SURJ

Stanford Undergraduate
Research Journal

Volume 11, Spring 2012

Stanford, CA 94305

Journal Staff

Executive Board

Editors-in-Chief
Vir Choksi
William Bassett

Editors Emeritus
Jennifer Levy
Daniel Shaffer
Karis Tang-Quan
Shinjini Kundu
Justin Solomon
Andy Ho
Peter Pawlowski
Emily Gustafson
Victor McFarland
Kevin Gao
Vineeta Agarwala
Jeff Park
Julian Hong
YunXiang Chu

Production Board

Susan Tu
Joel Dominguez
Bryce Cronkite-Ratcliff
Sheila Ramaswamy

Editorial Board

Natural Sciences
Section Editor
Daniel Ibarra
Associate Editors
Chen Lossos
Eric Liaw
Shannon Wu
Devan Diwanji

Humanities
Section Editor
Melissa Hesselgrave
Associate Editors
Angelina Garron
James Nie
Vivek Choksi

Engineering
Section Editor
Bryce Cronkite-Ratcliff
Associate Editors
Arzav Jain
Christopher Guthrie
Sunil Deolalikar

Social Sciences
Section Editor
Jodie Ha
Associate Editors
Laurie Rumker
Dmitriy Brezhnev
Evelyn Chang
Isaac Madan

The mission of the Stanford Undergraduate Research Journal is to encourage, recognize, and reward intellectual activity beyond the classroom, while providing a forum for the exchange of research and ideas. SURJ encourages students to become interested in research by displaying examples of what is studied and by offering the means of communicating knowledge between these disciplines to achieve an holistic effect.

A Letter from the Editors

Dear Reader,

It is with great pride and anticipation that we present the eleventh issue of the Stanford Undergraduate Research Journal. SURJ continues its commitment to presenting the best of undergraduate research both on campus and worldwide. With over 100 submissions received for this issue, it is clear that our undergraduates have continued their commitment to extend human understanding at every frontier.

From this rich corpus, SURJ's editors chose to publish 18 manuscripts that together exemplify the extreme breadth of research in which undergraduates are involved, and the depth to which their curiosity drives them. While staying committed to our mission to showcase the exciting and substantial research conducted by undergraduates at Stanford, nearly half of this year's articles come from students at other institutions around the world. Together, this year's authors truly embody intellectual vitality, that elusive but essential quality possessed by those who wake each morning seeking to understand more about humanity, nature, and the universe than they did the previous night. Our authors are held to the highest standard in this regard, and this issue of the journal is an undeniable testament to their dedication.

These contributions to the fields of Humanities, Social Science, Natural Science, and Engineering are brought to you thanks also to the tireless efforts of our team of editors and designers, Stanford undergraduates all. In particular, we would like to recognize section editors Daniel Ibarra ('12), Bryce Cronkite-Ratcliff ('13), Jodie Ha ('14), and Melissa Hesselgrave ('12); and head of design, Susan Tu ('15), whose leadership allowed production to run seamlessly from start to finish. We would also like to thank Professor Jonathan Hunt from the Hume Writing Center, whose guidance was greatly beneficial in the selection and refinement of articles. Lastly, we owe enormous gratitude to Dean Eyre III ('80) and the entire office of Undergraduate Advising and Research for their support over this year, and for the entirety of SURJ's run thus far. Of course, SURJ could not flourish without the full support of the Stanford undergraduate research community, and we hope that this edition of SURJ encourages its readers—be they undergraduates, graduates, high school students, or alumni—to think critically about the world around them, ask probing questions, and persevere in deepening human understanding.

Please join us in congratulating our authors and staff. This issue of the journal—their work product—serves as a reminder that young minds can be a source of knowledge and inspiration for their fellow students and researchers the world over.

Sincerely,

Vir Choksi
William Bassett
Editors-in-Chief

Want to learn more about Stanford Undergraduate Research Journal?

surj.stanford.edu

-Additional figures from this year's papers
-Archives of every previous journal
-Announcements for ongoing production
-Submit a manuscript, or get involved in editing and producing the next edition!

SURJ

A Letter from the Editors 3

NATURAL SCIENCES

Raman spectroscopy investigation in the NWA 3118 meteorite: Implications for planet formation 6
B. Ghimire, A. Dall'Asén, J. Gerton, I. Ivans, B. Bromley

Effects of CO_2 driven ocean acidification on sea hare (*Aplysia dactylomela*) larval development 9
B. Yanovsky

Virtual screening identifies novel inhibitor scaffolds of human carbonic anhydrase VII 15
D. Ljuboja

Physical properties of habitats occupied by platypus (*Ornithorhynchus anatinus*) in undisturbed and recently disturbed sites along rifle creek, Far North Queensland 21
W. LaGrandeur

An efficient $FeSO_4$-mediated synthesis of methyl-4-(ethoxymethyl)-benzoate and basic conformational analysis of the same using computational tools 27
S. Bhakat

HUMANITIES

Redemptive existentialism and Berkeleian metaphysics: a synthesis in Beckett's plays 31
H. Kim

Nature or nurture? Perspectives of Japanese manga on an age-old debate 35
S. Wang

Fichte the fascist? The misappropriation of a republican philosopher in Weimar, Germany 1918-1933 38
M. Albada

Terrible lizards and the wrath of God: How 19th century Christianity and Romanticism affected visual representations of dinosaurs and our perceptions of the ancient world 43
A. Peterson

ENGINEERING

The Bulb Vacuum System (BVS) – Negative Pressure Therapy for Healing Acute and Chronic Wounds in Underdeveloped Areas 50
D. Ljuboja

A thermal model for analysis and control of drilling in icy formations on Mars 56
A. Aggarwal

Coarse head pose estimation using image abstraction 61
A.V. Puri

Politics, Twitter, and information discovery: Using content and link structures to cluster users based on issue framing 67
M. Sudhof

SOCIAL SCIENCES

Masdar City: A model of urban environmental sustainability 77
A. Lau

The sordid economy: Gender and disease in nineteenth century labor struggles 83
Z. Leavitt

The significance of mixed-race: Public perceptions of Barack Obama's race and its effect on his favorability 87
S. Sinyangwe

United States drug policy: The scientific, economic, and social issues surrounding marijuana 95
M. Winterbourne

Land and labor in the Negev: Palestinian Bedouin citizens of Israel 101
R. Zuabi

Raman spectroscopy investigation in the NWA 3118 meteorite: Implications for planet formation

Bhuwan Ghimire, A. Dall'Asén, J. Gerton, I. Ivans, B. Bromley

A hierarchical model for the early stages of planet formation involves the coagulation of micron-scale dust particles orbiting the sun into larger, centimeter- or meter-scale objects. Computational simulations can produce planetary distributions that agree with astronomical data, but only when they start with planetesimals of ~1 km or more in size. Further investigation is needed to understand how these planetesimals arose from the adhesion of dust aggregates. One possibility is that the small-scale structure of interstellar dust enhanced adhesion. At very least, this structure, as preserved in stony meteorites commonly known as chondrites, may provide clues to the physical processes during this critical juncture in planet formation. Confocal Raman microscopy was used to investigate microstructures on the NWA 3118 meteorite. It is a carbonaceous type CV3 meteorite composed of micrometer scaled chondrules and inclusions embedded in the matrix material. Olivines and graphitic structures were identified in the NWA 3118. Raman mapping of meteorite samples was carried out, with a focus on the interfaces between chondrules and the matrix.

The formation of terrestrial planets and the icy cores of gas giants is driven by coagulation and accretion. The dust seen in the protoplanetary disks around many young stars is concentrated into larger planetesimals, which in turn merge to form full-fledged planets. While this scenario is compelling, it is incomplete in one regard: The growth of kilometer-sized planetesimals from micrometer sized protoplanetary dust particles remains an unsolved problem. Several theories have been proposed, including simple electrostatic adhesion of dust aggregates and the gravitational collapse of the material concentrated in turbulent eddies in the sun's protoplanetary disk.

To find clues for discriminating between these models, researchers have used Raman spectroscopy of chondrules – small grains formed in the early solar system found in several meteorites like Allende, Axtell & Murchison, NWA 3118, Nakhla, and Vaca Muerta.[1,2,3] Raman scattering produces an optical spectrum, which when compared to a reference spectrum, can yield information on specific minerals present in the chondrules and the matrix, the material in meteorites that binds constituents together. To focus on the dust-to-planetesimal problem, this study concentrates on processes by which meteorites were assembled. Specifically, we examine the interface between the chondrules and matrix of a type CV3 meteorite is examined in this review. The interface between the chondrules and the matrix may hold the answer to the question of whether kilometer-sized planetesimals were formed by collisional growth or gravitational instabilities. This study focuses on a sample from one meteorite, NWA 3118, and presents Raman spectra with the corresponding images of identified olivines and graphitic structures within this sample. The findings complement and corroborate the recent findings on this meteorite, and represents first efforts in a more detailed analysis.

Chrondrites

Chondrites are stony meteorites primarily composed of chondrules, rare refractory inclusions and the matrix material within which the chondrules and the inclusions are embedded. The millimeter-sized mineral aggregates of chondrites are the "chondrules." These are made from several kinds of minerals, mainly olivine and pyroxenes, and are the deposits of dust aggregates from the early solar system.[4] Refractory inclusions are composed of either minerals like calcium and aluminium or amoeboid aggregates of the matrix minerals, such as olivine with proportions of metals like iron, nickel, or both.

The matrix of the chondrites is full of fine-grained silicate minerals like forsterite, anorthite, and pyroxene. These minerals play a vital role in the formation of chondrites. They aid to fill the gap between the chondrules coating them, acting as an interface between them and the inclusion.[4] Understanding of these components with their formation will be simpler through the thermodynamic equilibrium diagram for solar nebula. Scott (2007) discusses mainly two types of minerals at a pressure of 10^{-3} bar: minerals that are stable below 1400 K, which are mainly present in chondrules and the matrix material, and the ones that are stable above 1400 K that are found in the refractory inclusions. The thermal processing of this silicate dust in the solar nebula into the refractory inclusions, chondrules in chondrites, and sub-micrometer silicate crystals present in chondrite matrices and comets needs to be understood. These materials aid in the accretion of asteroids, comets, and planets.[4]

The thermal and chemical history of meteoritic material reflects the conditions and processes that gave rise to the solar system. The formation of planetesimals, the kilometer-size building blocks of planets, begins with growth of fractal dust aggregates, followed by compaction processes; with the increase in the size of the dust aggregates, the mean collision velocity increases, which in turn stops the growth as the dust fragments reach decimeters in size.[5] The question is: what succeeds this stage? Several hypotheses for further growth have been proposed, including but not limited to sticky materials, enhanced growth at snow line (the region outside of Mars' orbit where refractory material can condense into ices), and cumulative dust effects with gravitational instability.[4,6,7] These theoretical ideas may be realistic, but they need to be tested in some way. The approach here is to seek evidence of these theoretical scenarios in the small-scale structure of meteorites.

Experimental

Material

The NWA 3118 was obtained from Erfoud, Morocco. It was first found in 2003. It is a type CV3 meteorite. CV3 is named after the type Carbonaceous '*Vigarano'* meteorite with petrologic type 3. Here, the type refers to the degree of the effects of thermal metamorphism and aqueous alterations,

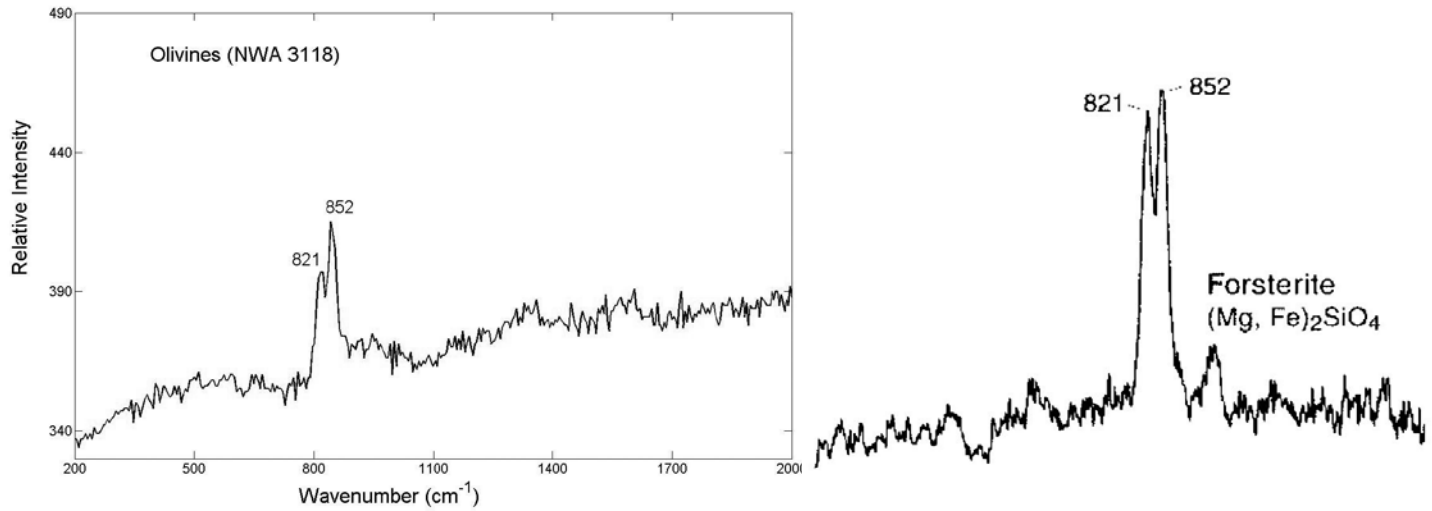


Fig. 1 (a) Raman spectra for olivine (forsterite) with the characteristic Fermi doublet at 821 and 852 cm^{-1} observed in the NWA 3118. **(b)** Raman spectra of forsterite on Martian rocks (Sharma 2002).

with 3 generally given to the most pristine chondrite with minimal aqueous alteration and low metamorphism.

Instrumental

This study used a laser excitation wavelength at 488 nm to investigate carbon structures and different meteoritic compounds present on the NWA 3118 meteorite. The analysis system used was a WITec alpha 300 S scanning near-field optical microscope. The detection gear included a WITec UHTS 300 Spectroscopy system equipped with a thermoelectric-cooled, back-illuminated CCD detector chip. The grating used in the spectrometer was 600 g/mm and the spectral resolution was around 3 cm^{-1}. Lenses 10x and 20x were used as objectives. Data acquisition was evaluated through the in-built WITec Project software. Spectral analysis was performed through Matlab 7.11.

Results and Discussion

The olivine micro-structures were verified on the NWA 3118 chondrite with the findings of Fries & Steele (2008). A strong signature of olivine detected from Raman bands at 821 and 852 cm^{-1} is seen in the spectrum of the NWA 3118 sample (Fig. 1a). Olivines of forsterite $(Mg,Fe)_2SiO_4$ have been identified to exhibit this doublet as a fingerprint in Martian surface minerals.[8] Fig. 1c represents the optical image of the surface with fragments of olivine chondrule under 20x objective lens, and Fig. 1d represents the Raman map of an olivine chondrule fragment.

Several line scans around the chondrule rims were taken. A plot for a line scan with spatial separation of 10 µm around the chondrule rim generated through Matlab is shown in Fig. 1e. The Olivine fingerprints show strong peaks at 821 and 852 cm^{-1} with a lot of background. The background was from space weathered materials and highly reflecting particles.

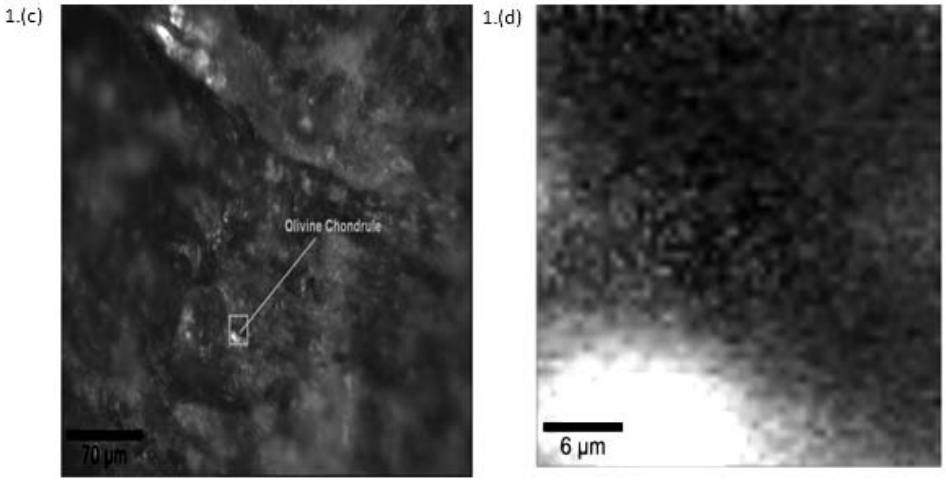


Fig. 1 (c) Photomicrograph of an olivine chondrule fragment taken with a 20x objective lens. **(d)** A sum filtered Raman image of the marked region in Fig.1 (c) with the olivine fragment at the bottom left .

Graphite

Graphite-like phases, lonsdaleite-hexagonal diamond, and graphite whiskers have been predicted and recorded on the NWA 3118, marking it as a carbonaceous chondrite.[2,9] Fig. 2a gives a clear spectrum of the presence of bulk carbon in the form of imperfect graphite near an edge of the NWA 3118 chondrite. The most intense features were identified as the D peak at ~1350 cm^{-1}, the G peak at ~1580 cm^{-1} and a weak G' peak at ~2682 cm^{-1}. The G' peak should be the second most intense peak in graphitic samples. It is the second order of zone-boundary phonons; but since Raman fundamental selection rules are not satisfied by zone-boundary phonons, they are usually not seen in first order Raman spectra of defect-free graphite. However, these phonons give a rise to a peak at ~1580 cm^{-1}, known as the D peak in defected graphite.[10]

Graphene was not identified in the structures, as a D peak should not be observed on pure graphene layers. A significant change in shape and intensity should be found in the G and G' peaks to indicate any graphene presence compared to the spectrum of defected graphite.[11] This study corroborates the findings of Karczemska *et al*. (2006) and Fries & Steele (2008). The photomicrographs of

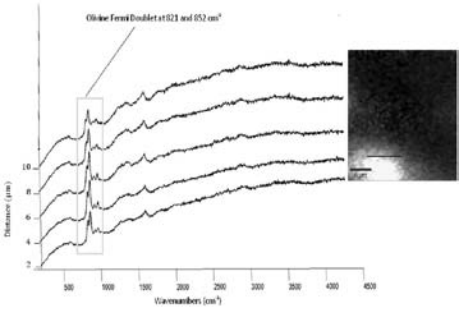


Fig. 1 (e) A plot of the line scan around the chondrule rims represented by the black line in the Raman image. The spatial separation is 10 μm.

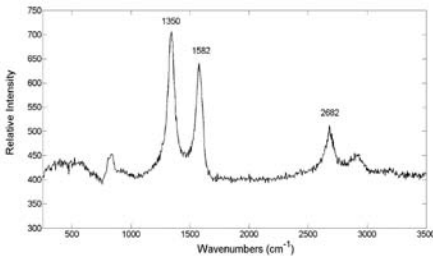


Fig. 2 (a) Raman Spectrum of defected graphite with peaks at 1350, 1582 and 2682 cm^{-1} observed in the NWA 3118. A doublet indicating the presence of olivine in around 816-855 cm^{-1} wavenumbers.

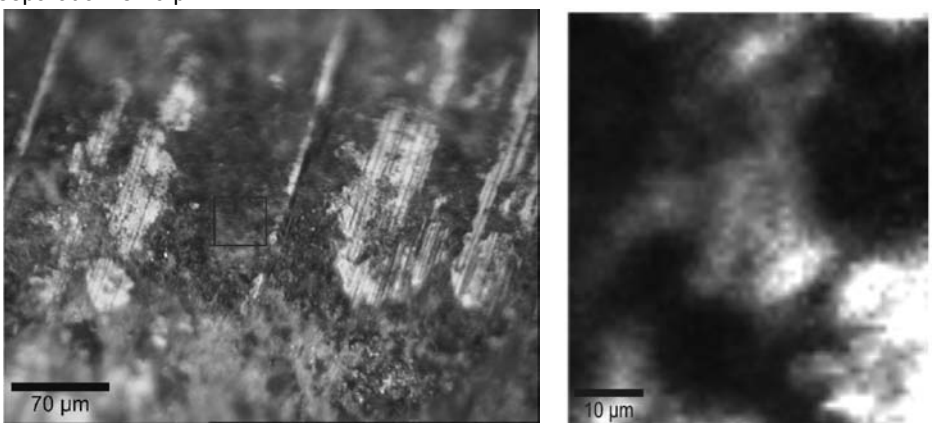


Fig. 2 (b) Photomicrograph of bulk carbon in defected graphite form. **(c)** Raman image of the spot with brighter regions representing carbon structures.

the spots of bulk carbon and Raman image of various defected graphite structures are shown in Figure 2b.

Conclusions

This study presents a comparative Raman spectroscopic study of a carbonaceous meteorite NWA 3118 in this review. The findings corroborate the recent discoveries: graphite-like phases, lonsdaleite-hexagonal diamond and graphite whiskers on the NWA 3118.[2,9] We also confirm that our technique and equipment is capable of making valuable analytical measurements on the scales necessary to glean information about the interfaces between constituents of meteorites.

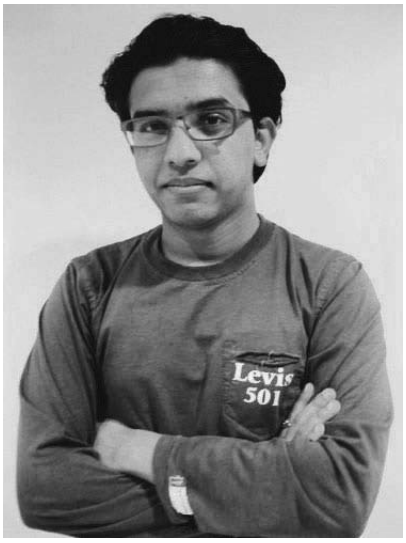
Olivines were identified in the form of forsterites $(Mg, Fe)_2SiO_4$ on the NWA 3118 in the form of micrometer chondrules embedded in the matrix. Defective graphites were identified near the edges of the sample specimen. These microstructures may answer some of the questions raised on planet formation. It is critical to understand the processes with which planetesimals are formed. Several theories have been proposed including sticky properties of the grains, dust aggregate collisions and gravitational instability. The coating of the chondrules and the matrix with the amoeboid micro and nanostructures discussed by Scott (2007) explain the possible coagulation of pre-solar grains and dust. However, it is not yet clear if or how this process leads to growth of kilometer-sized structures like planetesimals. The interface between the chondrules and the matrix may yield the answer to the question of whether kilometer-sized planetesimals were formed by collisional growth or gravitational instabilities. Until it does, the dust-to-planetesimal problem will remain "one of the great unsolved problems of planet formation."[4]

Acknowledgement

We would like to thank Randy Polson from The University of Utah for his valuable technical support with the facilities.

