payil Group of Tulum Red Ware excavated at the sites of Tipu and San Pedro. While these two types appear stylistically very similar, do they have the same chemical signatures? By analyzing the chemical makeup of these pottery samples we hope to see if the stylistic classification corresponds with the chemistry. The ultimate purpose of this study is to assess and compare Tipu and San Pedro pottery sherd samples through stylistic classification and chemical analysis to further understand trade and exchange routes in the Maya Postclassic.

Classification of Pottery

Why is it important and necessary to categorize pottery sherds? Classifications help to organize and structure data. Pottery can provide clues to the culture, history, trade and exchange routes, and time period of the people who used it. Pottery, when classified consistently, can provide a reliable method of achieving relative or correlated dates for time periods across Maya settlements. This can help us to better understand chronology and be used a basis for inferences about Maya life (Smith, Wiley, & Gifford 1960:330). It is important to note that there is no "correct" way to classify pottery—we categorize ceramics in order to establish potential answers to analytic questions. In this study, we used type-variety to categorize the ceramic sherds. Type-variety has been the dominant classification system for the analysis of Maya pottery for many years.

Essentially, the type-variety system is a hierarchical system of classification used to stylistically characterize pottery (Rice 1982:282). It was first used in the American Southwest and then adapted to Maya pottery. The type-variety system has four major levels: ware, group, type, and variety. Ware is a broad and inclusive category defined by the pottery's finish or paste/fabric. Groups are used to aggregate or "lump" types that are very closely related and have similar attributes, such as the Payil Group under investigation here. The third category, type, is where classification gets more specific and is what our study is concerned with. Types are defined by distinct clusters of attributes that are indicative of a particular class of pottery produced during a specific time interval in a specific region (Smith, Willey, & Gifford 1960:333). Type names are usually binomial with a place name as the first part of the type and a descriptive term as the second part of the type name. For example, the type name Payil Red follows these rules of a location and a descriptive word. Finally, varieties are the smallest meaningful unit in classification. They are subsets of types (or subtypes) based on one or more minor attributes. Varieties may be very local or limited in time. Previously, all classification work on pottery sherds at Geneseo was done through type-variety classification.

However, this study uses elemental and mineralogical analytical techniques in an attempt to examine the assigned classifications. Mineralogical analysis involves looking at the larger crystalline components of the ceramic paste (Rice 1987:372). X-ray diffraction is one method used to determine the mineralogical composition of samples. X-rays are shot at a powdered

sample and the angle of reflection is measured and used to determine the minerals within. While X-ray diffraction is a proposed part of the study, the pottery samples have not yet been run.

This study instead focused on assessing the chemical composition of the ceramic sherds. Chemical analysis identifies the chemical elements or compounds constituting the ceramic (Rice 1982:372). X-ray fluorescence (XRF) and instrumental neutron activation analysis (INAA) are the two main methods of determining elemental composition. However, as INAA is extremely costly and needs a nuclear reactor, which is not available at SUNY Geneseo, this study used XRF to analyze the composition of the sherds. X-ray fluorescence is a technique used to identify the chemical composition of a sample. First, a beam of x-rays is aimed at the sample. The electromagnetic radiation of the x-rays can cause the elements in the sample to become ionized. The energy from the radiation causes an electron to be emitted, and energy is released. This energy, also emitted as radiation, is called fluorescence. The amount of energy emitted is different for each element, thus allowing us to analyze the levels of fluorescence to identify the elemental composition of the sample. This method works well for elements Na (atomic number 11) to antimony (atomic number 51) (Shackley 2010). By studying the levels of energy emitted, we can determine which elements are present in a sample.

Nineteen samples from Tipu and 20 samples from San Pedro were included in this study. The majority of the samples, as previously mentioned, were of the closely related Red Payil or Palmul Incised types. Payil Red is a plain red type of Maya pottery and Palmul Incised is simply its incised version. These two types are believed to have been produced along the coast of the Mexican state of Quintana Roo at the sites of Ichpaatun, Tancah, and Tulum (Sanders 1960). These similar types were chosen in order to compare stylistic classifications with the elemental signatures of the ceramics. These types are so consistent that there are only so many stylistic, macroscopic details that can be assessed, and this study is designed to assess the chemical variability within these two related types. As these types are so closely related, there are no defined varieties underneath the types. It will be interesting to see if the chemical analysis will allow us to determine separate technological varieties under the stylistically similar types. Three extra samples from Tipu that fell outside the Red Payil Group were also analyzed. These samples, T-17A and T-17B, of the Augustine Red type, and T-18, which is Pozo Unslipped, were included to see how much the chemical results varied for sherds that are macroscopically quite different from Red Payil Group sherds.

Why Study Pottery Sherds?

The overarching question remains—what is the purpose of this study? Why is it important to understand similarities and differences in ancient Maya pottery? What is the grand significance?

Scientifically speaking, the purpose of the study is to determine whether the assigned ceramic classifications correlate with the chemical composition of the sample sherds. This study is an effort to see if the stylistic and chemical classifications align. Because the pottery sherds under investigation here are common and widely distributed, it will be interesting to see if chemical subtypes can be created from the macroscopically visible main types.