References

1. Larsen KL, Nielsen OF. Micro-Raman spectroscopic investigations of graphite in the carbonaceous meteorites Allende, Axtell and Murchison. J.Raman Spectrosc. 37:217-222.
2. Fries M, Steele A. Graphite Whiskers in CV3 Meteorites. Science, 2005; 320:91-93.
3. Rull F, Martinez-Frias J, Sansano A et al. Comparative micro-Raman study of the Nakhla and Vaca Muerta meteorites. J.Raman Spectrosc, 2004; 35: 497-503.
4. Scott E. Chondrites and the Protoplanetary disks. Annu.Rev.Earth.Planet.Sci. 35: 577-620.
5. Blum J, Wurm G. The Growth Mechanisms of Macroscopic Bodies in Protoplanetary Disks. Annu. Rev. Astro. Astrophys, 2008; 46: 21-56.
6. Chiang E, Youdin A. Forming Planetesimals in Solar and Extrasolar Nebulae. Annu. Rev.Earth. Planet.Sci, 2010; 38: 493-522.
7. Youdin AN, Shu FH. Planetesimal Formation by Gravitational Instability. The AstroPhysical Journal, 2002; 580: 494-505.
8. Sharma S, Lucey P, Manash G, et al. Raman spectroscopic detection of minerals on planetary surfaces. Spectrochimica Acta Part A, 2002; 59: 2391-2407.
9. Karczemska A, Kozanecki, M, Szurgot M, et al. Raman spectroscopy studies of meteoritic diamonds. Diamond $ Related Materials, 2007; 16: 781-783.
10. Remanich RJ, Solin SA. First- and second-order Raman scattering from finite- size crystals of graphite. Physical Review B, 1979; 20: 392-401.
11. Ferrari AC, Meyer JC, Scardaci V et al. Raman Spectrum of Graphene and Graphene Layers. Physical Review Letters, 2006; 97: 1-4.



Bhuwan Ghimire was born in Kathmandu, Nepal, and now studies physics at Westminster College in Missouri. He is particularly interested in observational cosmology and instrumentation, and after graduation plans to research galactic archeology and cosmology. In his spare time he likes to stay active, and particularly enjoys playing soccer and tennis.

Effects of CO_2 driven ocean acidification on sea hare (*Aplysia dactylomela*) larval development

Beckie Yanovsky

Increasing the partial pressure of CO_2 in the atmosphere affects the calcification rates of marine organisms by lowering aragonite concentrations in oceanic water. This experiment considered the development of the gastropod mollusk *Aplysia dactylomela* in its planktonic larval stage during which the organism calcifies an exterior shell, a crucial step in development. The effects of acidification on sea hare larvae development were analyzed using an ocean acidification system to submerge *A. dactylomela* egg masses in three different pH treatments. The variables considered include differences in shell length, shell smoothness, inner morphology, and larval movement. The experiment revealed that larvae reared in acidified conditions had significantly shorter shells and a higher proportion of unhealthy inner morphologies than the ambient control. Larvae also had a higher observed proportion of pitted shell morphologies and an overall decrease in mobility, though differences were not significant. Physiological changes could potentially cause functional and behavioral discrepancies and have an effect on the overall fitness of the organism. Disrupting the calcification process of sea hares and other mollusks through environmental acidification could have widespread consequences for coral reef ecosystems.

Rising CO_2 concentrations in the Earth's atmosphere are being taken up by one of the largest natural carbon sinks – the ocean. The fluctuation of the biogeochemistry of the ocean is a natural process that has occurred for millions of years; however, the rate at which carbon dioxide emissions are affecting the partial pressure of CO_2 (P_{CO2}) in the atmosphere is unprecedented in Earth's history. Atmospheric P_{CO2} has risen 100 parts per million (ppm) since preindustrial times, an increase of 36% from partial pressure values during the 18th.[1] And over one-third of the anthropogenic emissions have been absorbed into the ocean.[2]

Excessive carbon uptake by the ocean results in a phenomenon known as Ocean Acidification (OA) – the drop in oceanic pH as CO_2 is absorbed into the water. Studies using projected carbon dioxide levels from the Intergovernmental Panel on Climate Change (IPCC) predict a pH drop of .3-.4 pH units by 2100.[3] OA reduces the calcium carbonate concentration in the ocean, negatively affecting calcifying organisms by limiting their ability to grow, and develop.[4]

Calcifying organisms depend on calcium carbonate ($CaCO_3$) to build their skeletons and shells. Lowering the pH of surface waters can compromise carbonate accretion or cause $CaCO_3$ skeletons to dissolve into the carbonate-unsaturated waters. Previous studies have demonstrated negative effects from ocean acidification on reef-building corals – the structural backbone of reef communities – as well as other organisms such as pteropods, coralline algae, and mollusks.[5] Many marine invertebrates undergo a planktonic larval stage in which they develop a calcified exoskeleton. Planktonic larvae play important roles in the benthic–pelagic life cycle of marine organisms by connecting and sustaining existing populations and contributing to biodiversity. The calcifying larval stages of marine invertebrates are likely to be more vulnerable to OA, in part because they form their internal skeletons out of amorphous calcite which is more soluble than other forms of carbonate.[6] There is relatively little information on the long-term effects of pH changes on marine invertebrates, in particular their early developmental stages. The embryonic and larval stages are believed to be more sensitive to environmental disturbances and play critical roles in selection, recruitment and population maintenance.[7]

The spotted seahare, *Aplysia dactylomela*, is one such marine invertebrate that calcifies a spiral $CaCO_3$ shell in its planktonic larval stage. As an integral part of the marine symbioses within the coral reef community, the spotted sea hare plays a vital role in nutrient cycling, and also has implications for human medical sciences. Perhaps the sea hare's most significant role in the reef community is that of an avid herbivore. Space competition between algae and corals is an important ecological process underlying coral reef dynamics. A decrease in herbivore efficacy can promote algal competitiveness and lead to local reef degradation. Ocean acidification has been shown to make reef communities more vulnerable to algal takeover. Processes that regulate algal abundance, such as herbivory, will play an increasingly important role in maintaining coral communities.[8]

The calcifying early-stage life development of the sea hare impels the question if exposing egg masses to increased pH levels during embryonic development would have a significant effect on larval morphology and functionality, either by reducing calcification or enhancing dissolution in the $CaCO_3$ shells. Results could shed light on organisms' potential to reach juvenile and adult stages and assume their role as reef herbivores.

This question was explored in *A. dactylomela* by looking at shell morphology, internal development, and mobility of sea hare larvae whose egg masses developed in varying pH treatments. A relatively exposed and prevalent herbivore in the Great Barrier Reef, the spotted sea hare is a relevant organism in maintaining coral-algal balance. Furthermore, the spotted sea hare is a good study system for analogous calcification processes of other marine invertebrates. Analyzing marine communities' ecological responses to environmental changes requires an understanding of the significance of larval-stage development modifications, since future environmental stressors are likely to be expressed as sublethal effects before direct mortality.[4] The developmental implications of ocean acidification on herbivorous marine invertebrates could play a key role in the future of coral-algal interactions in marine ecosystems.

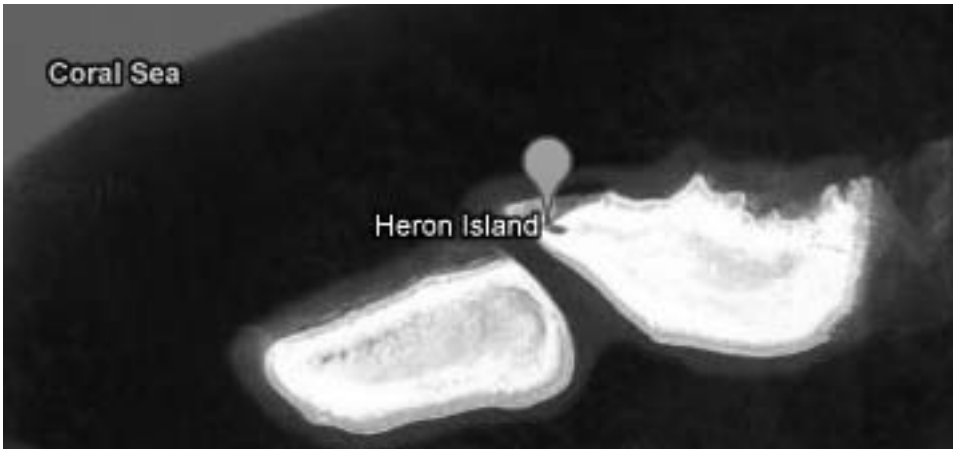


Fig. 1 Heron Island is a forested sand cay 72 km east of the Australian Coast. The island is located on the western side of Heron Reef, a platform reef of significant biodiversity, supporting 72% of the coral species found in the Great Barrier Reef. The reef was the collection site for the *A. dactylomela* specimen. cant (p=.13).

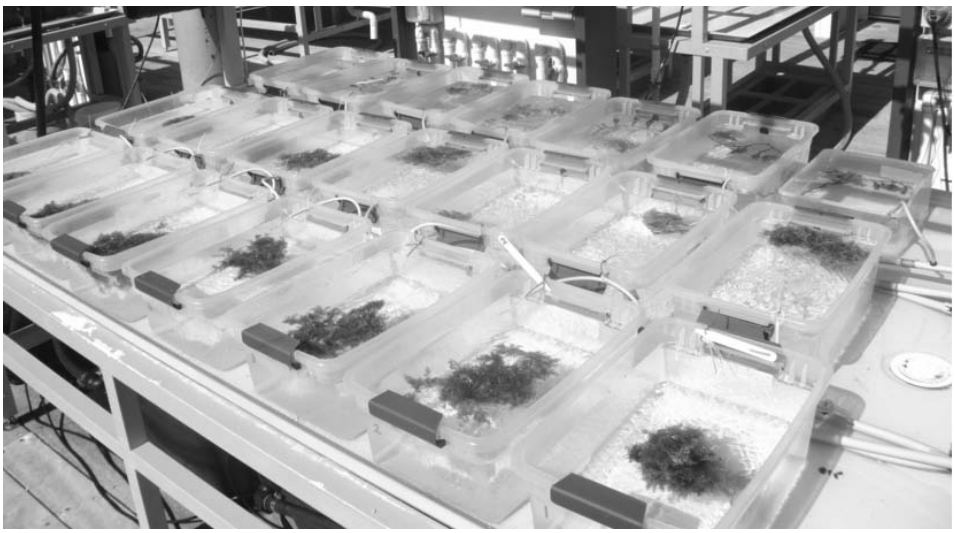


Fig. 2 This ocean acidification system on Heron Island Research Station was used to analyze the effects of rising oceanic pH on sea hare larvae. Six tubs housed 12 containers (2 each) of *A. dactylomela* egg masses in varying pHs from embryo stage to hatching – a period of 8 days.

Table 1 CO_2 Predictions based on the IPCC 2007 Report and Correlated pHs	
Current levels of CO_{2atm} (control)	380 ppm (pH = 8.2) (Ambient CO_2)
IPCC Scenario B1 of CO_{2atm} at year 2100	550 ppm (pH = 7.9) (Adding CO_2)
IPCC Scenario A2 of CO_{2atm} at year 2100	800 ppm (pH = 7.5) (Adding CO_2)

Methods

Study Site

Heron Island is a densely forested sand cay predominantly covered by Pisonia grandis and marginal grass and shrub vegetation. It is located on the edge of a coral reef platform in the Southern Great Barrier Reef, 72 kilometers northeast of Gladstone, Queensland, Australia (23° 26′ 31.2″ S, 151° 54′ 50.4″ E) (Figure 1).

Specimen Information and Larval Cultures

Found in shallow waters, tide pools, and sandy substrates, *A. dactylomela* belongs to the family Aplysiidae and class Gastropoda.[9] Due to its high fecundity and distinct calcifying developmental stage, the sea hare is a good model organism for this OA study. The egg-laying specimen was collected on the reef flat of Heron Island and the experiment was conducted in the laboratories of Heron Island Research Station in October of 2011.

The *A. dactylomela* egg masses were laid in a controlled, ambient (pH 8.1) touch tank at Heron Island Research Station. The sea hare subsisted on a diet of algae also collected from the reef flat, predominantly Laurencia. The egg masses were probably laid in two distinct episodes, as half of the masses were a bright green color and the other a duller brown. There is currently no definite information relating egg mass color to development differences in larval sea hares. Egg masses are known to vary in color depending on the sea hare's diet.[10]

The egg masses were divided and transferred into individual containers after being laid. Two sets of egg masses, one from each birthing period, went into each of 6 tubs with varying pHs. The plastic containers maintained fluvial input through two openings in the casings covered by 120-µm plankton mesh. Larvae were not fed throughout their tenure in the ocean acidification system to avoid algal photosynthetic effects on tub pH and because of the short time window of the experiment.

A single spotted sea hare laid the *A. dactylomela* egg masses to limit attribution of variable morphologies and survival chances to a mixed gene pool. The spotted sea hare has the ability to store sperm from multiple sexual encounters, so it is possible that egg masses laid at different times came from unique sea hare sperm, though this experiment has reduced the likelihood of this by sampling egg masses from the same individual at close time intervals; thus the gene pool is more standardized, and the variable is limited to ocean acidification.

Ocean Acidification System

Embryos of the spotted sea hare *A. dactylomela* were submerged in treatments of three different pHs – 8.1, 7.8, and 7.6 – representing an ambient and two projected scenarios of oceanic pH. The pH values were chosen to represent the stabilisation levels predicted by the Intergovernmental Panel on Climate Change (IPCC 2007). As pH is reduced in a predictable manner by elevated CO_2 concentrations, the CO_2 levels of the experimental, unfiltered seawater were controlled by adjusting the pH in separate, 200 L sumps. The pH of the seawater was maintained by a control unit (Aquatronica, AEB technologies, Italy) which opened solenoid valves that injected CO_2 into the seawater when the pH went

above the desired threshold. Probes were cleaned and calibrated using Mettler-Toledo calibration buffers to 0.01 pH units. Data were logged every 3 hours to provide average pH values for the duration of the project.

The following two scenarios are IPCC (Intergovernmental Panel on Climate Change) predictions for the year 2100. Treatment levels of pH 7.6 and 7.8 are likely to be conservative, as new data suggests ocean acidification is tracking along the worst case scenario (A2), and likely exceeding this.

The experiment used 6 of 18 available tubs in the ocean acidification system on Heron Island Research Station (Figure 2), with two tubs at each of the above pHs. Two containers of sea hare egg masses were placed in each tub. The embryos were transferred within 24 hours of being laid to maximize development time in their respective treatments. The egg masses were left in treatment for 7-8 days before data were recorded the evening of their hatching.

Morphological Observations

Sample photographs were taken from each treatment under 400x magnification to make detailed morphological observations on the larval specimens. Larvae were dispensed from their treatment onto a flat glass slide with a cover slip. An average of 15 individuals were photographed for analysis from each container. From these photos, the clearest 58 individuals were analyzed from each pH treatment for shell morphology, shell length, and internal composition.

Shell Morphology

A scale was developed to standardize assessments across the dependent variables (Table 2). Shell morphology was grouped into smooth and pitted; pitted shells had fragmentation and thinning in their $CaCO_3$ skeleton, while smooth shells did not.

The proportion of smooth to pitted shells was recorded for each container and then averaged across similar pH treatments to compare shell development.

Shell Length

A scale was developed to assess the proportions of shells that developed to full length in each acidification treatment (Table 3). Full-length shells extended at least to the tangent of the inner curve. Shells classified as underdeveloped were significantly shorter than their counterparts. Full length and shortened shells were tallied and then converted to percentages of the total individuals analyzed per container. Averages for containers in the same pH treatments were calculated to analyze shell length trends.

Inner Morphology

A scale was also developed to assess the inner development of the specimens to determine if increasing P_{CO2} had significant effects (Table 4). Classified individuals were grouped into one of three categories – healthy, unhealthy, and empty. Healthy individuals had well defined organelles and clear separation from their calcium carbonate shells. Unhealthy individuals were more muddled, often with a less clear separation between their inner morphology and calcium carbonate shell. Organelles either had not developed completely or were indistinguishable from each other, and there was a clear deviation from their healthy counterparts. Empty shell casings were simultaneously classified as dead. Empty shells consisted of an incomplete internal development that was a drastic deviation from the busier internal structure of the healthy sea hare larvae.

Percentages of each classification were then calculated for each container and averaged across similar pHs to get a mean proportion of healthy, unhealthy, and empty shells from each treatment.

Movement Observations

Larvae were placed under a 40x magnification the evening of their hatching, and videos were taken to analyze the functional impact of elevated P_{CO2} exposure through the microscopic lens. Approximately 140 larvae were in each frame of the video footage and analyzed from each of the 4 containers per treatment, totaling roughly 4*140=560 larvae observed per treatment. These larvae were pipetted onto a glass slide and studied for at least 2 minutes of swimming time.

Larvae were not individually quantitatively assessed; instead, they were studied as a group and qualified based on two factors – the proportion of larvae swimming and their swimming speed. The standardized scale used to measure across treatments is listed below (Table 5).

Intermediate values (of 1.5 and 2.5) were given based on the swimming speed of the higher value with a lower proportion of individuals moving. Each of the four containers per treatment were judged on this scale and then averaged to compare the functional impact of elevated CO_2 concentration on their mobility. Only movement of the living larvae was taken into account.

Data Analysis

Data analysis was done by sorting the photos and videos from each container based on pH treatment for analysis. Both qualitative and quantitative methods were then employed to obtain statistical data.

Photos from each container were qualitatively sorted based on the scales

Table 2 Shell Morphology Scale

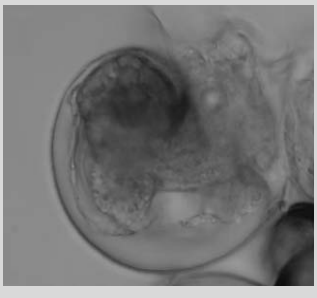
Smooth	Pitted
	

Table 3 Shell Length Scale

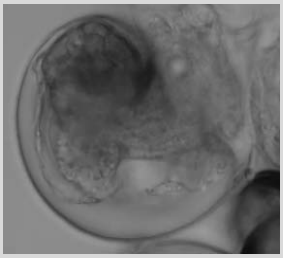
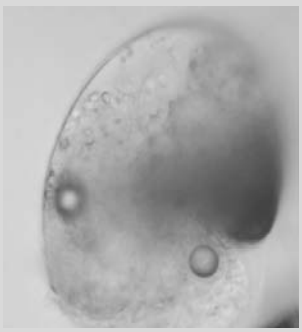
Full Length	Shortened
	

Table 4 Inner Morphology Scale

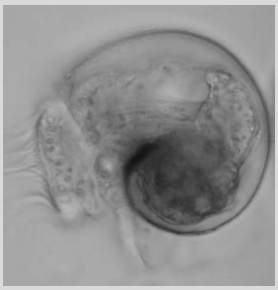
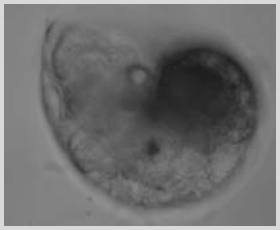
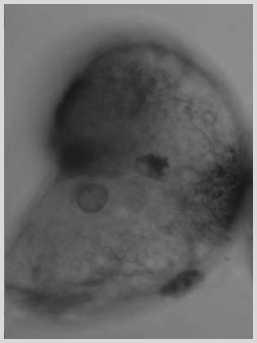
Healthy	Unhealthy
	
Empty	
	

Table 5 Standardized Motion Scale across pH Treatments

1	<10 % swimming at slow to nonexistent speeds.
2	Larvae swimming at intermediate speeds. >20% of larvae in motion.
3	Larvae swimming at high speeds with consistent movement. >30% of larvae in motion.

established in the methodology as Pitted or Smooth; Full Length or Shortened; and Healthy, Unhealthy, or Empty. The numbers were subsequently tallied for each container. These numbers were converted into a percentage based on the individuals in each category divided by the total individuals analyzed in the container. For example, the percentage of pitted individuals would be calculated with ($\%_{pitted}=N_{pitted}/N_{total}$). The same was done for smooth shells ($\%_{smooth}=N_{smooth}/N_{total}$), length, and internal development. These percentages were averaged across the 4 containers in each pH treatment to obtain a single average percentage for each variable. The results were studied and graphed using the standard deviation of the averages to obtain error bars. A single variable analysis of variance (ANOVA) was conducted to judge the significance of the data.

Results

General Observations

All of the containers had at least portions of their egg masses hatch. By the evening of the first day of hatching, all of the containers had larvae swimming around. None of the treatments produced an entirely immobile or mortal population. Changes were only visible when examined under a microscope. Almost all of the individual larvae in the control pH treatment were healthy – classified as smooth shelled, healthy inner morphology, and full length. There were no gross morphological dysfunctions at first glance, whereas in the acidified treatments, many of the larvae had dark masses in their shells, had blurred inner morphologies, and were swimming with less gusto. There seemed to be a larger difference between the 8.1 and 7.8 treatment (perhaps due to the dominant health of the larvae in the ambient pH) than between the two acidified solutions.

Statistical Significance

All of the statistical significance tables are the results of single variable ANOVA tests performed on the corresponding data. In order to be considered a significant result, ($p < .05$).

Shell Length Results

The amount of shortened shells in the 7.6 pH treatment increased by 18.25% from the proportion found in the ambient treatment (W). Although the relationship between pH change and percentage of larvae with shortened shells is not linear across the treatments, both of the acidified treatments showed declines in fully developed larvae shells by at least 18%. The differences between the two acidified treatments were smaller, resulting in a difference of only 10%.

The single variable ANOVA of the shell length morphology demonstrated statistical significance of the data set ($p = .029$), indicating that OA affects shell length in larval development.

Shell Smoothness

There was a decline in the proportion of smooth shells as pH decreased (Figure 4). The majority of larvae had non-pitted shells in the ambient treatment; moving to the 7.8 treatment, pitted shell proportions reached a high of 53%, then dropped slightly to 51.9% in the most acidic treatment.

Based on the single variable ANOVA test performed on these results, the shell morphology did not vary significantly with pH ($p = .13$); however, qualitatively there was a difference in shell pittance was noticed from the ambient ocean water to water with increased CO_2.

Internal Morphology Results

The amount of empty larvae shells increased from 0% to 22% across decreasing pH levels (Figure 5). There was an observed increase in unhealthy inner morphologies across decreasing pH treatments as well; in the ambient pH the majority of larvae had full internal development. A pH change of .3 units resulted in a significant portion of these larvae with unhealthy, muddled morphologies, often interfering with the shell development as well.