Yet there is much more to learn from the pottery than just the chemical composition and how that changes classification. Through studying the results of the elemental analysis we also hope to discover a "recipe" for the ceramics, which might shed light on the original potters' decisions regarding their work. This allows archaeologists to gain insight into the consumption and production of Maya ceramics. Through this study, we hope to further understand Maya consumer choices. How were certain pots made? How many different recipes were used? How far were they traded? Through results from this study and further analysis we hope to answer questions like these.

A final application of this study is to assess exchange and trade routes throughout the modern-day country of Belize as seen in the Maya Postclassic period. Ethnohistorical research indicates that Tipu was closely linked with various independent sites of the Peten Lakes region of Guatemala and other burgeoning Postclassic sites (Aimers personal communication 2011). These connections indicate extensive trade in a time period once called a "decline" or "collapse." As archaeological material from this time is relatively rare, the Tipu and San Pedro collections offer great potential for research and analysis. So, this research is one part of a much larger investigation that will involve multiple types from many sites, and eventually comparison with the distribution patterns of obsidian.

Sample Preparation

Geneseo's XRF spectrometer requires samples to be ground into powder and fused into either a glass bead, for identifying major elements, or a wax pellet, which is used to identify trace elements. The first step for each sample involved removing all materials from the surface of the sherd. This was done in order to analyze the elemental composition of the actual pottery body or

paste, and not any other foreign substances on the surface or the applied slip. For each sample, I used a silicon carbide burr to "sand" and remove the slip or any other materials from the surface. Slip is a fluid suspension of fine clay and water used to coat a ceramic vessel before firing and is often considered a stylistic classification attribute (Rice 1987:482). After this was completed, I broke off a piece of the sherd of about 12 grams using pliers. I cleaned the broken pottery to remove any soil and debris, and then broke up the sample into smaller pieces.

The next step involved grinding the sample into a fine, homogenized powder. The broken sample pieces were placed into a metal shaker with two metal balls inside to crush the pottery. The sample was placed into a ball mill and run for about four minutes. After that time I checked the sample to see if all pieces had been ground into a fine powder. If there were any solid pieces left, I ran the sample for another four minutes to ensure that the sample was fully powdered. The powdered sample was then added to a labeled plastic vial, and the ball mill shaker cleaned with sand before a new sample was run.

After powdering, I created the glass bead and wax pellet samples that would be run in the XRF spectrometer. Glass beads are used in XRF to identify the major elements composing the pottery, while wax pellets are used in identifying the trace elements. For the glass beads, six grams of flux was measured and added to a vial, and then .5 grams of powdered sample was measured and blended in with the flux. While most of the Tipu samples did successfully create glass beads, the fluxer machine was nonfunctional during my time in the lab and I personally never had the chance to create any glass beads.

However, this was not a total failure, as it is the trace elemental analysis that tells us more about the pottery. Rice notes, "Trace elemental analyses are based on the principle that most naturally occurring substances contain minute amounts of certain rare or trace elements. The combination and concentration of these elements are unique to their source and will vary considerably from the trace components in similar materials having a different source" (1978:513). To create the wax pellets used in trace elemental analysis, I measured out six grams of powdered sample and blended it with 1.5 grams of binder. Then using the die press machine in the lab, I put the sample under 20 metric tons of pressure for five minutes to create the wax pellet.

The finished samples were then run in the XRF spectrometer. Samples were carefully loaded with the homogenous side facing the laser beam and analyzed using SuperQ Manager

software, a process that took about two hours to complete. The results were then loaded into an Excel spreadsheet for further analysis and interpretation.

Potential Problems With XRF Analysis

Despite the best efforts and intentions, experiments can go wrong. There are several important issues to consider when analyzing elements in ceramics. The first, and most prevalent, issue with obtaining accurate results is inexperience. Shackley (2010) writes that with XRF technology there is the "…issue of poor scientific training by American archaeologists." Archaeologists are not chemists, and many archaeologists who undertake XRF studies do not know how XRF works or how to properly prepare samples. Shackley continues, "…archaeologists are interested in science, but are not necessarily trained in a scientific discipline." This cautionary statement addresses the fact that archaeologists have been guilty of rushing into XRF studies without proper knowledge and training. This whole study has been an exercise in learning, as the techniques were new to both my mentor and me.

This especially comes into play when Portable X-Ray Fluorescence (PXRF) is taken into consideration. PXRF machines have, within the past 10 years, frequently come into the hands of inexperienced archaeologists. Compared with INAA and XRF, PXRF is nondestructive and can be used on ceramics and archaeological material finds without causing any damage. These machines allow archaeologists to simply point a laser at an object and obtain quick results without the hassle of "traditional" sample preparation. However, the lack of experience and knowledge can lead to skewed results.