A single variable ANOVA test was run on these data and statistical significance was found between the relationship of pH and inner morphologies ($p = .017$).

Movement Results

Acidified larvae had lower average movement than those reared in present day

treatments. Ambient pH larvae averaged a 2.125 speed on the standardized scale, with the 7.8 treatment scoring 1.75 and the most acidic treatment 1.125 (Figure 6). Morphological differences in shell development and total mortality could have also affected the number of individuals swimming and their speeds. Experiment setup – a cover slip on a flat glass slide – could have also had an impact on larval movement, as well as shock from the pipettes and transfer.

Although negative observational trends were observed in the movement analysis, the above data were not found to be significant following a single variable ANOVA test ($p = .22$).

Discussion

Larvae that developed in lower pH had significantly higher proportions of individuals with shortened shells and unhealthy inner morphologies. A higher proportion of pitted shells and reduced larval movement were also present in acidified treatments; despite observational trends, data collected were not found to be statistically significant. In essence, embryonic development in acidified conditions caused both morphological and functional changes in dactylomela larvae. Some marine invertebrate larvae are vulnerable to environmental changes and will likely be affected by increased ocean acidification. These changes may result in various implications on the coral reef ecosystem and the numerous symbioses within.

One potential implication for the morphological changes in larval development is a negative effect on swimming performance.[11] In the experimental pH treatments, fewer larvae were swimming at the speeds observed in the ambient treatment. Larvae were swimming less emphatically and covering smaller distances. Decreased mobility is just one of many factors that may influence the overall fitness and competitiveness of *A. dactylomela* larvae.

Impeding the calcification of a protective shell during the larval state of marine invertebrates could affect the life cycle and prevalence of the species, and have a compounding effect across the entire food web.[6] An improperly developed shell could leave sea hare larvae exposed to harmful UV and make them more vulnerable to predation, lowering their chances of survival. There are few studies done on this particular species of sea hare; however, there have been a number of research projects undertaken on other calcifying marine invertebrates such as.[11] The majority of these experiments were done in short time windows of 2-4 days, similarly exposing their marine embryos to differing pH treatments and not feeding their larvae when they hatched. A longer run experiment of larval sand dollars was done over a period of 10 days that included larval feeding also resulted in smaller skeletal masses and morphologies different to those raised in present day P_{CO2} conditions. Overall, both larval sea urchins and sand dollars have been shown to suffer reduced calcification when reared in lower pHs.[4, 12]

Ocean acidification has been shown to induce morphological changes that may decrease the competitiveness of multiple species of planktonic larvae. These organisms play critical roles in coral reef ecosystems by connecting existing populations, contributing to nutrient cycling, and maintaining biodiversity.

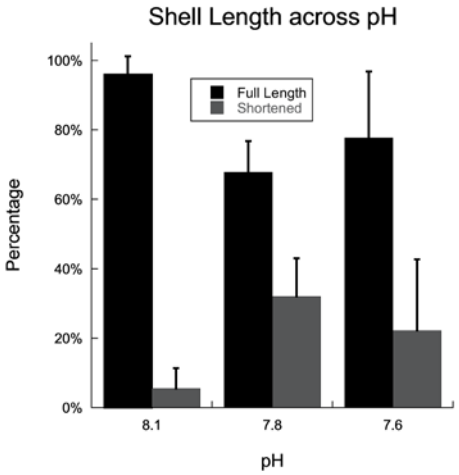


Fig. 3 The proportion of full-length shells in acidified treatments was significantly lower than in ambient pH (8.1). Acidified treatments had higher proportions of shortened shells.

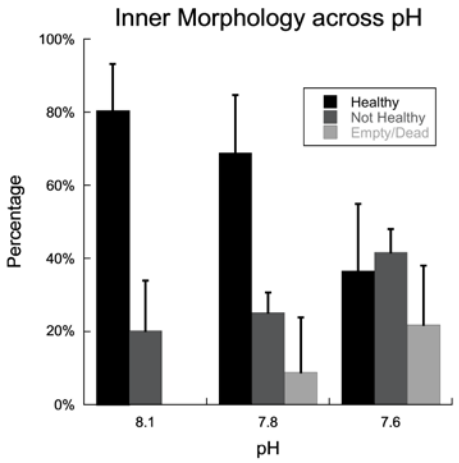


Fig. 5 The percentage of healthy inner morphologies in sea hare larvae dropped significantly as pH was lowered. The amount of empty shells increased from 0% to 22% across pH levels ($p>.05$).

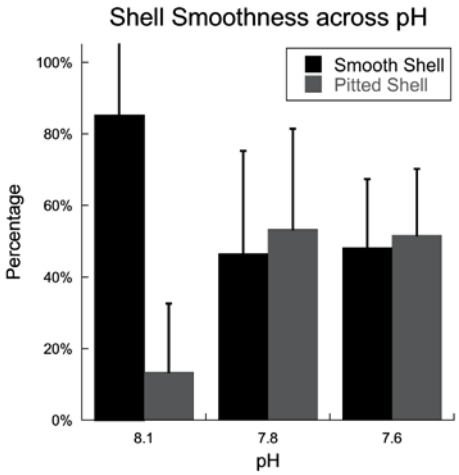


Fig. 4 The proportion of pitted shells increased from ambient pH (8.1) to acidified treatments. The acidified treatments had similar proportions of around 50% pitted and smooth shells. Although observational trends were present, the results were not found statistically significant ($p=.13$).

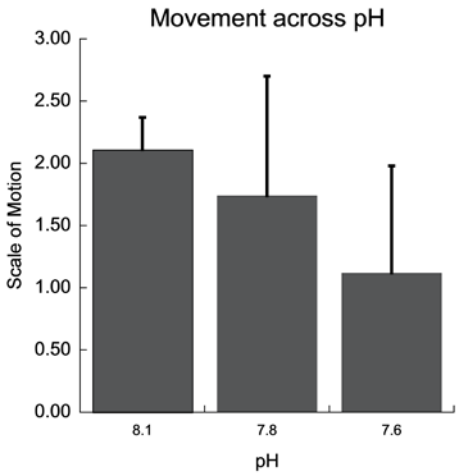


Fig. 6 Average larval speed and proportion of the population swimming declined with further deviation from ambient pH (8.1) by 53%. The scale of motion incorporated swimming speeds and proportion of larvae in motion. This data set was not found to be statistically significant; however, clear patterns across treatments were observed.

Furthermore, longer exposure to acidified conditions beyond embryonic development could cause these organisms to suffer even more drastic effects from the decreasing oceanic pH. These changes could ultimately reduce the number of individuals that move on to the juvenile life stage and eventually adulthood. Changing their population structure could have wide impacts for the health of coral reefs.

It is therefore important that further research be conducted on the effects of ocean acidification on larval development. Increasing exposure times in treatment could shed light on long-term impacts. Furthermore, a detailed study on mortality could contribute to the overall picture of future population structure on reef flats. Quantitative data on shell length, smoothness, and movement may illuminate the scale of disturbance ocean acidification could cause. Further studies should also focus on the adaptive capability of marine organisms, which will be crucial to forecast how marine organisms and ecosystems will respond to impending ocean acidification.

Understanding the future of sea hares in acidified oceanic conditions is an important element in maintaining healthy coral reefs. Many calcifying organisms will face falling oceanic pH levels within the century.[7] Corals themselves, the building blocks of reef ecosystems, will suffer lower calcification rates and slower accretion, increasing their vulnerability.[8] OA impacts on other marine organisms that contribute to coral health will further deteriorate the fragile ecosystem. The loss of herbivore grazing is a well-known process that can shift reef communities from coral to algae-dominated regimes.[8] Marine herbivores graze on one of corals' most threatening antagonists - algae. Reducing the viability of marine herbivores such as the spotted sea hare can increase the competitiveness of algae and pose a threat to coral reefs.[13] Corals and algae compete for space and light, and fleshy macroalgae may out-compete corals particularly under situations of low herbivory.[8] An algal dominated coral reef is a sign of unhealthy and stressed ecosystem, as well as limited biodiversity.

Ocean acidification has demonstrated visible effects on calcifying marine invertebrates such as dactylomela. Developmental disturbance in planktonic larvae could have reverberating impacts in the juvenile and adult stages of many reef organisms, as well as alter the food web and nutrient cycling in coral reef communities. Although ocean acidification is a complicated process with many future changes that will be difficult to predict, the effects of ocean acidification on the early life stage development of many calcifying organisms may ultimately decrease the resilience of reefs to algal takeover and undermine the overall health of these ecosystems.

References

1. Treut L et al. "Historical Overview of Climate Change Science." Earth, 2007; n. pag.
2. Sabine CL et al. "The Oceanic Sink for Anthropogenic CO2." New York: Science, 2004; 305.5682 : 367-71.
3. Solomon, S. Technical Summary. 2007.
4. Yu K et al. "Effects of Ocean-acidification-induced Morphological Changes on Larval Swimming and Feeding." Journal of Experimental Biology, 2011; 4: 3857-3867.
5. Orr JC et al. "Anthropogenic Ocean Acidification over the Twenty-first Century and Its Impact on Calcifying Organisms." Nature, 2005; 437.7059: 681-6.
6. Vézina, Alan, and Hoegh-Guldberg O. "Effects of Ocean Acidification on Marine Ecosystems." Marine Ecology Progress Series, 2008; 373: 199-201.
7. Dupont, S and Thorndyke MC. "Impact of CO2 -driven Ocean Acidification on Invertebrates Early Life-history – What We Know, What We Need to Know and What We Can Do." Biogeosciences Discussions 2008 (2009) : n. pag..
8. Diaz-Pulido G et al. "High CO2 Enhances the Competitive Strength of Seaweeds over Corals." Ecology, 2011;156-162.
9. Anthony KRN et al. "Ocean Acidification and Warming Will Lower Coral Reef Resilience." Global Change Biology 17.5; 2011 : 1798-1808.
10. Blankenship, J E, and R E Coggeshall. "The Abdominal Ganglion of Aplysia Brasiliana: A Comparative Morphological and Electrophysiological Study, with Notes on A. dactylomela." Journal of Neurobiology, 1976; 7.5: 383-405.
11. Todgham AE, Hoffmann GE. "Transcriptomic Response of Sea Urchin Larvae Strongylocentrotus Purpuratus to CO2-driven Seawater Acidification." The Journal of Experimental Biology, 2009 212.Pt 16 : 2579-94.
12. Clay, Tansy, and Grünbaum D. "Swimming Performance as a Constraint on Larval Morphology in Plutei." Marine Ecology Progress Series, 2011; 423 : 185-196.
13. Anthony K R N et al. "Ocean Acidification Causes Bleaching and Productivity Loss in Coral Reef Builders." Proceedings of the National Academy of Sciences of the United States of America, 2008; 105.45 : 17442-6.

Beckie Yanovsky is a Junior Class President and pursuing her B.S. in Biology at Stanford University. Beckie loves engaging with the Stanford community through her activities as a campus Tour Guide, Camp Kesem counselor, and member of Kappa Alpha Theta Sorority. Having recently returned from a quarter abroad through Stanford's Australia program, she is currently pursuing research in the Weissman Stem Cell Lab on campus and considering various graduate school opportunities for the future.

Virtual screening identifies novel inhibitor scaffolds of human carbonic anhydrase VII

Damir Ljuboja

Human carbonic anhydrase VII f is directly involved in an array of biological processes, including acid-base regulation, bone restoration, respiration, calcification, and the formation such compounds as saliva. Prior work has suggested the neural activity of CA7 in the mammalian brain, particular its high-frequency stimulation of GABAergic excitation, is directly correlated to seizures in epilepsy. This research was focused on the discovery of novel inhibitors to decrease the global impact of epilepsy. Ligation independent cloning, expression in E. coli cells, and affinity chromatography techniques were used to obtain carbonic anhydrase 7 protein for in vitro testing and analysis. Protein function was verified in enzyme assays varying enzyme and substrate, and assays varying 100% DMSO concentration were utilized to determine optimal conditions for inhibitory tests. Virtual drug screening with the GOLD docking genetic algorithm was used to collect top-scoring compounds from eight libraries (122,670 compounds). The top-scoring compound from an aggregate run, in addition to a top-scoring compound for another researcher, was tested using inhibitory assays. These compounds, ChemBridge 5107893 and 5115900, both displayed inhibition with projected IC50 values in the high micromolar (fractional millimolar) range. These data support the use of virtual screening for the discovery of novel inhibitors and have important implications for the arrangement of scaffolds and skeletal orientations of potential synthetic inhibitors of carbonic anhydrase VII and, ultimately, the global impact of seizures and epilepsy.

Epilepsy is one of the most common neurological disorders known, impacting roughly 1% of the planet's population, about 67,752,358 people[1]. carbonic anhydrase VII (CA7), a cytosolic member of the CA family, is a zinc metalloenzyme that participates in a wide variety of biological processes, including acid-base balance, bone restoration, respiration, calcification, and the formation of several compounds in the body (saliva, etc.). The enzyme is primarily known to catalyze the Zinc 2+ ion-dependant hydration of carbon dioxide to form bicarbonate. A primary difference between CA7 and other carbonic anhydrases is its location in the cortex, hippocampus, and thalamus regions of the mammalian brain and role in neuronal excitability, leading to direct involvement in seizures and epilepsy. It is due to this significance, particularly in regard to its generation of high-frequency stimulation, or HFS-induced GABAergic excitation, that carbonic anhydrase VII is a drug target. Although it has many names and isoforms/isozymes, CA7 is titled "carbonate dehydratase" by the BRENDA enzyme database. Its enzyme number (E.C.) is 4.2.1.1. and its only cofactor is Zinc. It molecularly functions as a lyase and the reaction mechanism CA7 catalyzes may be explained by the following chemical equation:

$$H_2CO_3 \leftrightarrows CO_2 + H_2O$$

The source organism in question is humans, although the protein is ubiquitous in prokaryotes and eukaryotes alike. In the lab it is commonly overexpressed in *E. coli*. Aside from Zinc, ligands include glycerol and 6-ethoxy-1,3-bezothiazole-2-sulfonamide (EZL, chemical formula $C_9H_{10}N_2O_3S_2$). Even though no CA7 inhibitor drugs can be found in the marketplace, some drugs designed for carbonic anhydrase (as a whole) include Diamox (the only one said to specifically target Epilepsy via Drugs.com), Topamax, and Zonegran.

The shortcomings of these drugs,

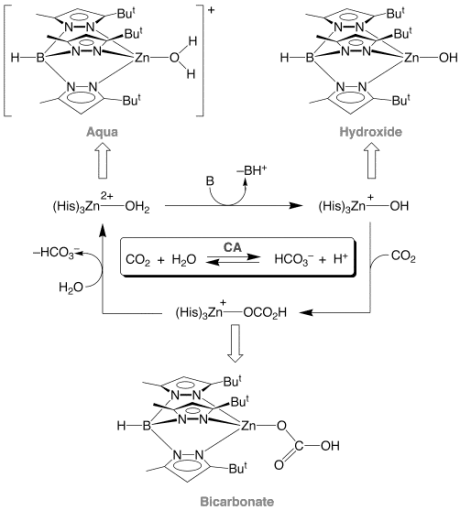


Fig. 1 Pathway featuring the central involvement of carbonic anhydrase VII to catalyze a vital reaction to the formation of several products.

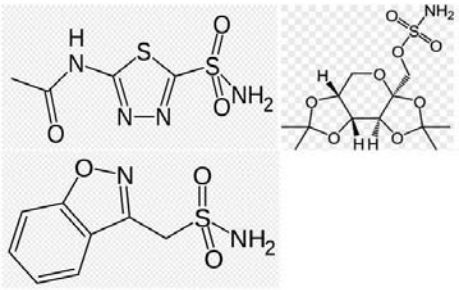


Fig. 2 Structural diagram of Diamox (top-left), Topamax (top-right), and Zonegran (bottom).

Table 1 Lipinski Data for Diamox, Topamax, and Zonegran, three inhibitors of CA7 (or the CA family in general).

Name	logP	Weight (Daltons)	HBA	HBD
Diamox (acetozolamide)	-.39	222	6	2
Topamax (topiramate)	1.29	339	9	1
Zonegran (zonisamide)	.67	212	5	1

however, are in their side effects. Diamox is said to cause vomiting, fever, confusion, blood in the urine, painful urination, yellowing of the skin and/or eyes, seizures, and unusual bleeding. Topamax had a list of approximately 30+ side-effects, some of the worst including memory problems, depression, uncontrollable shaking of body parts and eyes, diarrhea, constipation, swelling of tongue, bloody urine, skin color changes, blurred or double vision, chest pain, irregular heartbeat, and back, muscle, or bone pain,. According to NCBI, Zonegran may cause vomiting, confusion, difficulty with memory, pain and numbness, worsening or longer-lasting seizures, bloody urine, severe weakness and muscle pain, fast/shallow breathing, loss of consciousness, and more. As evident by this, the need for novel, safer inhibitors for carbonic anhydrase is outstanding. According to the EXPASY Protparam program, the protein sequence for carbonic anhydrase VII has a pI isoelectric point of 5.05, a molecular weight of 66244.9 Daltons, and a molar extinction coefficient of 15625 $M^{-1}cm^{-1}$ at 280 nm in water. The three-dimensional structure of CA7 exists in the literature under the Protein Data Bank identifier 3MDZ and was determined by x-ray crystallography to 2.32 angstroms resolution[2].

Materials and Methods

Cloning

The plasmid vector bearing the CA7 coding DNA (hsCD00041026) sequence was obtained from the Arizona State University PSI and the Harvard PSI. However, according to the obtained sequence of the clone, there does not appear to be a His6 tag on it although it is expected to be present (ASU PSI website, Harvard PSI website). The DNA was amplified via PCR and sticky ends were generated on the PCR inserts to clone the DNA into a pNIC-Bsa4 expression vector containing a His6 tag. The unique PCR settings involved 66.7 °C for 15 seconds during 20 annealing cycles in a final concentration of 1.5 mM $MgCl_2$. The forward primer used was TACTTCCAATCCATGACCGGCCAC CACGGCTGG and the reverse primer was TATCCACCTTTACTGCTACAAGGCCC GGA AGGAGGCC. DpnI treatment was done on the gene to remove the Kanamycin resistance and the BsaI enzyme (10 units) was used to cut the gene and *E. coli* plasmid for cloning.

T4 DNA polymerase and dNTPs (final concentration 2.5 mM) were added to the PCR inserts. The sticky ends were mixed with a final concentration of 5 mM DTT and a final concentration of 0.1 mg/mL 100x BSA, mixed for 30 minutes at 22 °C and heat inactivated at 75 °C for 20 minutes. The accepting vector was combined with dGTP final conc. 2.5 mM, 100 mM DTT final conc. 5 mM, and 100x BSA final concentration of 0.1 mg/mL, kept at 22 °C for 30 minutes, and heat inactivated at 75 °C for 20 minutes. The two were then mixed in the following ratios: 1 tube was 1:2 pNIC-Bsa4: CA7 and the other was 1:3 pNIC-Bsa4: CA7. Construct was cloned in chemically competent DH5α *E. coli* calls and successful ligations were indentified using pLIC-for and pLIC-rev primers. (Figure 3) The confirmed clones were then scaled up and expressed in protein via transformation in BL21 (DE3) Novagen competent cells (EMD Chemicals, Gibbstown, NJ).

Expression and Purification Conditions

In the literature, methods for human CA7 purification involved agarose resin column chromatography, carbonic anhydrase inhibitor affinity chromatography, Ni-NTA resin affinity chromatography.[3] In terms of expression, articles in the literature have verified success resulting from cDNA expressed in *Escherichia coli* BL21 (DE3).[4] The colonies were transformed overnight and kanamycin was added the next day for a final concentration of 50 µg/mL. The colonies were transferred to a 500 LB culture, which was grown until the OD_{600} was between .6 - .8. Protein (T7 polymerase) expression was induced with the addition of IPTG to a final concentration of 500 µM and the culture was grown for 4 hours at 37 °C on a shaking incubator and then harvested by centrifuging for 20 minutes at 4 °C in a Beckman JA10 rotor at 6000 g. The next day, the pellet was resuspended with 5-10 mL lysis buffer (100 mM NaH_2PO_4, 300 mM NaCl, 10 mM Imidazole) and sonicated with output at 5.5 using the following cycles to break open the cells – 1x30 seconds at 70%, repeated after 1 minute, 1x30 seconds at 90% after another minute. The mix was then centrifuged for 30 minutes in a JA17 rotor at 20,000 *g* and then His-tag purified with a Bio-rad chromatography column and 5-prime Ni-NTA resin to obtain the protein. The column was cleaned with nanopure water, washed with 5 mL of wash buffer (pH 8 100 mM NaH_2PO_4, 300 mM NaOH, 30 mM Imidazole), and stripped of protein with 5 mL of elution buffer applied twice (pH 8 100 mM NaH_2PO_4, 300 mM NaOH, 300 mM Imidazole). The final product was characterized via PAGE-SDS gel and quantified using 260 nm wavelength on a NanoDrop ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE). For storage, the protein was combined with glycerol to a 10% final concentration and stored in a -20 °C refrigerator. The presence of the glycerol prevented freezing. Alternatively, BRENDA Enzymes indicated that storing the protein at 4 °C would keep it stable for up to a year.

Virtual Screening

Evidence of prior virtual screening done with CA7 is limited and doesn't appear frequently in the literature. Thus, any virtual screening done with CA7 could provide novel results that could pave the way for future testing. Although virtual screening experiments have been performed with other members of the carbonic anhydrase family (for example, carbonic anhydrase II), it was difficult to locate any verification that such tests had been carried out with carbonic anhydrase VII. The proposed screening with GOLD (Genetic Optimization for Ligand Docking, Cambridge, UK)used several different libraries to determine top scoring compounds and is independent of research done on other isoforms of CA7. Selectivity against these isoforms was verified through conducting screens using the top ranked compounds and binding them into structures of the said isoforms (obtained from the Protein Data Bank). A validation dock was performed aligning the provided ligand in the crystal structure of CA7 with its highest-scoring docked pose as determined by virtual screening to verify the functioning of the docking program. This was done with an autoscale of 1.0 and the top 10 poses were saved and then aligned to the original conformation with RMSD values shown in the molecular visualization program, PyMol (Schrodinger, Portland, Oregon).

The following libraries were screened using GOLD's (Genetic Optimization of Ligand Docking) genetic algorithm to randomly sample conformational space and evolve a more likely configuration for the ligand being evaluated:

HF9_180_Plates_1um3D_catnum.sdf, KINASet3d.sdf,

ChemBridge-diversity3D.sdf,
ION_Channel3D.sdf,
MolecularWeightSet.sdf,
FragmentLibrary.sdf,
zincSigmaLibrary3D.sdf,
and CB-kin_UT.sdf.

Each was run through an initial job using autoscale at .5, testing 10 poses for each ligand, and saving the top 10% of compounds in an output file. The Gold Fitness Score was used as an overall means of ranking the results – subscores included Shb_ext – protein-ligand hydrogen bond score, Shb_int – intramolecular bonds in the ligand, Svdw_ext – the protein-ligand van der Waals score, and Svdw_int – the contribution due to intramolecular strain in the ligand. A secondary job was performed with autoscale set to 1 with the input file set as the output file from the first run. The top 10% of ligand solutions from this run were saved and the top 10 compounds were evaluated for affirmation of Lipinski's rules of 5 and those which met all requirements were visually examined in PyMol for docking confirmation. The Lipinski rules of 5 adhere to the following limitations – not more than 5 hydrogen bond donors, not more than 10 hydrogen bond acceptors, molecular weight under 500 daltons, logP less than 5. These results were combined into an .sdf input file for an aggregate run using high accuracy (autoscale = 1) and the top 30 results were examined (Table 2). The top five were selected on the basis of drug-likeness and gold scoring function and the top one was ordered for inhibition assay testing.

Docking was carried out using parallel processing methods on the TI3D computer cluster at the University of Texas at Austin. The cluster has 16 HP Proliant BL35P blades, each of which retains two dual core AMD Opteron 2.4 GHz processors and 8 GB of memory. The compounds were ranked by fitness score and filtered (modified to include compounds with logP values <2.5). Between all the libraries screened, 122,670 compounds were evaluated and reduced to the top two compounds for *in vitro* testing (Figures 4-6).

Enzyme Assays

In previous cases, 96-plate inhibition assays were performed using various isoforms of CA7 for optimal results. The proposed method of assay conditions is explained in the following. The substrate used was PNPA (4-Nitrophenyl Acetate) (Acros Organics, New Jersey) in acetone and the buffer conditions were 50 mM Tris Buffer mixed in such a fashion that 20 ng/µL of enzyme was placed in 10 µL of Tris. The PNPA also served as the chromophore, becoming colorless when hydrolyzed[5]. This assay was performed at room temperature with pH adjusted to 7.6. Automated readings did not require for a separate reaction for each time point and thus a stop solution was not necessary. A Red Tide USB-650 spectrophotometer (Ocean Optics, Dunedin, Florida) was used with LoggerPro to record a kinetic measurement of absorbance at 348 nm wavelength every 15 seconds for 5 minutes, which directly evaluated the ability of the PNPA to bind into the enzyme active site (Figure 10a). The second enzyme assay involved the same conditions as the first but substrate amount was varied instead of enzyme amount (which was kept constant at 100 µL). (Figure 10b) The third assay involved varying the 100% DMSO concentration to determine optimal conditions (Figure 10c). This was done by regulating CA7 amount at 100 µL and PNPA at a 3 mM concentration. PNPA was combined with acetone and nanopure water in all cases to hydrolyze it and was kept at room temperature for use. Each of the aforementioned was performed a total of n =2 times and average changes in absorbance were noted in Microsoft Excel.

Inhibition Assays

Two potential inhibitors (ChemBridge ID 5107893 and 5115900) were tested, both by the same protocol, to determine their ability to bind to and inhibit the CA7 enzyme. These were performed under the same conditions as the enzyme assays (348 nm kinetic measurements every 15 seconds for 5 minutes, pH 7.6, room temperature) but there was an additional 5 minute incubation at room temperature for each cuvette prior to mixing in the PNPA substrate to allow time for the inhibitor to bind to the enzyme. The stock inhibitor (ChemBridge 5107893 - MW 311 daltons, LogP -2.52) was diluted from a 50 mM conc. to a 10 mM final conc. (with 20.4 µL stock, 30.6 µL 100% DMSO, and 51 µL Nanopure) after solubility tests determined it was water-soluble at that concentration. PNPA substrate was kept at 3 mM final conc. for all cuvettes and inhibitor concentration ranged from 0 mM to .6 mM final concentration. The amount of CA7 used in all cases was 100 µL (in a final volume of 500 µL). There was also a 9 mM final concentration of Tris buffer in each cuvette. The measurements were taken sequentially and while one cuvette was measured, the next was prepared (in ascending order by inhibitor concentration). This was done a total of two times and overall results were determined in Microsoft Excel (Figure 11). This same protocol was repeated for the other inhibitor selected (Figure 12).

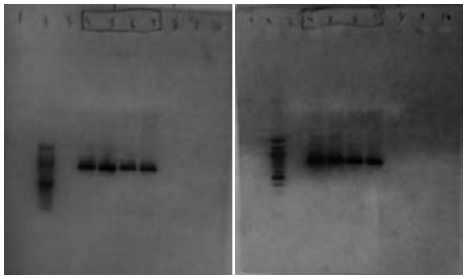


Fig. 3. PCR gel check attempts 1 (left) and 2 (right) with lanes correlating with the following in both cases 1) skip 2) 100 bp ladder 3) 6 ng CA7 0 Mg 4) 6 ng CA7 1.5 Mg 5) 6 ng CA7 2 Mg 6) 6 ng CA7 4 Mg 7) 6 ng CA7 6 Mg 8) control 9) skip 10) skip

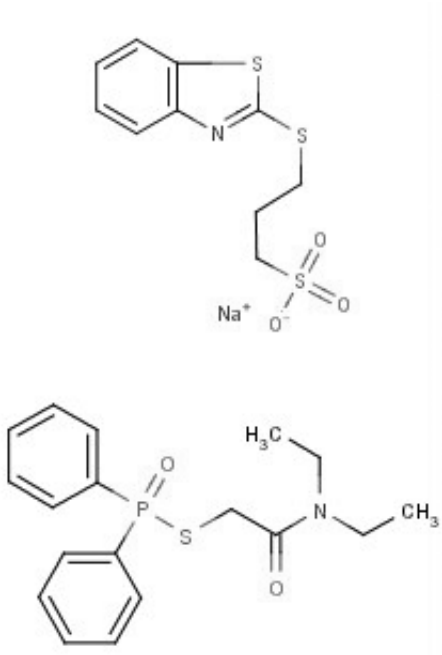


Fig. 4 2D representations of compounds 5107893 (top) and 5115900 (bottom).

Results

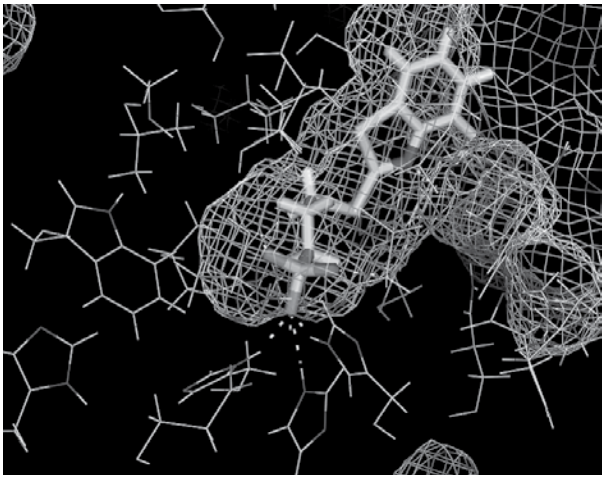


Fig. 5 Compound 5107893 (sticks colored by element) docked into the CA7 active site (mesh) with side chains visible (lines) and polar contacts (dashed lines).

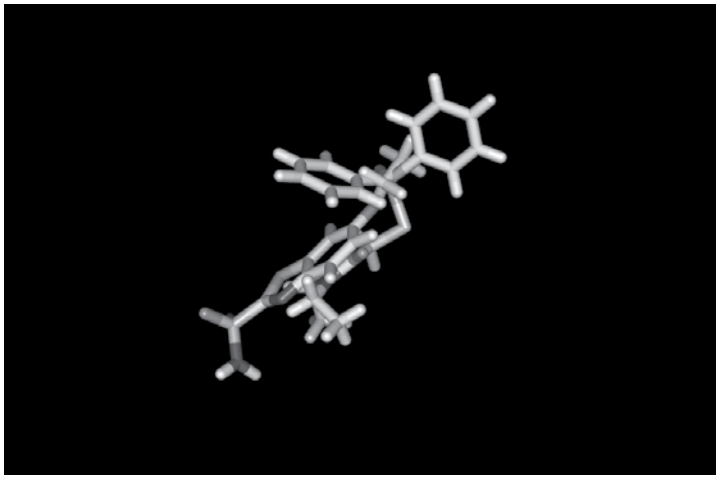


Fig. 7 Alignment between EZL known ligand (lower left) and Compound 5115900 (upper right) shown with 10 to 10 atoms overlaid and an RMS value of 3.661.

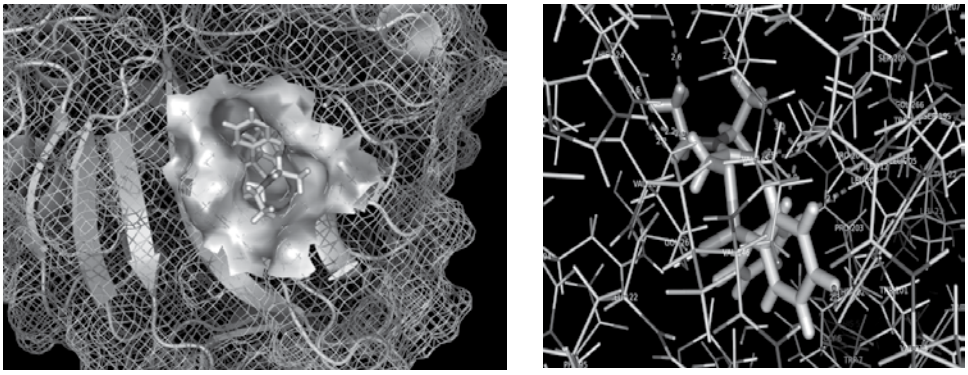


Fig. 6 (a) Compound 5115900 (shown as sticks) docked into the active site of CA7 (surface) with polar contacts shown as dashed lines. **(b)** Compound 5115900 (sticks) in the active site of CA7 (lines) with hydrophobic interactions shown as dashed lines. Hydrophobic interactions are shown primarily with PHE 133 (aromatic rings), PRO 204, and LEU 200, with a few between the compound and ALA 137, GLY 134, THR 101, and with VAL 145.

Rank	Fitness	S(	hb_e	xt)	S(vdw_	S(hb_i	S(int)	Ligand name	ClogP	MW	HBD	HBA	Re-Rank
1	79.13	2	0.34	44.19	0	-1.97		ZINC12405084	3.49	425	3	15	
2	78.44	2	1.51	43.65	0	-3.08		ZINC12405083	4.48	479	4	17	
3	72.84	3	0.9	32.41	0	-2.63		ZINC01809738	0.45	288	0	4	1
4	71.21	2	6.01	37.2	0	-5.95		ZINC03869814	1.56	329	3	11	
5	71.18	3	1.44	30.26	0	-1.87		ZINC12503609	3.76	401	3	14	
6	70.82	1	3.02	45.68	0	-5.01		5230098	2.92	354	2	3	
7	70.03	2	2.91	34.16	0	0.14		ZINC03869848	2.73	345	4	12	
8	69.93	3	3.06	32.88	0	-8.34		ZINC03869815	1.56	329	3	11	
9	69.78	2	7.32	35.81	0	-6.78		ZINC03869815	1.56	329	3	11	
10	69.11	1	1.48	50.88	0	12.33		5159191	5.42	495	3	1	
11	68.56	1	5.23	45.53	0	-9.29		5331849	1.82	356	2	4	2
12	68.11	2	4.09	34.44	0	-3.33		ZINC12296728	2.72	345	4	12	
13	67.95	2	2.38	34.6	0	-2.01		5135086	0.48	284	0	4	3
14	67.57	1	4.19	42.35	0	-4.86		5229993	2.33	360	1	5	4
15	67.56	1	3.39	47.66	0	11.36		5143462	3.14	433	4	2	
16	66.61	1	9.75	39.2	0	-7.04		5226748	3.43	356	2	2	
17	65.11	1	2.94	41.77	0	-5.26		5352822	4.05	393	1	4	
18	65.00	1	5.97	38.44	0	-3.83		5185876	3.82	407	1	4	
19	64.93	1	0.55	42.06	0	-3.46		5648183	1.25	315	1	6	5
20	64.81	1	1.83	43.3	0	-6.57		5355734	3.51	296	2	1	

Table 2 Virtual Screening Top Results (ranked on the right after filtering with Lipinski's rules) with Shb_ext – protein-ligand hydrogen bond score, Shb_int – intramolecular bonds in the ligand, Svdw_ext – the protein-ligand van der Waals score, and Svdw_int – the contribution due to intramolecular strain in the ligand. light gray denotes satisfying all filtering, dark gray does not satisfy Lipinski rules, and white satisfies Lipinski's rules but not subsequent LogP filtering (less than 3).

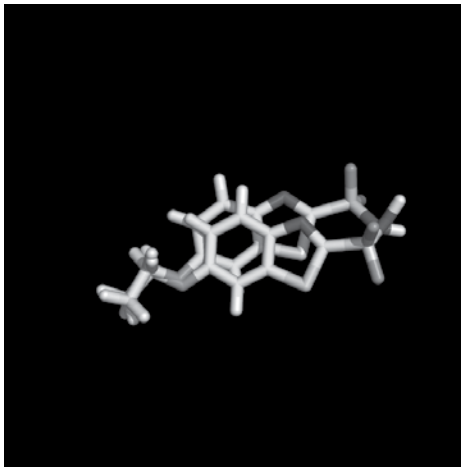


Fig. 8 Redocked EZL (sticks) binding in CA7 (lines) with two polar contacts displayed as dashed lines - pose determined by GOLD with a fitness score of 87.39 .

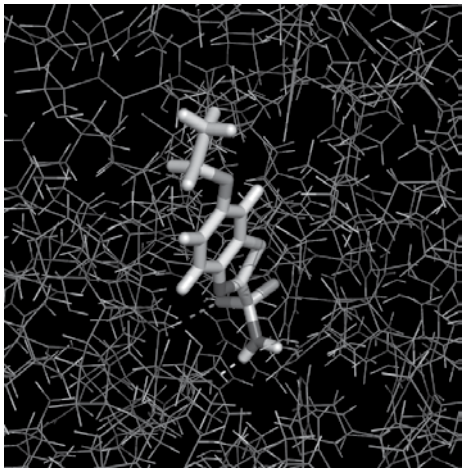
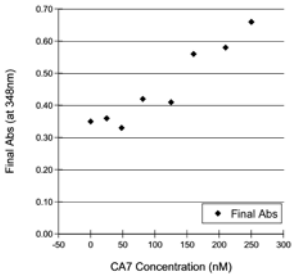
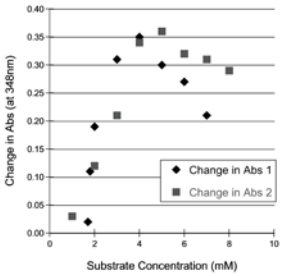


Fig. 9 Superimposition of validation dock EZL (in front) and actual EZL (behind) originally extracted from CA7) shown as sticks with an RMS Value: 1.837 (10 atoms to 10 atoms).





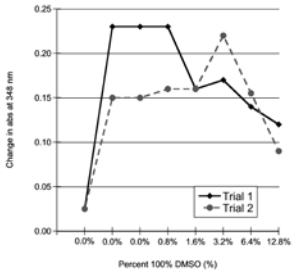
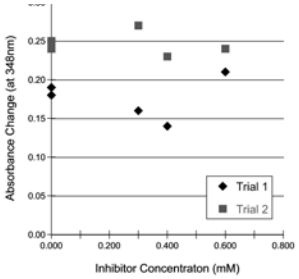


Fig. 10 (a) Varying enzyme amount assay results with kinetic measurements taken at 348 nm every 15 seconds for 5 minutes, pH 7.6, room temperature, and 1.5 mM PNPA substrate. **(b)** Varying substrate amount assay results with kinetic measurements taken at 348 nm every 15 seconds for 5 minutes, pH 7.6, room temperature, and 2 mM CA7 per cuvette. **(c)** Varying 100% DMSO amount assay results with kinetic measurements taken at 348 nm every 15 s for 5 minutes, pH 7.6, room temperature, 2 mM CA7 and 3 mM PNPA substrate.

Discussion

After the assays were performed and data was analyzed, it was determined that both compounds showed some degree of inhibition.

Compound 5107893 displayed virtually no inhibition when diluted to a stock concentration of 10 mM (Table 5) but, when brought down to 5 mM, the compound seemed to inhibit CA7 a great deal more (Table 6). Given the data in Table 7, it was estimated that the IC_{50} of compound 5107893 was somewhere in the high micromolar/fractional millimolar range of about .8 mM (800 µM). Compound 5115900 showed a similar level of inhibition (Table 7) but did so when diluted to a stock concentration of 10 mM. It was also estimated to have an IC_{50} in a similar range – roughly .9 mM (900 µM). The desired IC_{50} values for potential novel inhibitors would have been, ideally, in the nanomolar (nM) or low micromolar (µM) ranges. Although these two compounds fall just outside of this limitation, they attest to the ability of virtual screening to find compounds with inhibitory capabilities



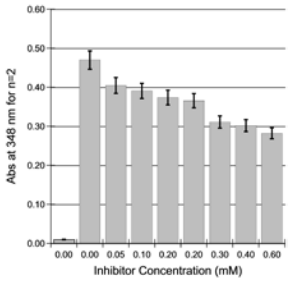
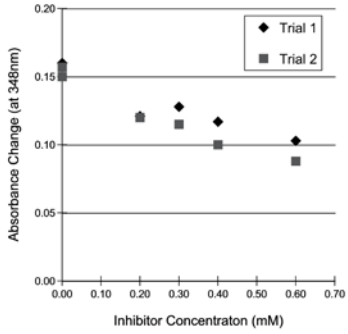


Fig. 11 Inhibition assay results of compound 5107893 with kinetic measurements taken at 348 nm every 15 seconds for 5 minutes (pre-incubation of 5 minutes before PNPA addition), pH 7.6, room temperature, 2 mM CA7, and 3 mM PNPA substrate. Compound was **(a)** diluted to 10 mM **(b)** diluted to 5 mM with standard deviation error bars shown.



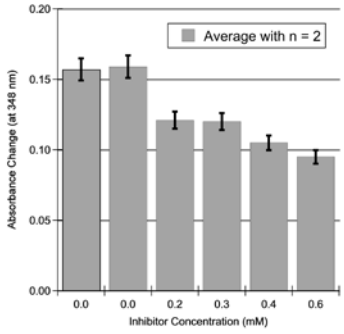


Fig. 12 Inhibition assay results of compound 5115900 with kinetic measurements taken at 348 nm every 15 seconds for 5 minutes (pre-incubation of 5 minutes before PNPA addition), pH 7.6, room temperature, 2 mM CA7, and 3 mM PNPA substrate. Compound was **(a)** diluted to 10 mM **(b)** diluted to 5 mM with standard deviation error bars shown.

that could be modified (via methylation, acetylation, bond creation/cleavage, etc.) and tweaked to effective inhibitors that are able to outcompete natural substrate.

There were some concerns regarding the differences observed in the concentration of the stock used for compound 5107893 in correlation with its ability to inhibit CA7. There are several factors that may have impacted this, including 100% DMSO concentration, enzyme concentration, and correct protein folding during the expression step. Also, it is possible that the affinity chromatography did not fully purify the protein, leaving behind other proteins that were able to catalyze a similar reaction. Protein folding mistakes could have made the active site different from that predicted by the x-ray structure, and thereby less accessible to the inhibitors. The presence of only one PBD file of CA7 also created a dependency on the structure's correct representation of the actual protein. All binding accuracy and potential success was based on this model and, in the case that there were inaccuracies with its design, they could have result in failure regarding ligand binding.

Although this was not the case and the assays proved successful, there are uncertainties regarding the accuracy of the values obtained (inhibition could in actuality be much greater or much lower). This stems from many variables and thereby creates an array of sources of error, amongst which are assay conditions, correct protein function and folding, and purification. This could detract from the overall prediction of the GOLD program, which considered ideal conditions when it made its evaluation.

Difficulties were encountered mainly in the PCR and expression steps of protein obtainment and failed several times. This could have led to a lesser degree of confidence in the presence of the correct protein prior to the assays. However, taking this into consideration, success was still obtained in verifying the overall reliability of the virtual screening at predicting effective inhibitors, particularly involving the discovery of the groundwork for synthesizing more effective inhibitors via modification of bonds. The work done here provides the basis for further optimization of two separate classes of compounds to inhibit carbonic anhydrase VII.

Conclusion

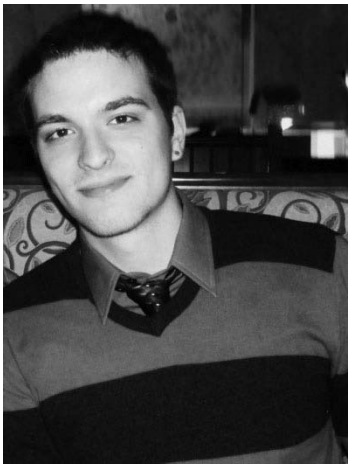
Using the GOLD genetic screening algorithm with 8 libraries of 122,670 compounds total, several compounds were determined to have potential inhibitory interactions with carbonic anhydrase VII. Using a combination of PCR ligation-independent cloning, transformation, affinity purification, and enzyme assays, CA7 protein was obtained and conditions were optimized for inhibition assays. Two compounds, 5107893 and 5115900, displayed inhibition with projected IC_{50} values in the high micromolar (fractional millimolar) range. Although these values were not ideal, they signified the effectiveness of virtual screening at predicting novel inhibitors and, more significantly, produced two compounds which, with further optimization (via functional group addition, methylation, etc.) could be made to inhibit CA7. This highlights the ability of docking programs to discover skeletal compounds which could be subsequently modified into effective inhibitors with IC_{50} concentrations in the desired nanomolar to low micromolar range. Future research entails optimizing assay conditions and the protein obtained (via expression and affinity purification) to validate and improve the accuracy of these results. Also, the use of derivatives of these compounds could be tested via virtual screening and in assays to evaluate their potential as CA7 inhibitors. Compounds 5107893 and 5115900 provide a good foundation for further drug development.

Acknowledgements

My sincerest gratitude goes to Dr. Josh Beckham, Adam Nguyen, Charles Nwaogu, and Rebecca Nekolaichuk. Without them, this project would not have been possible.

References

1. Ray A, Pathak-Ray V, Walters R et al. Driving after epilepsy surgery: effects of visual defects and epilepsy control. British Journal of Neurosurgery. 2002; 16:456–460.
2. Montgomery J, Venta P, Eddy R, et al. Characterization of the human gene for a newly discovered carbonic anhydrase, CA VII, and its localization to chromosome 16. Genomics 1991; 11:835-848.
3. Grüneberg S, Stubbs M, Klebe G. Successful virtual screening for novel inhibitors of human carbonic anhydrase: strategy and experimental confirmation. J Med Chem. 2002; 45:3588-602.
4. Zhao X, Liu X, Guo W et al. Expression of carbonic anhydrase IX in NSCLC and its relationship with VEGF and Ki67 expression. Zhongguo Fei Ai Za Zhi 2010; 13:861-866.
5. Armstrong J, Myers D, Verpoorte J et al. Purification and properties of human erythrocyte carbonic anhydrases. J. Biol. Chem. 1966; 21:5137-5149.