Although PXRF is theoretically non-destructive, to correctly analyze ceramics the surface must be clear of paint, glaze, dirt, etc. While an acceptable surface may be found on the pottery, better and more accurate results will be obtained when the surface has been abraded or the sample has been prepared as a pressed pellet or fused disk (Speakman et al. 2011:3484).

Another issue that archaeologists must contend with is the accuracy of PXRF machines. In a study by Speakman et al., ceramic pottery samples from the American Southwest was analyzed by a PXRF instrument and through INAA. All samples were carefully prepared (INAA samples were cleaned and powdered and PXRF samples abraded on the surface and then cleaned) and chosen from a collection of well-studied pottery. The researchers found that some elements correlate well between the two analytical methods, while others had lower correlations. Rubidium (Rb), Strontium (Sr), and Iron (Fe) all correlated well. However, Thorium (Th) and

Aluminum (Al) had low correlation ratios. When the major elemental data was separated into groups, the PXRF data did not fall into distinct, chemically separate groups. Groups based upon chemical makeup were much clearer with INAA. The study concluded, "Although there is not a direct 1:1 correlation between the two datasets, data from the portable XRF study are somewhat in agreement with the corresponding INAA data… Despite the generally acceptable results from portable XRF, it is clear that INAA has significantly greater analytical precision. In addition, the ability of INAA to measure trace and rare earth elements has proved critical to being able to effectively identify compositional groups that are useful for provenance studies" (Speakman et al.).

Another problem that crops up with ceramic chemical analysis is the idea that archaeologists make certain assumptions when dealing with ceramics. Rice challenges the basic assumption in ceramic studies is that "…at no time during pottery manufacture or use…is the trace element configuration characteristic of a raw clay deposit significantly altered" (1978:514). In order to assign pottery to a specific manufacture area, archaeologists assume that the geochemistry and availability of clay deposits have not changed in the intervening time, that the potters have not significantly altered the clay, that firing time and temperature will not affect trace element composition, and that post-depositional alteration does not introduce or leach out elements (i.e., the ceramic chemical makeup does not change once it has been discarded due to environment) (Rice 1978:515–517). While archaeologists can safely work with some of these assumptions, special consideration must be taken into account for others and extra work must be carried out in order to counterbalance such assumptions.

Conclusion

While all samples have been powdered, made into pellets, and run through the XRF spectrometer, we do not yet have conclusive results to share. As my mentor and I both are new to this process, we have decided to consult an XRF specialist to help us make sense of our results. Aaron Shugar, an inorganic chemistry conservation scientist at Buffalo State University, use-tested our created samples using a Bruker handheld TRACER III-SD XRF spectrometer. Shugar ran the prepared samples with the PXRF machine in order to compare those results with our results and to help us understand the data. Shugar, geologist Dori Farthing, and my mentor Jim Aimers have used these preliminary results to create a calibration system for PXRF in order to

contend with some of the PXRF problems mentioned above. While we do have raw data, we have not yet had the opportunity to analyze and interpret this data in relation to our goals.

In addition, as this is a new direction of study, we did not anticipate how long it would take to create and test samples and to analyze those results. Subsequently, there is a need to continue testing more samples and to assist in the analysis of the results. My mentor and I are making plans to continue working on this study and to obtain and analyze further samples. With these results we hope to understand the chemical makeup of the pottery and to determine if chemical classification corresponds with our stylistic classifications. Ultimately, through our work with classification we hope to better understand the Maya trade and exchange routes of the Postclassic period.

Works Cited

Rice, Prudence.
 1978 Clear Answers to Vague Questions: Some Assumptions of Provenience Studies of Pottery. In *The Ceramics of Kaminaljuyu, Guatemala*. R. K. Wetherington, ed. pp. 513–542. The Pennsylvania State University Press, Philadelphia.

Rice, Prudence.
 1987 *Pottery Analysis: A Sourcebook*. University of Chicago Press, Chicago.

Rice, Prudence.
 2012 Type-Variety: What Works and What Doesn't. In *Ancient Maya Pottery: Classification, Analysis, and Interpretation*. J. J. Aimers, ed., in press. University Press of Florida, Gainesville.

Sanders, W. T.
 1960 Prehistoric Ceramics and Settlement Patterns in Quintana Roo, Mexico. *Contributions to American Anthropology and History* 12, Carnegie Institution of Washington Publication 606, Washington, D.C.

Shackley, Steven M.
 2010 Is There Reliability and Validity in Portable X-Ray Fluorescence Spectrometry (PXRF)? *The SAA Archaeological Record* 10(5):17–20.

Shugar, Aaron.
 2009 Peaking Your Interest: An Introductory Explanation of How to Interpret XRF Data. *WAAC Newsletter* 31(3):8–10.

Smith, Robert E., Willey, Gordon R., and Gifford, James C.
 1960 The Type-Variety Concept as a Basis for the Analysis of Maya Pottery. *American Antiquity* 25(3):330–340.

Speakman, Robert J. et al.
        2011 Sourcing Ceramics With Portable XRF Spectrometers? A Comparison With INAA Using Mimbres Pottery From the American Southwest. *Journal of Archaeological Science* 38:3483–3496.