Damir Ljuboja is a junior from Sarajevo, Bosnia, double majoring in Biology Honors and Humanities Honors at the University of Texas at Austin. Alongside exploring his scientific research interests, Damir enjoys rowing in the sweltering southern weather, producing electronic music, and writing poetry. He will apply to medical school this summer but remains open to the possibilities of the future, whether they lie in an academic hospital setting, a translational research lab, or a clinic abroad.

Physical properties of habitats occupied by platypus (*Ornithorhynchus anatinus*) in undisturbed and recently disturbed sites along rifle creek, Far North Queensland

Weston J. LaGrandeur

The platypus, (*Ornithorhynchus anatinus*), is an endemic species to Australia that spends all of its life in and around riverine areas. This study compared physical characteristics of four different sites (two undisturbed, two disturbed less than one year ago) located on Rifle Creek in Far North Queensland. The site with more platypus sightings was grazed by cattle until January 2011. The two previously disturbed sites had a higher average velocity, bank angle closer to ninety degrees, taller bank heights and depths between one and three meters. The results indicated that the platypus selects foraging and living sites based upon characteristics such as average velocity, bank angle, bank height and stream depth. Furthermore, the construction of riparian fencing along this stretch of land in January 2011 showed rapid improvement regarding platypus habitat. There was no correlation between platypus sighting and higher macroinvertebrate abundance suggesting that the platypus responds to physical properties rather than invertebrate abundance.

The platypus, (*Ornithorhynchus anatinus*), is a semi-aquatic, egg laying mammal that is found in the freshwater ecosystems of Australia.[1] The platypus is one of three living species in the class monotremata, but is physiologically unique as the other two species are echidnas.[2] *Ornithorhynchus anatinus* has a long list of unique physiological characteristics including the electromagnetic sensitivity in its bill, a poisonous barb and its egg laying ability. Platypuses are endemic to Australia, more specifically eastern Australia, where they can only be found in freshwater where they nests and rests in burrows situated in banks along the freshwater in which they feed.[3,4] They are crepuscular and feed upon aquatic macroinvertebrates. The platypus is sensitive to electric waves in the water created by escaping prey which allows them to hunt in water with little visual aid.[5] The platypus can cover a net distance of up to 3.4 kilometers in any given feeding period.[2] Pursuant to its endemism to Australian freshwater, the platypus occupies many habitats at varying elevations, inland and coastal areas, glacial runoff, and manmade bodies of water. They have also been reported to inhabit backwaters and isolated pools filled by flood.[3]

Despite being listed on IUCN's Red List at a level of "Least Concern", many threats in to the platypus exist: primarily resulting from human-based actions, these threats include hunting, habitat loss, and rising global temperatures. Up until the early 1900s, fur hunting along with drowning in freshwater fishing nets drastically reduced the numbers of platypuses.[3,6] Human development, pollution, construction and the removal of freshwater sources due to the industry of an increasing human population still threatens much of the platypus habitat. Further studies suggest that the main factor correlated to available platypus habitat switched in the 1960s from riparian and aquatic morphology to differences in temperature sparked by global warming.[3,7]

Studying river and bank morphometrics assays several main attributes to the lifecycle and survival including the burrows where platypuses spend the time while not feeding. The macroinvertebrate food source of the platypus is mainly found in river sediment. Contradictory studies exist concerning the platypus' preference to differing river sediment habitats. Milione *et al.* 2009 found that there were less platypus sightings in rivers with a high gravel density whereas Serena *et al.* 2001 found that platypus were more prevalent in areas with more coarse grained soil.[8,9] River velocity is important when considering platypus habitat because a faster current often means more effort on behalf of the platypus. Because of this, they may avoid fast flowing water if calmer areas are available.[1] Furthermore, bank height and sediment composition can alter a platypus' preference for burrow construction and habitation. In another study done by Serena *et al.* in 1998, 77% of burrows were located in banks that were 1-2.5 m tall.[4] Many factors within river and bank morphology appear throughout numerous studies and create a common thread of importance that impels the study at hand.

Human disturbance, varied river morphology and contradicting data in past studies lead to this question: Do physical river properties and river morphology, including past disturbances by human activity, influence the preference of platypuses to inhabit one area over another? Furthermore, is riparian fencing is an effective rehabilitation technique with concern to platypus habitat? In concordance with background research, it is hypothesized that the river properties influence the factors of an optimal platypus habitat. Specifically, slower rivers and sediment density influence platypus abundance. Finally, physical river properties influence optimal platypus habitat more than resource abundance of invertebrates.

Materials and Methods

In Far-Northeast Queensland, four observation/data collection sites were defined (Figure 1). The sites named Wetherby A and B were located upon Rifle Creek on the property of Wetherby Station. The other two sites were located in the Sheoak Ridge Nature Reserve and are upstream of the Wetherby sites on Rifle Creek. Using transect measuring tapes, 25 meter transects were set up on the trail side of the river at each of the four sites. To mark the sites, pink marker tape was tied to riparian growth at 0 m, 12.5 m and 25 m along the transect. The observation of platypus activity was conducted along this transect at each of the four sites from

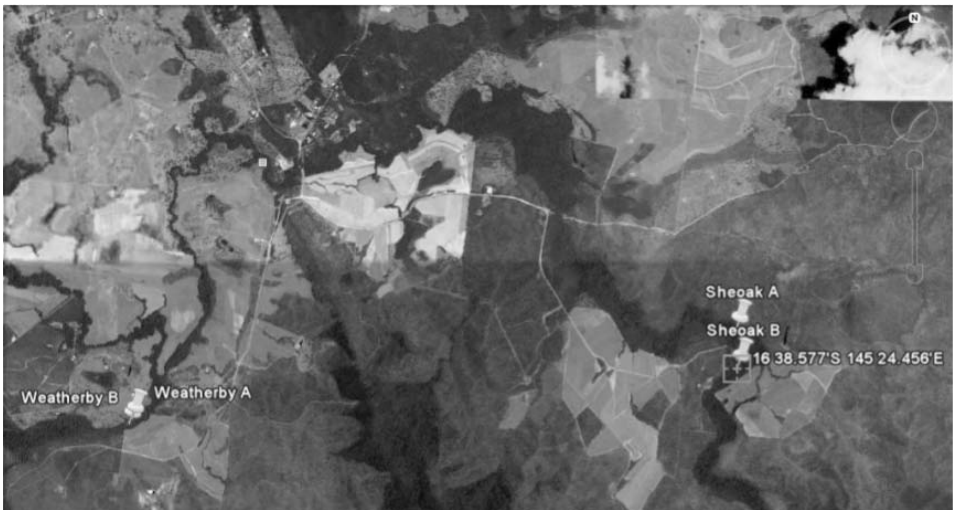


Fig. 1 A satellite image of the four sites. The Wetherby sites are surrounded by cattle pasture and sugarcane fields. The Sheoak sites are surrounded by thicker, undisturbed riparian vegetation with sugar cane paddocks adjacent to the river.

5:45 to 6:45 AM and 5:00 to 6:00 PM on October 14, 2011 through to the morning of the 17th. The first morning observation on October 15, 2011 was taken from 5:30 to 6:30 AM but was adjusted to the later time due to the lack of visibility for the first 15 minutes. Observers noted when the platypuses emerged and, if possible, how long the platypuses were at the surface, taking note of dive time. All four sites were observed at the same time. In total, the research team tallied 24 hours of observation.

At each of the four sites, river sediment samples were taken at the 0 m, 12.5 m and 25 m marks along the transect. The sediments were collected in a plastic specimen jars at the deepest point of the river along the respective distances of the transect. The river sediments were separated using a progression of sieves. First, the total amount of sediment was weighed using a metric scale. Then, the sediment was sifted using a 2 mm sieve. Gravel (>2 mm) was then weighed and put back into the sample jar. The remaining sediment was sifted in a 1 mm sieve. Sand (.5-1 mm) was weighed and set aside. Finally, the last of the sediment was determined to be sediment (.5-.25 mm) or fine sediment (<.25 mm). Sediment has visible separated grains whereas fine sediment is not grainy to the touch and has no visible grains. The weights were averaged among the three samples at each site, which resulted in the overall river sediment composition average. Bank sediments were collected with the same method as the river sediment. The sediments were taken 50 cm from the current water level. The bank sediments were grouped holistically using color, texture and general soil composition as the criteria.

The bank angle was measured using a clinometer at water level. These measurements were taken at the three sites along the transect (0 m, 12.5 m and 25 m) on both sides of the river. The angles were then averaged for each respective side of the river. If a bank measured greater than 45°, the bank height was recorded using a transect measuring tape. The measurement was taken from the water level to the height at which the riverbank angle tapered to less than 45°.

The maximum and minimum widths of the creeks were measured using a transect tape stretched across the widest and narrowest parts of the observation site respectively. Maximum depth of the creeks at each of the four sites was also measured using a transect tape. Maximum velocity was measured in the area of the river with the most visibly strong current. Maps were drawn of each site with current, fallen trees, erosion and amount pooled taken into account. To assess the quantity of submerged logs, a member of the research team walked perpendicular to the 12.5 m mark on the transect across the river while counting the amount of logs they tread upon. Foliage Projective Cover (FPC) was measured every 5 meters along a transect set perpendicular to the 12.5 m mark on the original transect. The percentages were then averaged for each site. Tree girth was measured using a tailor's tape along this same perpendicular transect for trees that were taller than 2 m. Again, the values were averaged to obtain a representative value.

Results

Maps of the four individual sites: Sheoak A (Fig. 2), Sheoak B (Fig. 3), Wetherby A (Fig. 4) and Wetherby B (Fig. 5) give a visual interpretation of the stream and bank morphology. Sheoak A and Sheoak B are dominated by pool regions which explains the slow currents through the transect site. Wetherby A and B show more signs of erosion than the sites at Sheoak. Wetherby B has many more fallen logs that are visible from the bank, whereas Wetherby A and Sheoak A have major parts of the banks covered in roots from riparian trees.

Prior to riparian fencing, the Wetherby sites were not fenced off to keep out cattle. Cattle were allowed to walk down and into Rifle creek causing bank erosion as well as decreased water quality with bovine waste. In January 2011, riparian fencing was placed along the Wetherby portions of Rifle Creek, which led to 100% cattle exclusion. However, the fencing does not keep feral animals such as pigs and cats out of the area. Furthermore, upstream from the Wetherby sites, Bushy creek joins with Rifle Creek which increases catchment input. The water level can rise up to four meters past the current water level during the wet season.

There is much dissimilarity across the river metamorphic data; however, differences in maximum velocity, maximum width and average bank angle factor heavily into the variations between Sheoak and Wetherby. The maximum velocities for Wetherby A and Wetherby B are 5 and 16 times faster than the flow rates in Sheoak, respectively. Similarly, the maximum widths of Wetherby A and B are larger by 2.2 and 1.5 times compared to the Sheoak sites. The minimum widths, however, differ from site to site with Wetherby A and Sheoak B sharing the largest minimum width. Wetherby B is an outlier in the average tree girth with a measurement nearly 30 cm larger than the next closest. This is, in part, due to several large gum trees located within the transect. Wetherby A and B both possess a bank with at least a 60° average bank angle whereas Sheoak B only contains one bank that is barely greater than 45°. The tallest average bank heights vary from site to site. Sheoak A had no bank heights because there were no banks with an angle of greater than 45°.

The Wetherby A measurements came from the far side of the creek, whereas Wetherby B and Sheoak B had taller banks on the trail side. No glaring disparities appear with the rest of the shared data, although attention should be drawn to the FPC in Sheoak. They register more foliage projection cover because of the undisturbed rainforest ecosystem through which Rifle Creek travels.

River sediment analysis shows a stark contrast between the dominant sediments in Wetherby compared to the sediments in Sheoak. Both Wetherby sites have over 65% sand and gravel sediment whereas the Sheoak sites have less than 30% sand and gravel compositions. Individually, Wetherby A has nearly an equal amount of sand and gravel, whereas Wetherby B contained 9% more sand than gravel. Wetherby A also has sediment contrasting the fine sediment that exists in Wetherby B. Moreover, sediment and fine sediment dominate the creek composites at the Sheoak sites with 70% at Sheoak A and Sheoak B respectively. Sheoak A has only sediment with a nearly equal mixture of sand and gravel, whereas Sheoak B contains fine sediment with more gravel than sand. Organic matter only appears in the Wetherby A sediments. Unaccounted sediment refers to the sediment lost during the sifting process. This percentage (by weight) was lost in the form of water and sediment sticking in the sieves. These percentages, however, do not drastically alter the percentages gathered.

The analysis of riverbank sediments at each site shows three dominant soil types: sedimentary, sand, and clay. Wetherby B has the one exception with a bank full of a unique condensation of sedimentary soil into stalactite type shapes. For the most part, bank sediments remain consistent with the river sediments. Wetherby A and B have dark brown sand. However, the sand is darker and much more condensed than the sand sampled in the actual river bed. Sheoak B is all chocolate-brown clay-type soil that is homogenous and looks very similar to the fine sediment found in the creek bed. Sheoak A contains light tan soil with fine sedimentary particles that clumps together. This site also contains condensed sandy soil that resembles the Wetherby sites except a lighter color.

All previously collected results ultimately factor into the amount of platypus sightings. No sightings were recorded at Sheoak B throughout the observation period. There were also no sightings at any site in the evenings of October 14th or 15th. However, Wetherby A had the most individual sightings of platypuses with two, three and four each of the respective morning observations totaling 9 sightings. Wetherby A had a 100% sighting rate in the morning observation periods. Wetherby B ranked second with 7 sightings, all during the morning observation periods as well. Sheoak A had 4 total sightings, but was the only site to record any evening observations. The average time at the surface varied from site to site and from day to day. However, Wetherby B had much shorter average times when compared to the averages of Wetherby A and Sheoak A.

Discussion

This study suggests a possible connection between platypus sightings and a number of factors determining habitat structure. Stream sediment analysis showed a large difference between the compositions of the streambeds at Wetherby and Sheoak (Fig. 6). Several studies have been conducted correlating platypus presence to stream sediments.[8,9] However, these studies contain an evident contradiction. Some studies found that fine stream sediments, such as silt, encourage platypus inhabitance, whereas other studies strongly suggest that stream sediments composed mainly of gravel and larger particles attract platypus populations.[6,8,9,10] The data collected in this study supports the latter research. The two Wetherby sites had much more gravel and sand, whereas the Sheoak sites were dominated by sediment and fine sediment. It is possible that platypuses prefer the larger substrate for several reasons. The presence of macroinvertebrates in one substrate more than the other could suggest why platypuses were spotted more at Wetherby.

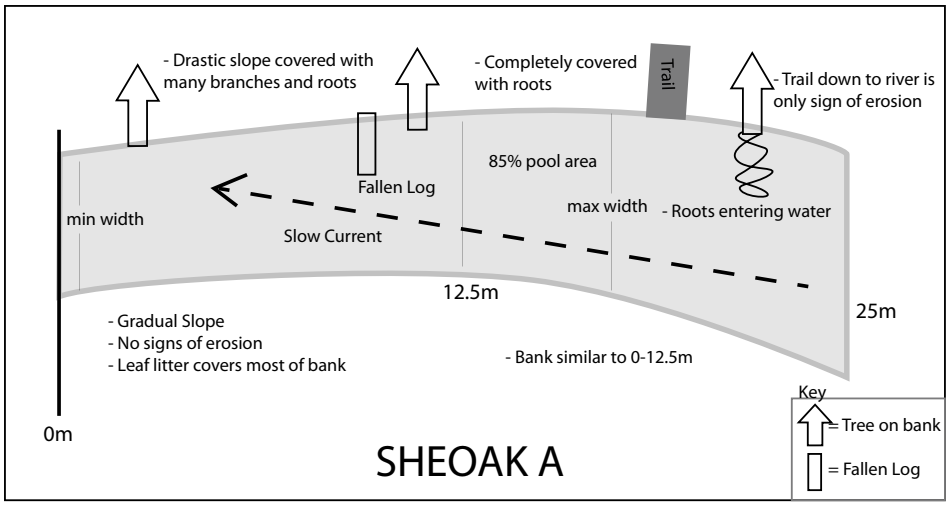


Fig. 2 Sheoak A

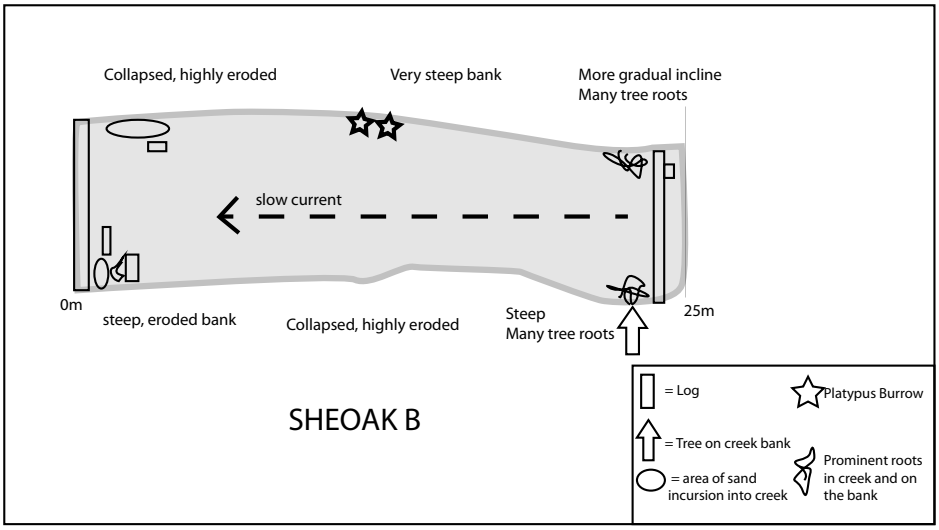


Fig. 3 Sheoak B

Coarse substrates such as gravel and sand provide more surface area for invertebrates to inhabit as well as a more stable substrate on which to attach.[6,9] However, a concurrent study focused on invertebrate diversity and abundance indicated that both invertebrate abundance and diversity was higher in the Sheoak sites (Guerra, unpublished data). Another study done in New South Wales supports the current findings that there is greater invertebrate abundance in the finer sediment areas of platypus habitats.[10] Milione *et al.* 2009 suggests that invertebrate diversity is not as important as abundance due to the adaptable diet of the platypus.[8] The many contradictions with concern to correlating platypus presence to invertebrate abundance leads to a possible conclusion: the platypuses at Wetherby inhabited that specific area for reasons other than invertebrate abundance and diversity. The greater presence of invertebrates at Wetherby further suggests that as long as there is sufficient food, factors such as average flow rate, river depth and bank morphometrics factor into habitat preference.

The differences between Wetherby and Sheoak concerning average stream velocity suggest that this is a physical factor that platypuses detect when inhabiting and foraging in certain sites (Table 1). Platypuses tend to avoid fast flowing streams and sites with large riffle areas due to the large energy expenditure required to hunt in such areas.[8,10] However, stagnation and stratification decrease the concentrations of dissolved oxygen and increase nutrient concentrations which further increase the risk of toxic algal blooms' negative impacts on aquatic invertebrate and fish communities.[7] Rifle Creek at Wetherby had a stronger flow rate due to the input of Bushy Creek upstream of the observation sites. More sightings at the Wetherby show that the riffle areas in the Wetherby sites were not too strong to completely discourage platypus foraging. Also, as seen in the site profiles, Wetherby has pooled areas which suggest that an area with enough current to create a more gravel/sand dominated substrate along with pooled area provides an ideal foraging habitat for platypuses. The flow rate at Sheoak B was slow enough to show evidence of stagnancy (Table 1). Platypuses were not seen at all at here, possibly due to the stagnancy of the pool which had a film that formed over the area.

Maximum stream depth did not vary much across the four sites. However, the Wetherby sites did have deeper maximum depths than those of Sheoak. 98% of platypus dives were in waters less than three meters, with the average dive being 1.28 meters.[6] All sites have maximum depths less than three meters and greater than 1 meter. The larger maximum depths at Wetherby could suggest that platypuses prefer pool depths of greater than 1.5 meters.

Stream bank morphometrics are important to the platypus because this is where burrows are made. Platypuses use burrows not only to rest in between foraging activities, but are also used for mating and egg laying.[11] Stream bank morphometrics present more differences in terms of bank angle and bank height between Wetherby and Sheoak. The Wetherby sites have at least one bank each with an average angle greater than 60 degrees. Serena *et al.*1998 showed that a positive correlation existed between platypus burrows and vertical banks.[4] Although the Wetherby banks are not vertical, they have an angle closer to 90° than the Sheoak sites. Platypus presence is evidence of a burrow as close as 100 meters because 50% of platypus activity is spent in a 100 meter range of the nearest burrow.[1]

No burrows were observed at any of the sites besides Sheoak B, but platypus burrows are often underwater or hidden in root systems.[11] Sheoak B was the only site with no platypus observations. Yet, it was the only site with visible platypus burrows, indicating that bank height and soil composition can be seen as an example of where the platypus actually dig burrows. The sites with platypus observations recorded bank heights of 1.85 meters or greater. However, there is no evident pattern concerning stream bank sediments

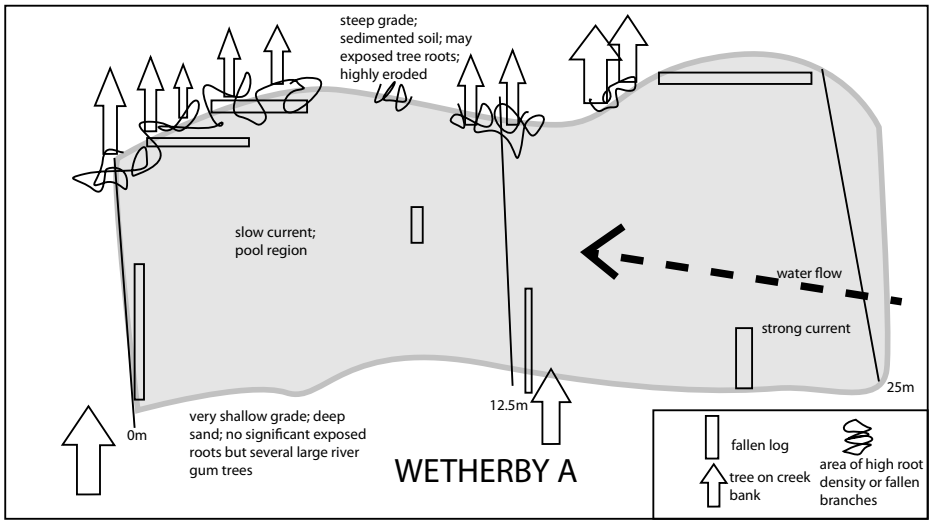


Fig. 4 Wetherby A

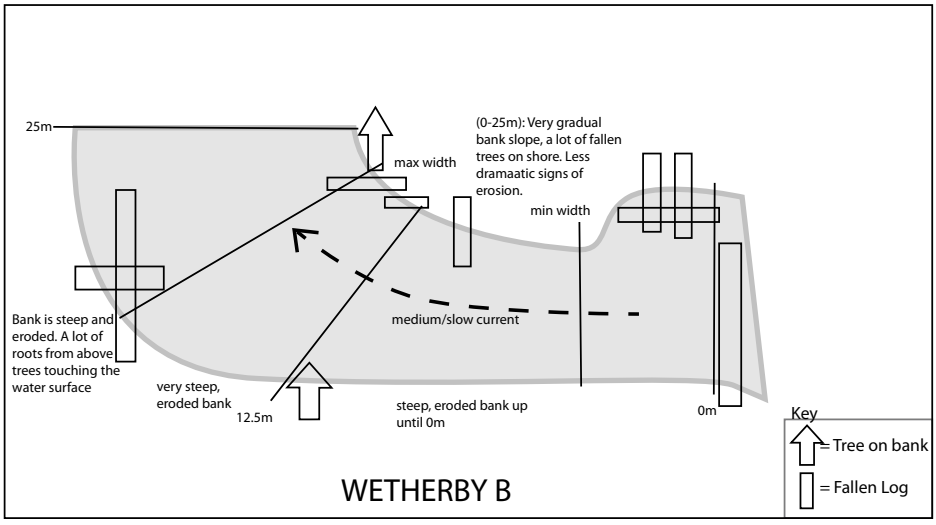


Fig. 5 Wetherby B

	Wetherby A	Wetherby B	Sheoak A	Sheoak B
Max Depth (m)	2.7	1.5	1.4	1.15
Max Width (m)	24.5	17.2	10.9	10.9
Min Width (m)	9.5	4.8	5.6	9.5
Avg. Velocity (m/s)	.2	.64	.04	.04
# Submerged Logs	5	10	8	7
FPC %	50.5	65	74.7	74.5
Avg. Tree Girth at 12.5 m (cm)	34	74.7	45.8	46
Avg. Bank Angle: Trail side	36.3°	60°	39.6°	47.6°
Avg. Bank Angle: Far Side	63.6°	35.6°	28°	42.3°

Table 1 Bank and creek characteristics across all four sites.

which could mean that burrowing merely requires a certain angle and height as opposed to a certain sediment composition.

A very similar study to this research was conducted in October of 2010 using the same sites at both Wetherby Station and Sheoak Ridge. However, several differences between the data from a year ago and these data exist. While the study last year noted a 0% sighting at either of the Wetherby sites, there was 100% AM sightings at Wetherby A and 50% AM sightings at Wetherby B in this year's study (Carey *et al.* unpublished data). These differences could be attributed to the riparian fencing constructed in January 2011 along the Wetherby sites which keeps cattle out of the portion of Rifle Creek that was studied. A year ago, the creek was observed to be heavily impacted by cattle mainly with concern to water quality and heavy bank erosion. The unrestricted grazing of cattle in stream areas can result in loss of riparian vegetation, erosion and catchment filling.[6] Riparian vegetation is important due to its two-fold effect on both burrows and water quality.[7] The riparian fencing has already shown a direct impact within eight months of its construction. Eliminating disturbance allowed for platypuses to move back into the Wetherby sites. There is no way of speculating whether these are the same platypuses observed in the Sheoak sites last year, but the presence of more platypuses in the Wetherby sites suggests that it is a more physically suitable environment for the reasons earlier discussed.

The contradictions in trends between platypus presence and macroinvertebrate abundance suggest that the platypus is more perceptive of physical differences within its habitat. These differences include an ideal velocity, bank angles close to vertical, creek depths between 1 and 3 meters and bank heights around 1.85 meters or greater. Furthermore, riparian fencing installed along Rifle Creek located in the Wetherby sites in January 2011 highlights the differences between this year's study and the study done prior to the fencing. More platypus sightings at Wetherby and the noted changes in stream physiology suggest that Wetherby is a preferred platypus foraging ground.

The protection of platypuses are important because not only are they a unique cultural marker of Australia, they play an integral role in stream ecosystems. Given the platypus' ability to adapt and its resilience, it can be a good indicator of positive trends in habitat conservation and rehabilitation because it will be one of the first species to move back into a previously destroyed habitat. Furthermore, the success of the site rehabilitation at Wetherby is a good example of successful conservation efforts and can be used as encouragement for other land owners interested in conservation. The more knowledge gained about platypus habitat, the better the conservationist can manage and prepare for threats to the platypus instead of just reacting to the issues as they come. The conclusions of this paper offer a possible solution to the river sediment contention. However, it would be helpful for a study to do a direct comparison between sites, one site with its substrate majority being sediment and another site with a coarse substrate majority. Also, a study devised to ascertain the minimum amount of food necessary for platypus foraging to occur could be useful for reintroduction efforts. Creating a yearly study at the Sheoak and Wetherby sites would be an invaluable asset to the topic of platypus research because such a longitudinal study does not exist. Certainly, continued environmental research into the platypuses' habitat and preferences is essential to ensure ongoing health of this fascinating and iconic species. Although just a small piece of the puzzle, this research brings light upon how a platypus chooses its environment to inhabit. The more knowledge gained about the platypus, the better able ecologists will be to conserve one of the world's most unique mammals.

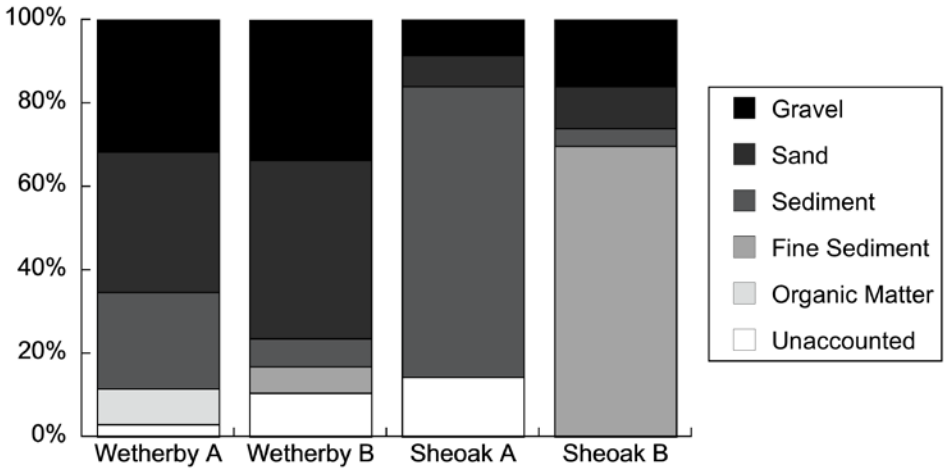


Fig. 6 A side by side comparison of all sites to illustrate differences in river sediment composition. Three samples were taken at each site along the transect in the middle of the stream. Compositions were averaged to give a comprehensive view of overall sediment composition.

Acknowledgements
Special thanks to: Mr. John and Kathy Colless, Dr. Claire Baker, Kristen Malinak, Ana Guerra, Tess Morgridge, and Mr. Bill Anderegg.

References

1. Gust N, Handasyde K. Seasonal Ranging Behaviour of the Platypus (O. anatinus) on the Goulburn River, Victoria. Australian Journal of Zoology,1995; 43:193-208.
2. Gardner JL and Serena M. Spatial organisation and movement patterns of adult male platypus, Ornithorhynchus anatinus (Monotremata: Ornithorhynchidae). Australian Journal of Zoology, 1995; 43(1): 91-103.
3. Grant TR, Temple-Smith PD. Field biology of the platypus (O. anatinus): historical and current perspectives. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 1998; 353(1372): 1081-1091.
4. Serena M, Thomas JL et al. Use of stream and river habitats by the platypus, Ornithorhynchus anatinus, in an urban fringe environment. Aust J Zool, 1998; 46:267-282.
5. Manger P, Pettigrew J. Electroreception and the Feeding Behaviour of Platypus (Ornithorhynchus anatinus: Monotremata: Mammalia).Philosophical Transactions: Biological Sciences, 1995; 347:359-381.
6. Grant TR, Temple-Smith PD. Conservation of the Platypus, Ornithorhynchus anatinus : Threats and Challenges, Aquatic Ecosystem Health & Management, 2003; 6(1), 5-18.
7. Klamt M, Thompson R et al. Early response of the platypus to climate warming. Global Change Biology, 2011; 17(10): 3011-3018.
8. Milione M, Harding E. Habitat use by platypus (O. anatinus) in a modified Australian Wet Tropics catchment, north-eastern Queensland. Australian Mammalogy, 2009; 31:35-46.
9. Serena M, Worley M et al. Effect of food availability and habitat on the distribution of platypus (O. anatinus) foraging activity. Aust J ZoolW, 2001; 49(3): 263-277.
10. McLachlan-Troup TA, Dickman CR et al. Diet and dietary selectivity of the platypus in relation to season, sex and macroinvertebrate assemblages. Journal of Zoology, 2009; 280:237-246.
11. Hawkins M, Battaglia A. Breeding behaviour of the platypus (O. anatinus) in captivity. Aust J Zool, 2009; 57:283-293.

		14.X.2011 PM	15.X.2011 AM	15.X.2011 PM	16.X.2011 PM	16.X.2011 PM	17.X.2011 AM
Wetherby A	# of times surfaced	0	2	0	3	0	4
	Average Time at surface (s)	NA	14	NA	25	NA	12.5
Wetherby B	# of times surfaced	0	1	0	0	0	6
	Average Time at surface (s)	NA	8	NA	NA	NA	4
Sheoak A	# of times surfaced	0	0	0	1	3	0
	Average Time at surface (s)	NA	NA	NA	24	12.3	NA
Sheoak B	# of times surfaced	0	0	0	0	0	0
	Average Time at surface (s)	NA	NA	NA	NA	NA	NA

Table 2 Platypus sightings by number of time surfaced. Times at surface averaged to obtain value in table. No sightings were recorded at the Sheoak B site. No sightings recorded at any site on October 14th or 15th in the PM observations.

Weston LaGrandeur is currently a Junior majoring in Biology and minoring in History at Stanford. Weston enjoys playing rugby for Stanford, working at Stanford Sierra Camp in the summers and doing anything in the outdoors. Once he graduates, Weston plans on taking a year off to travel and work as an EMT before he goes to Medical School where he hopes to become involved in Wilderness Search and Rescue.

An efficient $FeSO_4$-mediated synthesis of methyl-4-(ethoxymethyl)-benzoate and basic conformational analysis of the same using computational tools

Soumendranath Bhakat

The synthesis of Methyl-4-(ethoxymethyl)-benzoate from Methyl 4-(Bromomethyl) benzoate and ethyl alcohol using $FeSO_4$ as a recoverable and reusable mediator has been described without use of base and cosolvent under mild conditions and conformational analysis of the synthesized compound has been performed by computational tools using ArgusLab software .

Methyl-4-(ethoxymethyl)-benzoate is a substituted benzyl ether. Benzyl and substituted benzyl ethers are the most versatile protecting groups, useful in the protection of hydroxy groups in multi-step synthesis of complex natural products, especially in synthesis using carbohydrate as chiral templates.[1,2] Further, Methyl-4-(ethoxymethyl)-benzoate is used pharmaceutically as a drug with carboxylic acid derivatives having thiazole rings (mentioned in the patent (WO/2007/126043)). The synthesis of benzyl alkyl ethers include Williamson ether synthesis under basic or acidic conditions using trichloroacetimidates and benzyloxypyridinium triflates with alcohol in the presence of additives.[3,4] The literature evinces that only a few reports have been published for the synthesis of substituted benzyl ethers like Methyl-4-(ethoxymethyl)-benzoate without base and acid additives.[5] Specifically,in the patent (WO/2007/126043) Methyl-4-(ethoxymethyl)-benzoate was prepared using methyl 4-(hydroxymethyl) benzoate, iodoethane, N, N dimethylformamide and potassium tert-butoxide. The reaction was carried out in inert atmosphere such as a nitrogen balloon because potassium tert-butoxide is easily destroyed by atmospheric moisture. This synthesis is very difficult to carry out because the use of environmentally hazardous compounds (e.g., potassium tert-butoxide) and use of co-solvents (N, N dimethylformamide). Therefore, a more effective synthesis for substituted benzyl aromatic compounds is needed. Herein, it is demonstrated that the synthesis of Methyl-4-(ethoxymethyl)-benzoate is possible via use of environment friendly method using reusable $FeSO_4$ and use of co-solvent using the work of Joshi and Adimurthy.[6] This synthetic method avoids the above complications such as use of co-solvents, and environmental interaction with reagents, and provides a more manageable synthetic approach to the synthesis of substituted benzyl aromatic compounds such as Methyl-4-(ethoxymethyl)-benzoate that can be used in future syntheses.

Materials and Methods

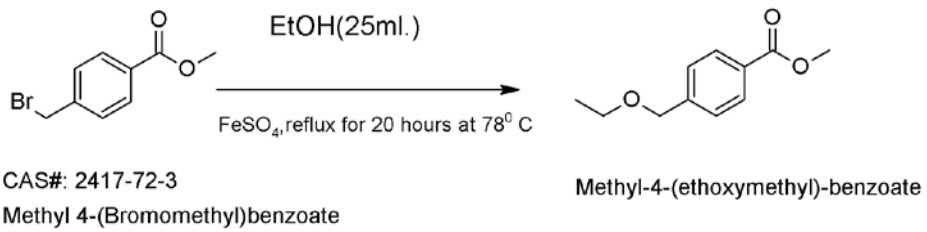
Synthesis of Methyl-4-(ethoxymethyl)-benzoate

A mixture of Methyl 4-(Bromomethyl) benzoate (CAS#: 2417-72-3) (0.458 g, 2.0 mmol) and $FeSO_4$, $7H_2O$ (2.0 mmol, 0.556 g) were refluxed (magnetic stirrer was used) in ethanol (25.0 mL) at 78 °C for a period of 20 hours. The progress of the reaction was monitored by thin-layer chromatography (TLC). TLC was monitored by TLC Silica Gel 60 F_{254} under UV light. After the complete disappearance of benzyl bromide, the reaction mixture was cooled to room temperature (25 °C). $FeSO_4$ was removed by filtration and the residue was washed with 10 mL of cold ethanol.

The characteristic red color crude product was dried out by RotaVap and Vacuum pump.

The ethanol was stripped out, and the remaining residue was purified by column chromatography on silica gel (100–200 mesh) to obtain the pure product as colorless liquid (as previously mentioned in the patent (WO/2007/126043)) with 86% yield.

This method of synthesis of the above mentioned compound is not only very efficient but is considered to be a "green synthesis" as the $FeSO_4$ used in the reaction can be recovered without any change from the reaction condition and further can be used in further reactions like in the synthesis of (Methoxymethyl) benzene from benzyl bromide and methanol, (Ethoxymethyl)benzene from benzyl bromide and ethanol.[5]



Synthesis scheme for Methyl-4-(ethoxymethyl)-benzoate

Experimental

^{1}H spectra were recorded on a spectrometer operating at 500 MHz in $CDCl_3$ unless otherwise stated. Column chromatography for purification was carried out on silica gel (100–200 mesh). Analytical thin-layer chromatography (TLC) was performed on an Aluchrosep silica gel 60/F_{254} plate under UV light. Mass spectroscopy of the sample was recorded on WATERS-Q-Tof Premier-HAB213 micro electrospray ionization (ESI) positive mode.

Results and Computational Study

Spectral data of the compound Methyl-4-(ethoxymethyl)-benzoate

^{1}H NMR 500 MHz($CDCl_3$): d(ppm) 1.235 (1, 3H, t, J=6.997), 3.815 (2, 3H), 7.331 (4, 1H, ddd, J=8.498, J=3.530, J=1.341), 7.324 (5, 1H, ddd, J=8.490, J=3.132, J=1.341), 7.924 (6, 1H, ddd, J=8.495, J=3.132, J=0.000), 7.924 (7, 1H, ddd, J=8.490, J=3.530, J=0.000),3.447(8, 2H, q, J=6.997),4.548(9, 2H);MS(ESI)($C_{11}H_{14}O_9$) =predicted[M+H$^+$]=195.1015,observed=195.1029,predicted[M+Na$^+$]=217.08341,observed =217.0851

Computational Method

A computational conformational analysis and geometry optimization study of Methyl-4-(ethoxymethyl)-benzoate was performed on a window based computer using Argus lab and ACD Lab Chem Sketch software.[7,8] The Methyl-4 (ethoxymethyl)-benzoate was used to determine the 3D structure of the molecule (Fig. 6). The Methyl-4-(ethoxymethyl)-benzoate structure was generated by ArgusLab 4.0.1 and geometric optimization was performed with the semi-empirical RHF/Austin Model 1 (AM1) parameterization.[9] The minimum potential energy was calculated by using geometry convergence function in Argus lab software. In order to determine the allowed conformation the contact distance between the atoms in adjacent residues was examined using criteria for minimum Vander Waal contact distance (Simons et al., 1983).[10] Surfaces created to visualize ground state properties as well as excited state properties such as orbital, electron densities, electrostatic potentials (ESP), and spin densities were used to generate the grid data used to make molecular orbital surfaces and visualize the molecular orbital (Fig. 3,4), making an electro static potential mapped and electron density surface (Fig. 5) map. The minimum potential energy was calculated for drug receptor interaction through the geometry convergence map. The final geometrical energy and SCF energy was found to be -59092.0491 kcal/mol as calculated by RHF/AM1 method, as performed by ArgusLab 4.0.1 suite.

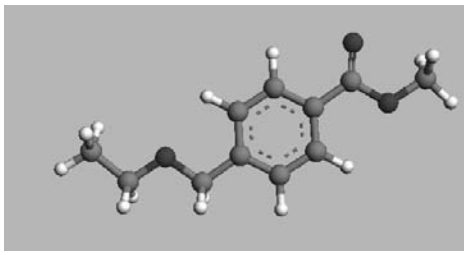


Fig. 1 Perspective view and active conformation of Methyl-4-(ethoxymethyl)-benzoate as optimized by ArgusLab 4.0.1 software

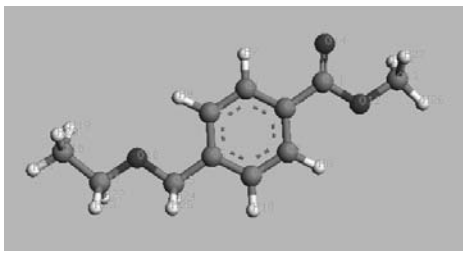


Fig. 2 Methyl-4-(ethoxymethyl)-benzoate molecule with labelled atoms

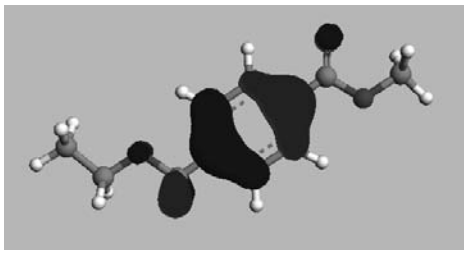


Fig. 3 HOMO (Highest Occupied Molecular Orbitals) of Methyl-4-(ethoxymethyl)-benzoate as visualised by ArgusLab 4.0.1 , blue shows positive and red shows negative charged molecular orbitals

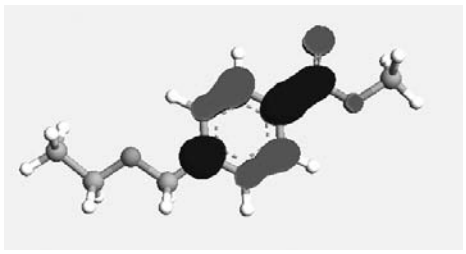


Fig. 4 LUMO (Lowest Unoccupied Molecular Orbitals) of Methyl-4-(ethoxymethyl)-benzoate as visualised by ArgusLab 4.0.1 ,blue shows positive and red shows negative charged molecular orbitals

Discussion

The present work demonstrates the synthesis Methyl-4-(ethoxymethyl)-benzoate under mild, efficient, environmentally friendly conditions mediated by $FeSO_4$ without any dry solvents and in an inert atmosphere. The most energetically favorable conformation of Methyl-4-(ethoxymethyl)-benzoate is found to have a heat of formation of -110.9143 kcal/mol via use of the Argus Lab software. Conformational analysis of bond angles, bond distances, mulliken charges, ZDO charges with minimum potential energy is crucial when establishing SAR/QSAR models using theoretically calculated descriptors, since the structure can depend on the molecular composition. Additionally, all geometric variables were completely optimized for the compound and the lowest energy conformation was used in molecular modelling studies. Current research is going on the preparation of a hydrazine derivative from Methyl-4-(ethoxymethyl)-benzoate. These hydrazine derivatives are predicted to have certain anti-cancer properties, as do hydrazine derivatives such as procarbazine analogues, methyl hydrazine derivatives, 2,6-Dimethyl-imidazo[2,1-b][1,3,4] thiadiazole-5-carboxylic acid (2-hydroxy-benzylidene)-hydrazide and others.[11]. This research is being presently conducted and its outcomes will be reported in the near future.

Acknowledgement

First, I would like to acknowledge Dr. Pratik Sen and his group members: Mr. Rajeev Yadav, Mr. Shahnawaz Rafiq, Ms. Bhaswati Sengupta, and Mr. Shradhey Gupta of the Department of Chemistry, IIT, Kanpur for their tremendous help and inspiration to make my project into reality. A special thanks goes to Dr. Somnath Bhakat and all my friends of Birla Institute of Technology, Mesra for their constant encouragement and pushing during my lows. Last but not the least, a special heartiest thanks to the Department of Chemistry of IIT, Kanpur for superb lab facilities and to all my teachers for keeping faith in me.

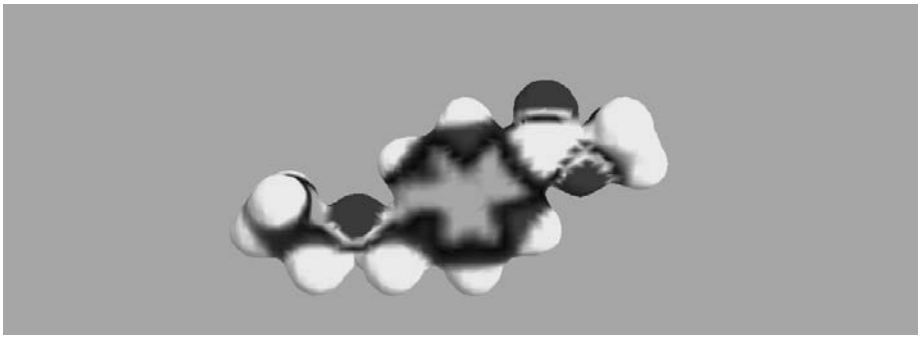


Fig. 5 The complete surface with the color map of ESP of Methyl-4-(ethoxymethyl)-benzoate

Table 1 List of Mulliken Charges and ZDO charges of Methyl-4-(ethoxymethyl)-benzoate as calculated by ArgusLab suite

Numbering of atoms as shown in the figure	Atoms	Mulliken Charges	ZDO charges
1	C	-0.1321	-0.0726
2	C	-0.1333	-0.1149
3	C	-0.1255	-0.0688
4	C	-0.1794	-0.1206
5	C	-0.0653	-0.0433
6	C	-0.2111	-0.1524
7	H	0.2267	0.1572
8	H	0.2224	0.1538
9	H	0.2305	0.1603
10	H	0.1991	0.1356
11	C	0.3972	0.3442
12	O	-0.3284	-0.2810
13	C	-0.1905	-0.0601
14	O	-0.3863	-0.3552
15	C	-0.0653	0.0125
16	O	-0.3171	-0.2824
17	C	-0.1034	-0.0184
18	C	-0.3532	-0.2157
19	H	0.1384	0.0894
20	H	0.1373	0.0884
21	H	0.1271	0.0813
22	H	0.1144	0.0661
23	H	0.1142	0.0660
24	H	0.1311	0.0807
25	H	0.1312	0.0806
26	H	0.1503	0.0992
27	H	0.1343	0.0840
28	H	0.1371	0.0863

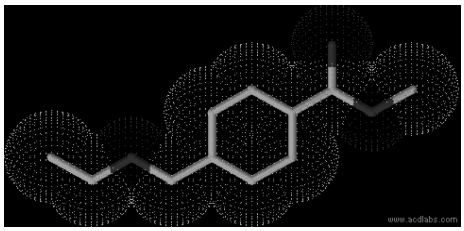


Fig. 6 3D view of Methyl-4-(ethoxymethyl)-benzoate visualized by ACD/3D viewer

References

1. Plante OJ, Palmacci ER, Seeberger PH. Automated solid-phase synthesis of oligosaccharides. Science, 2001; 291: 1523-1527.
2. Greene TW, Wuts PGM. Protective Groups in Organic Synthesis; John Wiley and Sons: New York, 1999.
3. Eckenberg P, Groth U, Huhn T et al. A useful application of benzyl trichloroacetimidate for the benzylation of alcohols. Tetrahedron, 1993; 49: 1619- 1624.
4. Poon KWC, House SE, Dudley GB. A bench-stable organic salt for the benzylation of alcohols. Synlett, 2005; 3142–3144.
5. Paul S, Gupta M. Zinc-catalyzed Williamson ether synthesis in the absence of base. Tetrahedron Lett, 2004; 45, 8825–8829.
6. Joshi G, Adimurthy S. New Method for the Synthesis of Benzyl Alkyl Ethers Mediated By $FeSO_4$, Synthetic Communications, 2011; 41: 720–728.
7. ArgusLab 4.0.1, Mark A. Thompson, Planaria Software LLC, Seattle, WA, http://www.arguslab.com
8. Downloaded from http://www.acdlabs.com/recources/freeware/chemsketch
9. Dewar MJS, Zoobisch EG, Healy EFet al. AM1: A new general purpose quantum mechanical molecular model. J. Am. Chem, 1985; Soc. 107: 3902-3910.
10. Simons J, Jorgensen P, Taylor H et al. Walking on potential energy surfaces. J. Phys. Chem, 1983; 87: 2745-2753.

Table 2 Ground State Dipole (Debye) of Methyl-4-(ethoxymethyl)-benzoate

X	Y	Z	length
-1.23106153	-3.33897775	-0.17334334	3.56291073

Table 3 Bond Angles between different atoms of Methyl-4-(ethoxymethyl)-benzoate

C18-C17-O16	106.94°
C17-O16-C15	111.85°
O16-C15-C5	109.52°
C2-C11-O14	125.06°
C2-C11-O12	114.64°
O14-C11-O12	117.30°
C11-O12-C13	116.43°

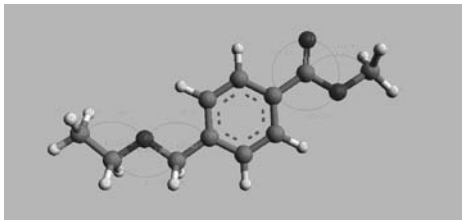


Fig. 7 Bond angles (in degree) in optimized geometry of Methyl-4-(ethoxymethyl)-benzoate as calculated by ArgusLab suite

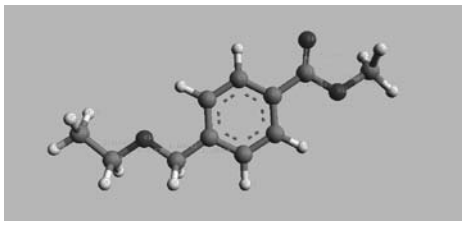


Fig. 8 Bond distance (in angstrom) in optimized geometry of Methyl-4-(ethoxymethyl)-benzoate as calculated by ArgusLab suite

Table 4 Bond distances (in angstrom) in Methyl-4-(ethoxymethyl)-benzoate

C18-C17	1.513
C17-O16	1.426
O16-C15	1.424
C15-C5	1.494
O14-C11	1.233
C11-C2	1.469
C11-O12	1.371
O12-C13	1.426

Soumendranath Bhakat is currently pursuing a B.Pharm course from Birla Institute of Technology, Mesra, India. His main interests are synthetic chemistry, medicinal chemistry, organic reaction mechanisms, and computational chemistry. He has published seven papers in various journals, and many of his papers have been accepted for presentation in different international chemistry conferences. He has also written articles for magazines related to chemistry. He describes himself as a crazy cricket player and a passionate novel reader.

Redemptive existentialism and Berkeleian metaphysics: a synthesis in Beckett's plays

Hannah Kim

Samuel Beckett's *Waiting for Godot* and *Krapp's Last Tape* are traditionally associated with Sartrean existentialism in that they deny any inherent purpose to life. However, a careful reading of the two plays shows that they need not be read as despair-ridden and pessimistic. Aware of the potentially devastating implications of Sartre's philosophy, Beckett offers a way for humans to find essence by highlighting George Berkeley's idealism, in which nothing exists without being perceived. Through the repeated motif of perception, Beckett's plays include in Berkeley's ideology, which holds that being recognized is at the heart of meaningful existence. The result is a Beckettian synthesis between existentialism and idealism wherein humans, thrust into a world with no essence, construct significance through perceiving and being perceived. *Waiting for Godot* and *Krapp's Last Tape* suggest that though meaning may not be inherent in the world, our sense of self and happiness is irreducibly tied to the way we are acknowledged, and that we are empowered only when we are accepted by ourselves and others. In the postmodern world where ambition, technology, and chaos often leave one to retreat into self-imposed loneliness, Beckett's dramas convey that introspection and interdependence are at the crux of purposeful life.

In *Waiting for Godot* and *Krapp's Last Tape*, both written in the 1950s, Samuel Beckett examines the emptiness of human existence, especially acute during the post-war era. His plays, which express the banality of life through seemingly meaningless repetitions, have been associated with existentialism in that they deny any inherent purpose to life. However, a careful reading of *Waiting for Godot* and *Krapp's Last Tape* reveals that Beckett's plays need not be interpreted as despair-ridden and pessimistic. A dialectical reading of the plays illustrates that Beckett offers a way for humans to find essence in life by showing his audience what *not* to do; in other words, by repeatedly exposing why his characters Vladimir, Estragon, and Krapp are unhappy, Beckett highlights their weakness and in turn suggests what one *should* do in order to overcome life's barrenness. Beckett's solution to existential despair derives from Berkeley's idealism in which nothing exists without being perceived. Beckett applies this ideology to the human psyche and dramatically conveys that although meaning may not be inherent in the world, humans can find essence in our relationships and interdependence.

After the Second World War, Samuel Beckett played a large role in shaping the literary movement that would later become known as postmodernism. Emerging from the horrors of genocide and destruction, many writers struggled to make sense of the atrocities they had witnessed and were further disappointed by the new Cold War. The apparent lack of progress in history inspired postmodernists to express the bleak human prospect in their works. In particular, playwrights reflected the transitory and incoherent nature of human existence in the Theatre of the Absurd, in which "men and women, as Shakespeare says, were viewed as mere actors in an absurd play, making their entrances and exists upon the stage of life and mouthing their tales 'full of sound and fury,' signifying nothing."[1]

Among the most influential postmodern plays are Beckett's *Waiting for Godot* and *Krapp's Last Tape,* both closely associated with absurdism and existentialism. Noting the barren yet symbolic details of his works, the philosopher Alain Badiou describes Beckett as "a writer of the absurd, of despair, of empty skies, of incommunicability and of eternal solitude − in sum, an existentialist."[2] Above all, Sartre's philosophy − which argues there is no intrinsic meaning in human existence − permeates *Waiting for Godot* and *Krapp's Last Tape.* Considered the father of existentialism and a contemporary of Beckett, Sartre held that existence is prior to essence and that humans come into being without inherent significance. Existentialism thus "places the entire responsibility for [one's] existence squarely upon [one's] own shoulders."[3] Though Sartre himself did not necessarily intend his philosophy to be pessimistic, Beckett showed that the seeming meaninglessness of life could lead individuals into despair. The seemingly purposeless lives of Vladimir, Estragon, and Krapp epitomize the Sartrean individual who struggles to make sense of life's seeming insignificance.

The Sartrean existential struggle is apparent even from the first line of *Waiting for Godot*. The bare setting of "A country road. A tree. Evening" immediately conveys emptiness; roads and trees, which are conventional symbols of life, are barren, suggesting a lack of abundance. The use of participle − the continuous and unfinished form of the verb − in the title of the play is also significant because it suggests that nothing worthwhile happens in life as Vladimir and Estragon idly wait for Godot to appear. Furthermore, the play opens with Estragon exclaiming that there is "nothing to be done," in response to which Vladimir confesses, "I'm beginning to come round to that opinion."[4] There really is nothing to be done in a play in which two plain characters wait, passing the time with verbal and physical repertoire. There is no coherent story, point, or design in the play, and Vladimir and Estragon's lives are reduced to meaningless repetition and banters. In fact, the two characters even dismiss existence as a sort of a problem to be solved:

> Vladimir: Suppose we repented.
> Estragon: Repented what?
> Vladimir: Oh… We wouldn't have to go to into the details.
> Estragon: Our being born?[5]

Being born becomes an occasion to repent for, as if existing is a sin, a fault, a regrettable fact. Vladimir also directly represents Sartre's notion of the universal individual, the idea that one man represents all of mankind because "in choosing for himself he chooses for all men."[6] For example, Vladimir convinces Estragon that they should help Lucky by claiming: "at this place, at this moment of time, all mankind is us, whether we like it or not. Let us make the most of it, before it is too late!"[7] Because one always chooses what one believes to be the best, each decision a

person makes serves to define the ideal for all of humanity. In this way, not only does *Waiting for Godot* capture the boredom and insignificance of an existential life, but it also relays Sartre's sentiment that the choices individuals make necessarily affect everyone else.

Similarly, *Krapp's Last Tape* harkens to Sartrean existentialism with its mindless repetitions and lack of a traditional plot. It is interesting to note that the play is set in the future; though the play takes place in late evening in the future (suggesting the end of time), Krapp possesses no intelligible sense of his past and thus fails to attribute meaning to his life. For instance, Krapp does not remember his journal's contents when he rummages through them, even when they involve something as grave as his mother's death: "mother at rest at last…. Hm…. The black ball…. [*He raises his head, stares blankly front. Puzzled*.] Black ball?"[8] Krapp repeatedly raises his head and stares blankly into space in hopes of remembering, but his life, too, has been reduced to fragments of random and redundant moments that have lost significance. Krapp's life lacks any sense of progress since what he thought was most important in his past years − discoveries, thoughts, and philosophy − are no longer valuable to him. In fact, they torment him to the point where he cannot even stand to listen to them again. When the younger Krapp in the tape shares what sounds like an epiphany and reads, "what I suddenly saw then was this, that the belief I had been going on all my life, namely—," Krapp angrily switches off the tape and curses as he fast-forwards through the next few minutes of the tape. He comments, "[I've] just been listening to that stupid bastard I took myself for 30 years ago, hard to believe I was ever as bad as that."[9] Instead of leading a fulfilling life, Krapp has spent his time on earth in a meaningless acquisition of years, exemplifying Sartre's philosophy that life is inherently meaningless.

Despite Sartre's belief that existentialism is empowering because it allows one to freely determine one's own essence, Beckett recognized how dangerous an inherently meaningless life could be for the individual; after all, Vladimir, Estragon, and Krapp are all unhappy. As such, Beckett employs strong currents of Berkeley's idealism in *Waiting for Godot* and *Krapp's Last Tape* to suggest a possible solution to the seemingly hopeless existential life. Building upon Sartre's notion of the universal individual, Beckett's plays endorse Berkeley's idea that one cannot exist happily on one's own because being recognized is at the heart of meaningful existence. Berkeley, an eighteenth-century idealist, famously argued that "*esse* is *precipi"*: that to be is to be perceived.[10] He posited that everything in the realm of awareness − form, colors, texture, taste, etc. − are ideas perceived by the senses and concluded that everything in the world, including individuals, exists only when acknowledged by a perceiving mind. When Berkeley wrote that "all the choir of heaven and furniture of the earth have not any subsistence without a mind" and that "their being is to be perceived or known," he contended that nothing truly exists until perceived by a conscious mind.[11]

When analyzed from a Berkeleian point-of-view, *Waiting for Godot* and *Krapp's Last Tape* show how life becomes reduced to a series of boring, meaningless, and fragmented moments when no one is really "looking." The plays reveal "a universe of insignificance, its tension created by the conflict between the insignificance and man's effort to give himself meaning despite everything."[12] However, difficulty arises when a man tries to give *himself* meaning; though Beckett may believe in the individual ability to determine essence, he suggests that we cannot conceive meaning on our own because we are relational beings.

One commentator agrees when he shares that "Beckett may be wishing to apply to mind Berkeley's notion of the relativity and dependency of the sensible object."[13] Vladimir, Estragon, and Krapp are all occupied with perception: they crave recognition and become depressed when they are deprived of it − this explains why Vladimir and Estragon endlessly wait for Godot and why Krapp only seems to recall people's eyes. Through the repeated motif of perception, Beckett provides a Berkeleian answer to the potentially devastating bleakness of existentialism. The result is a Beckettian redemptive existentialism in which our sense of self and happiness are constructed only when we are perceived, both by ourselves and others.

Evidence of Vladimir and Estragon's tireless desire for recognition pervades *Waiting for Godot*. Vladimir repeatedly wakes up his sleeping friend until Estragon exclaims, "Why will you never let me sleep?" Vladimir simply answers, "I felt lonely,"[14] as if he cannot stand not being looked at even for a minute because he is afraid he will cease to exist when no one is aware of him. Even Pozzo, a master and therefore arguably the most powerful character in the play, repeatedly asks for attention when he utters, "is everybody ready? Is everybody looking at me? Will you look at me, pig!"[15] However, the correlation between being seen and existing becomes most clear when Vladimir converses with the Boy upon learning that Godot won't be coming that evening:

> Boy: What am I to say to Mr. Godot, sir?
>
> Vladimir: Tell him… [*he hesitates*]… tell him you saw us. [*Pause*] You did see us, didn't you?[16]

Here, Vladimir and Estragon seek recognition from Godot before Godot arrives, and this exemplifies Beckett's synthesis of existentialism – the philosophy that existence precedes essence – and idealism – that to be is to be perceived. A similar dialogue takes place when Vladimir learns once again that Godot won't be coming in Act II:

> Boy: What am I to tell Mr. Godot, sir?
>
> Vladimir: Tell him… [*he hesitates*]... tell him you saw me and that… [*he hesitates*]… that you saw me. […] [*With sudden violence*] You're sure you saw me, you won't come and tell me tomorrow that you never saw me![17]

There is a strong need for Vladimir to know that he had been seen; Berman explains that "[in Beckett's plays] minds become dependent, vulnerable beings; they need the support and comfort of being perceived."[18] Beckett's characters express their basic desire for acceptance through their obsession with being seen.

The Berkeleian "to be is to be perceived" axiom is also present in *Krapp's Last Tape* as Krapp ends his reclusive life in deep unhappiness. During the play, we learn that Krapp had been distancing himself from others because of his artistic ambitions. "As the tapes make clear," Gordon analyzes, "he has pursued the life of the mind, separating the 'grain' from the 'husks' following his 'vision' that would, so he thought, survive his infirmity."[19] The extent to which Krapp preferred solitude is clear when he records that he "celebrated the awful occasion [of his birthday], as in recent years, quietly at

the Wine-house" with "not a soul" around him.[20] Instead of choosing to take a break once a year to celebrate his birthday, Krapp took even that time to pursue the life of a writer. However, this commitment to achievement only leaves him restless, and he repeatedly confesses his loneliness: "[I] never knew such silence. The earth might be uninhabited."[21] We learn that even his career has been unfulfilling, as only "seventeen copies [of his book] sold, of which eleven [were sold] at trade price to free circulating libraries beyond the seas."[22] Though he invested himself in the life of a detached artist, he returns again and again on the tape to the moments he wants to relive, and to his disappointment, Krapp learns that beauty lies in moments shared and not in the achievement of brilliance. In accordance with Berkeley's philosophy, Krapp's greatest moments were those concerned with recognition.

Krapp, like Vladimir and Estragon, strongly desires to be seen, as evidenced through his obsession with eyes. His insatiable appetite for recognition, coupled with his misery due to his failure to connect with others, suggests that a meaningful existence cannot exist without others' perception. Krapp is particularly obsessed with the female gaze. For instance, though he calls his relationship with Bianca a "hopeless business," he fondly recalls her eyes as "very warm" and "incomparable"; he remembers "not much about her, apart from a tribute to her eyes."[23] Similarly, Krapp is drawn to a particular nurse because she happened to be looking at him: "whenever I looked in her direction she had her eyes on me… […] The face she had! The eyes!"[24] Krapp seems incapable of remembering people unless he feels acknowledged by them. The crucial boat scene to which Krapp returns time and time again confirms how important the female gaze is for Krapp's sense of self:

> "I said again I thought it was hopeless and no good going on and she agreed, without opening her eyes. [*Pause*.] I asked her to look at me and after a few moments − [*Pause*.] after a few moments she did, but the eyes just slits, because of the glare. I bent over her to get them in the shadow and they opened. [*Pause. Low*.] Let me in"[25]

Here, the last sentence − "Let me in" − though unmistakably sexual in suggestion, also functions as Krapp's metaphysical plea: the girl needs to help him belong to this world. In his analysis, Knowlson acknowledges that the girl's eyes not only serve as windows to the soul, but also as mirrors, reflecting and confirming that which is before them.[26] This is why Krapp cannot stand having the girl's eyes in the sunlight and creates a shadow so that she can look at him properly: he needs to know that he is being perceived thoroughly to feel grounded. The present-day Krapp repeatedly returns to this part of the tape because this instance of companionship, gaze, and intimacy was when he felt most alive.

Lastly, Beckett shows the importance of self-perception for meaningful existence through Krapp's inability to address his present self. In Beckett's adaptation of idealism, self-recognition is just as indispensable as others' recognition in shaping one's existence. Krapp is always recording to his future self or listening to his past self; he does not, or cannot, ever face his present self. To complicate matters further, the play is set in the future: Krapp "lives entirely outside time in a no-man's land," a place of "a repeated past and anticipated future."[27] The tape recorder limits Krapp to his past or future self, and his present self is preoccupied with everything but the now. "The words that Krapp had recorded so many years ago," Knowlson observes, "now represent the only form of contact that he can achieve in a depleted, solitary, almost totally barren existence."[28] If time is a relative term and reality is constructed of perceiving minds and ideas, Krapp never truly exists because he is never in the present.

Beckett uses the Berkeleian notion of perception in *Waiting for Godot* and *Krapp's Last Tape* to offer a solution to the problems that existentialism creates. The two plays echo Sartre's philosophy in their suggestion that there is no necessary or inherent purpose to life. However, Beckett provides a way out from existential despair in Berkeley's idealism, the philosophy that nothing in the world is made significant until chanced upon a by perceiving mind. Beckett's philosophy precludes passive and reclusive lives from being meaningful, and in this framework, Vladimir, Estragon, and Krapp's lives serve as the antithesis of the ideal lifestyle we should be pursuing. In short, Beckett's characters show what we should *not* do if we desire a purposeful life.

Ultimately, in the postmodern world where technology and ambition promote self-reliant individualism more than ever, Beckett warns that a meaningful existence can seldom be achieved alone. Despite life's seemingly inconsequential repetitions and routines, there is essence to be found, and, according to Beckett's drama, self-conscious trueness to the present self and the willingness to be perceived by others are at the crux of purposeful life. One must perceive oneself and one another in order to find meaning in this world. We cannot thrive on our own. Reconciling with the past, taking advantage of the present, and forming relationships with others ultimately point us to Beckettian redemptive existentialism − the philosophy that existence can be empowered when we are perceived, and that only when we are accepted do we find meaning and significance.

References
1. Fletcher J. About Beckett: The Playwright and the Work. London: Faber and Faber, 2003: 20.
2. Badiou A. On Beckett. Ed. Nina Power and Alberto Toscano. Manchester: Clinaman, 2003: xvii-38.
3. Sartre JP. "Existentialism Is a Humanism." Club Maintenant, Paris. October 29 1945.
4. Beckett S. Waiting for Godot. In: Samuel Beckett: The Complete Dramatic Works. London: Faber and Faber, 1955: 11.
5. Ibid p. 74.
6. Sartre JP. "Existentialism Is a Humanism." Club Maintenant, Paris. October 29 1945.
7. Beckett S. Waiting for Godot. In: Samuel Beckett: The Complete Dramatic Works. London: Faber and Faber, 1955: 74.
8. Beckett S. Krapp's Last Tape. In: Samuel Beckett: The Complete Dramatic Works. London: Faber and Faber, 1955: 217.
9. Ibid. p.222
10. Berkeley G. A Treatise Concerning the Principles of Human Knowledge. Ed. Jonathan Dancy. Oxford: Oxford University, 1998: PHK§3.
11. Ibid, PHK§6.
12. Guicharnaud J. Existence on Stage. In: Harold Bloom, ed. Modern Critical Views: Samuel Beckett. New York: Chelsea House, 1985; 116.
13. Berman D. Beckett and Berkeley. Irish University Review Spring 14.1 1984: 5.
14. Beckett S. Waiting for Godot. In: Samuel Beckett: The Complete Dramatic Works. London: Faber and Faber, 1955: 17.
15. Ibid, p. 30.
16. Ibid, p. 50.
17. Ibid, p. 47.
18. Berman D. Beckett and Berkeley. Irish University Review Spring 14.1 1984: 5.
19. Gordon L. Krapp's Last Tape: A New Reading. Journal of Dramatic Theory and Criticism Spring 1990; 100.
20. Beckett S. Krapp's Last Tape. In: Samuel Beckett: The Complete Dramatic Works. London: Faber and Faber, 1955: 217.
21. Ibid, p. 223.
22. Ibid, p. 222.
23. Ibid, p. 218.
24. Ibid, p. 219.
25. Ibid, p. 221.
26. Knowlson J. Krapp's Last Tape: The Evolution of a Play, 1958-75. In: Cathleen Culotta Andonian ed. The Critical Response to Samuel Beckett. Connecticut: Greenwood, 1998; 155.
27. Gordon L. Krapp's Last Tape: A New Reading. Journal of Dramatic Theory and Criticism Spring 1990; 104.
28. Knowlson J. Krapp's Last Tape: The Evolution of a Play, 1958-75. In: Cathleen Culotta Andonian ed. The Critical Response to Samuel Beckett. Connecticut: Greenwood, 1998; 150.

Hannah Kim is a junior at Emory University majoring in Philosophy and English. Outside of the classroom, she actively contributes to the campus music scene as a drummer/percussionist. Her academic interests include Hegel, theology, philosophy of time, and Metaphysics (after all, how could anyone do philosophy without questioning the nature of reality?). She hopes to one day help bridge the gap between metaphysics and theoretical physics. This paper was written while studying abroad in London, where she plans to return for a year before embarking on her postgraduate studies.

Nature or nurture? Perspectives of Japanese manga on an age-old debate

Sean Wang

For centuries, the question of whether heredity or the environment plays a more significant role in determining individual identity has been hotly contested.[1] Studies suggest that 50 percent of who we are psychologically is genetic, but in the world of manga, that proportion may be higher.[2] This analysis will examine undertones of the "nature versus nurture" debate in three of the most widely circulated manga series – Dragon Ball, Naruto, and The Prince of Tennis – by highlighting connections between each series and its implied perspective on personal growth and development, and drawing upon direct quotations and key plot points as evidence of genetic and external factors.[3] Finally, the results are applied to the context of manga as a developmental force and identify potential implications on the next generation of global readers.

Manga: a pop culture pandemic

Manga, the Japanese phrase for "humorous pictures," typically consists of highly stylized comics or print cartoons that tell dramatic tales. In Japan, the concept of illustrating these character-centered narratives dates back to as early as the 12th century; today, manga in Japan represents a thriving $3.6 billion industry.[4]

No longer isolated to East Asia, the popularity of this art form has spread throughout the world, reaching virtually all races and age demographics. Despite only gaining attention in North America starting in the 1970s, manga publications already command an impressive market, generating roughly $175 million of annual sales in the United States and Canada alone.[5] With a continuously growing online readership, it is no surprise that more business leaders and politicians are recognizing manga's expanding role in economics, entertainment, and international affairs. Even Japanese Prime Minister Taro Aso, a self-confessed "Otaku" (a Japanese term for a person obsessed with manga) and addict, geared his 2009 stimulus measures to exporting his nation's "pop culture influence" overseas, creating an estimated half a million jobs for his struggling economy.[6] Manga's potential impact on its readership, particularly young, developing children who comprise the largest audience segment, should not be underestimated.[6]

***Dragon Ball*: power through inheritance**

Originally introduced in 1984 as part of a weekly magazine, *Dragon Ball* follows the adventures and battles of Goku, a martial arts enthusiast born from a legendary alien race, the "Saiyans." In the manga's multi-planetary universe, Goku continuously faces challenges from elite and malevolent fighters threatening to destroy Earth, forcing him to constantly seek greater power in order to protect the people around him. Each chapter of the storyline progresses in this similar manner: the protagonist encounters allies to assist him on his journey against even stronger enemies whom he must defeat to defend the planet.[7]

Even though he is assisted in his conflicts by a team of incredibly competent "Z warriors," it is eventually Goku, and not one of his comrades, who fights and wins the majority of the 187 battles that take place in the series.[8] In fact, the superhuman character is responsible for saving the world from almost every major *Dragon Ball* villain, triumphing over evil when his teammates fail. This impels the question: why, out of all of Earth's fighters, can only Goku carry the fate of the world on his shoulders?

In determining the origin of Goku's strength, there are two primary candidates: the hero's external training environment and his internal "Saiyan blood." In support for the former, the most commonly cited example is the protagonist's combat preparation in the "100G" Gravity Machine, conditions which nearly triple his recorded power level and openly demonstrate a case of experience-triggered improvement.[9] Yet contrary to this pro-nurture argument, according to the manga, none of the other characters utilize this training device despite the availability of the machine and each of their desires to become stronger fighters. This brings up the likelihood that out of the Z warriors, the alien Goku is the only one genetically capable of enduring a hundred Earth gravities, a force far exceeding the limits of normal human beings. The nature-favoring perspective is further boosted by multiple references to the Saiyans as an "invincible warrior race," thus explaining his invulnerability on the battlefield through a likely hereditary transcendence of mankind.[10] Moreover, Goku is quoted to possess the innate advantage of a "Saiyan zenkai," defined by the manga as a genetic trait that allows his power to increase substantially after recovering from near fatal injuries.[7] Together, these observations validate that individual development in the series is based more on descent than environmental influences.

The instances in which Goku is defeated also provide convincing evidence for heredity as the ultimate source of fighting prowess. The most notable instance of this is the protagonist's duel against the computer-generated experiment of Cell, a genetically enhanced warrior created partially from collections of Goku's DNA:

> "Long ago, he began gathering the cells of fighting experts, and combined those cells. His artificial human research began, but since it was taking too much time, he abandoned it mid-way. However, his computer continued the work without rest. Son Goku['s]… cells were harvested at the battle"[11]

The entire character of Cell rests on the assumption that power originates from strong genetic material; this proposes that in *Dragon Ball*, individual combat skill is determined primarily by inherited nature. Following this thought, it makes logical sense that among the remaining Z warriors, the only one with the potential to subdue Cell is Gohan, the son and closest genetic equivalent of Goku.[7] If heredity is the most critical factor in assessing overall strength, then the best fighters theoretically should be those with the most superior genes, regardless of extrinsic factors. In accordance with this hypothesis, it is not a coincidence that Goku, Gohan, and Cell, who share much of the same genetic code,

represent the top three "power levels" during the saga.[12] From their rankings, it is clear that training is not the most important attribute in determining an individual's fighting proficiency. Instead, an upper limit exists that can only be surpassed by heredity greatness.

Naruto: genius trumps hard work

One of the best selling manga of all time, *Naruto* narrates the story of a rambunctious adolescent ninja, Naruto Uzumaki, and his dream to become the village leader, or "Hokage." In a fantasy parallel world of historic Japan, the title character and his ninja academy classmates embark on various missions to promote the welfare of their village, often fighting bandits and rogue criminals who pose threats to public safety. Additionally, the young warriors compete amongst themselves while training, pushing each other to improve their techniques and surpass prior limits to become more capable on the battlefield.[13]

A central theme from the early *Naruto* chapters is the idea that "Hard work can beat genius," a phrase frequently uttered by one of the main supporting characters of the series, the workaholic Rock Lee.[14] When the academy students and their combat styles are first introduced in the series, it is apparent that nearly all of them were born with "bloodline limit jutsus," supernatural fighting abilities genetically passed from parent to offspring.[13] Unlike most of his classmates, however, Lee possesses none of these heritable powers. Instead, the well-nurtured warrior relies solely on physical training and tenacity to keep up with his "genius" peers:

> "You know what? Here is what I believe. There are two different kinds of ninja. Those like you who are born with a natural talent and do not have to work at it. And those like me! The ones that have to train every day of our lives!"[14]

Initially, Lee's assertion that dedication can triumph over natural talent holds true when he defeats an alleged genius, Sasuke Uchiha, in a brief hand-to-hand skirmish. At this point in the manga, the hardworking ninja is clearly much faster than his gifted opponent, whose body cannot even react to the agility of the attacks.[13] Yet only a month after their confrontation, Sasuke acquires the speed that Lee trained a lifetime to reach through his inherited "Sharingan" ability.[15] Grudgingly, Lee acknowledges the superiority of the genius.

The debate is sustained by the mysterious lineage of the main character, Naruto Uzumaki. Orphaned from infancy, Naruto grows up without knowing the identity of either parent. Subsequently, he develops the reputation of a delinquent child, favoring vandalism over his ninja schoolwork and barely graduating from the academy at the bottom of his class.[13] During this early portion of the storyline, the title character's external training is still minimal compared to that of the other students; in fact, he has yet to achieve the lowest ninja rank of a foot soldier, or "genin." Despite his relative lack of practice and dedication, Naruto demonstrates signs of a natural gift when he flawlessly executes an extremely difficult "A-rank" technique to rescue his injured instructor.[16] If we take into account the limitations of his formal ninja education, there appears to be no reasonable justification for Naruto's sudden talent, implying a pro-nature explanation. Confirming these predictions, it is later revealed in the series that Naruto's father was the former "Hokage," or village leader, while his mother was a legendary "Jinchuuriki," or ninja of unparalleled strength and stamina, thereby exposing two inherited sources of the character's latent potential.[17] Because of his blessed ancestry, the protagonist carries an innate fate to be the "child of prophecy" and in time lead the ninja world.

Naruto develops into a tireless and hardworking hero; however, the key to his supremacy over his equally or more diligent classmates remains his genetic disposition. After analyzing Naruto's background in addition to those of Lee and Sasuke, it is evident that a poor heredity accompanied by continuous environmental training can match or even temporarily surpass the levels of intrinsic genius. However, although both hard work and natural talent may appear necessary for success, the required latter component plays a substantially larger role in forming the ultimate individual. Therefore, based on the growth and maturation of the village's strongest ninjas, analysis concludes that in *Naruto*, the predetermined genetics of genius trump all outside influences on development.

The Prince of Tennis: gifted by birth

Considered perhaps the most popular sports manga, *The Prince of Tennis* details the life of Ryoma Echizen, a junior tennis prodigy attending Seigaku Academy in Japan. During his first year on the school's competitive tennis club, Ryoma engages in dozens of matches against older and more experienced opponents, each with their own signature shots and abilities that pose new challenges for the young athlete. With support from his coach, father, and extraordinarily skilled teammates, the gifted "Prince" eventually leads Seigaku through multiple tournaments all the way to a national championship, emerging victorious after kindling his unconditional love for the game.[18]

As the series title suggests, Ryoma is blessed with royal blood, giving him a genetic basis for success at the sport. According to the plotline, the protagonist is the son of Nanjiro Echizen, a former professional tennis player who remained undefeated for his entire career. Not surprisingly, many of Ryoma's winning techniques can be attributed to this hereditary advantage, such as his ambidexterity or "nitōryū," which he reportedly inherited from his father.[19] Thus, from the manga's very name to its frequent parallelisms of playing styles between father and son, there seems to be a clear emphasis on the theme of inborn talent.

While Ryoma improves significantly faster than most of the other tennis players, there is no tangible evidence that he practices any harder than they do. For example, even though all members of the Seigaku team undergo the same drills and training regiments, the Prince alone is able to maintain an official record of zero defeats, illustrating the lack of connection between training and playing ability.[17] In addition, his gradual advancement during matches from the lowest position of Singles 3 at the season's beginning to the top ranking of Singles 1 by the national competition confirms the idea that Ryoma progresses at the highest rate.[19,20] To further dispel any notions about exterior sources of tennis ability, the manga implicitly states that Sadaharu Inui, a Seigaku upperclassman whom Ryoma easily defeats, actually subjects himself to the most rigorous form of training, taking on a supplementary workload totaling to 2.25 times more than that of his teammates.[21] In relation to the nature versus nurture debate, the scenario presents a testable hypothesis for the influence of the environment on individual development. If in *The Prince of Tennis* external experiences are the main contributor to a player's tennis skill, then Sadaharu should have the

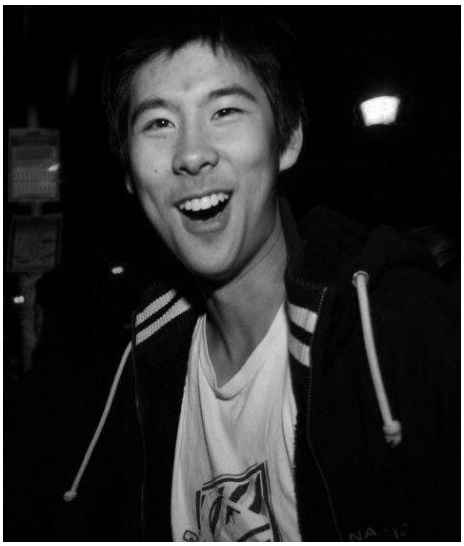
fastest improvement of any team member. However, as demonstrated by his imperfect match results and the exception of Ryoma, this is not a satisfactory explanation.

Of all the possible remaining candidates, innate and inherited factors appear the most likely, especially after considering explicit evidence from the manga's pro-nature naming conventions. According to the ending chapters of the series, Ryoma's final opponent is repeatedly addressed as the "Child of God," implying that truly mastering tennis requires some type of divine birth.[22] Even if these collectively perceived notions of genetic favoritism are all a coincidence, it is unquestionable that to its readers, *The Prince of Tennis* conveys an impression of the world in which sporting prowess is hereditarily predetermined. In agreement with this mindset, analysis deduces that the perspectives of this series and popular manga in general are strongly geared toward a reverence of intrinsic and natural talent.

Conclusions

After analyzing character development in three of the most widely read manga, a noticeable trend has emerged, one that extends well beyond the contexts of *Dragon Ball*, *Naruto*, and *The Prince of Tennis*.

What are the implications of all of these pro-nature perspectives on manga readers? For casual readers of *Weekly Shōnen Jump* or any other source manga publication, the impacts may be largely subconscious, ranging from unintended sensitivity concerning one's genetic makeup to self-limiting mentalities during childhood and adolescence. Particularly among younger audience segments, there is the possibility that manga restricts potential for growth by downplaying the importance of environmental factors such as education and routine practice in their daily lives, an effect additionally damaging for those perceived to be of humble birth. Although many of these anticipated side effects pose clear hazards to the work ethics of the younger readers, series authors cannot and should not be blamed for how their works are interpreted. Readers must instead learn to recognize that manga, along with all other forms of addictive yet fictional entertainment, are simply stories of human invention, fabricated for our enjoyment and not an accurate portrayal of reality.



References

1. Groff P, McRae L. "The Nature-Nurture Debate in Thirteenth-Century France." American Psychological Association, 1998.
2. Bouchard, Jr TJ, Lykken DT, McGue M et al. Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart. *Science* 250.4978 (1990): 223-28.
3. ComiPress. "Shounen Jump Manga Circulation Numbers." 1 June 2006. .
4. Wong WS. "Globalizing manga: From Japan to Hong Kong and beyond." Mechademia: an Annual Forum for Anime, Manga, and the Fan Arts, 2006. pp. 23–45.
5. Reid C. "2008 Graphic Novel Sales Up 5%; Manga Off 17%." Publishers Weekly. 6 Feb. 2009.
6. McCurry, J. "Japan looks to manga comics to rescue ailing economy." The Guardian. 10 April 2009.
7. Anime News Network. "Encyclopedia: Dragon Ball (manga)." 2011. Web.
8. Dragon Ball. "All Fights from Dragon Ball & DBZ." 2001.
9. Toriyama, A. "Son Goku's Spaceship (Son Gokū no Uchūsen)." Dragon Ball. Weekly Shōnen Jump. March 2003.
10. Toriyama A. "A Discovery (Buruma kara no Shirase)." Dragon Ball. Weekly Shōnen Jump. Dec. 2003.
11. Toriyama A. "Cell (Nazo no Kaibutsu, Tsui ni Shutsugen!!)." Dragon Ball. Weekly Shōnen Jump. Dec. 2003.
12. Planet Namek. "Dragon Ball Z Power Levels." Web. 22 Dec. 2011.
13. Viz Media. "Naruto Shippuden: Story." Web. 19 Dec. 2011.
14. Kishimoto M. "A Total Mismatch!!! (Saiaku no aishō...!!)." Naruto. Weekly Shōnen Jump. 7 Dec. 2004.
15. Kishimoto M. "The Great Flight!! (Ōi naru hishō!!)." Naruto. Weekly Shōnen Jump. 5 Dec. 2006.
16. Kishimoto M. "Uzumaki Naruto!." Naruto. Weekly Shōnen Jump. 6 Aug. 2003.
17. Kishimoto, Masashi. "The Conversation with Lord Fourth!! (Yondaime to no kaiwa!!)." Naruto. Weekly Shōnen Jump. 2 Feb. 2010.
18. Anime News Network. "Encyclopedia: The Prince of Tennis (manga)." 2011. Web. 22 Dec. 2011.
19. Konomi T. "A Suspicious Child Has Appeared (Kuse mono arawaru)." The Prince of Tennis. Weekly Shōnen Jump. 19 May 2004.
20. Konomi T. "Drive B (Doraibu B)." The Prince of Tennis. Weekly Shōnen Jump. 6 Sept. 2005.
21. Konomi T. "Endurance vs. Endurance (Seishin Ryoku Kakeru Seishin Ryoku)." The Prince of Tennis. Weekly Shōnen Jump. 5 Sept. 2006.
22. Konomi T. "Final Battle! The Prince vs. The Child of God (Fainarubatoru! Kami no ko tai Ōji)." The Prince of Tennis. Weekly Shōnen Jump. 5 April 2011.

Sean Wang is a freshman and prospective biology major from South Pasadena, CA who secretly enjoys reading manga when time allows. Although considering a future in medicine, he wishes to explore all the opportunities that Stanford has to offer and embody the spirit of entrepreneurship through organizations such as Social-E Capital and the Compass Fellowship.

Fichte the fascist? The misappropriation of a republican philosopher in Weimar, Germany 1918-1933

Michael Albada

Beginning in the late 19th century, the influential philosopher Johann Gottlieb Fichte came to be posthumously celebrated as a hero of the German nation. His renowned Addresses to the German Nation became a founding document of German nationalism that was frequently used to justify a wide range of political ideologies that differed from, and even contradicted, Fichte's own philosophy. This pattern reached its climax during the Weimar Republic, when figures representing the full political spectrum – from socialists to liberals to fascists – invoked the philosopher to justify their political agendas. A vocal minority of conservative writers trumpeted and misrepresented Fichte to challenge the newly-created democratic system, support their political agenda, and call for a fascist state. This paper offers a close study of the development of this anti-democratic portrayal of Fichte during the Weimar Republic, providing insight into the dangers of heroification, the intersection between philosophy and politics, and the profound malleability of nationalism as a political ideology.

Few cases provide such a fascinating glimpse into the mythology, grandeur, contradiction, and malleability of nationalism than that of German philosopher Johann Gottlieb Fichte, an individual who came to be used to justify a full range of political ideologies from socialism to fascism in the early twentieth century in Germany. Johann Gottlieb Fichte, who lived from 1762 to 1814, stood at the crossroads between cosmopolitanism and nationalism. The two coexisted within Fichte's thought, though certainly not without tension between them. An Idealist philosopher and disciple of Immanuel Kant, Fichte based his thinking on the Enlightenment-era universalistic and cosmopolitan conceptions of the eighteenth century. In his Characteristics of the Present Age, published in 1806, Fichte called on the enlightened mind to reject his government if it decayed and turn to "where light and justice are" in the world. This appeal represented a rejection of local patriotism and embraced a cosmopolitan ideal, which stands in sharp contrast to the countless nationalist invocations of Fichte that dominated the twentieth century.[1]

A staunch supporter of the French Revolution, Fichte fiercely criticized monarchy and the aristocracy, called for a republican government, and proposed a proto-socialist economic system. Fichte saw Napoleon's rise to power as a rejection of republican ideals and, in his legendary Addresses to the German Nation called on his countrymen to become the beacon of progress to the world, carry out a nationalist awakening, and resist Napoleon's armies. This perceived shift in Fichte's writing from cosmopolitanism to nationalism became one of the most contested points of his life and an endless source of dispute. Fichte's vision of nationalism came well before its time and he remained a largely obscure figure.

The rise of nationalist sentiment in the late nineteenth century brought Fichte to great prominence, venerating him as a prophet to the German nation and bringing him into the budding pantheon of national heroes. This celebration of Fichte as national hero continued through the late nineteenth and early twentieth centuries, placing particular focus on his Addresses to the German Nation, which writers mythologized as the pinnacle of patriotic fervor, moral conviction, and the standard for nationalist rhetoric.[2] This process intensified dramatically during the First World War, when writers invoked Fichte as a prophet of war and source of moral courage for soldiers and citizens alike.

After four years of devastating war and under tremendous strain, the German Empire crumbled. Cities broke into open revolt and the Kaiser and all ruling princes abdicated in November 1918. Meanwhile, a national assembly convened in the city of Weimar and announced the creation of a republic amidst ongoing revolts and tremendous social and economic pressures. These ongoing issues contributed to an unstable political situation, where no fewer than six major parties, and many smaller ones, vied for power. On the left side of the spectrum, the Social-Democratic Party focused on social welfare, the German Democratic Party represented intellectuals and small traders, and both parties supported the republic. The German People's Party and the Centre Party made up the middle and tended to be ambivalent about the republic and wanted a revision of the Treaty of Versailles. On the right side of the spectrum, the National German People's Party and the National Socialist German Workers' Party fiercely opposed the democratic system and supported either a return to monarchy or authoritarianism.[3] The Weimar Republic was marred by numerous problems including hyperinflation, political extremism, instability, and came to an end with the ascent of the Nazi Party and Adolf Hitler to power in 1933. Despite its flaws, however, the Weimar Republic's collapse was by no means inevitable.

During the time of the Weimar Republic, a vocal minority of conservative writers and political activists used Fichte to call for a powerful nationalist awakening, challenge the newly-created democratic system, and assert German power abroad. While this result was, in a certain regard, the culmination of the long development of national thought in Germany, this anti-democratic, nationalist movement overwhelmingly contradicted Fichte's own position on the state and the nation. Republican, democratic, and socialist writers also celebrated Fichte, but for very different reasons. Some invoked Fichte to support the new democratic government in Berlin, others to build support for the newly-created League of Nations, and still others to call for a more social-democratic structure of society. In this bizarre moment between the World Wars in Germany, individuals of varying ideologies invoked the same man to justify a range of contradictory opinions across the political spectrum in Weimar Germany.

One must ask why Fichte's work is so malleable to such a wide range of political ideologies, even to contradictory positions.

The answer appears to be twofold: first is what historian Rainer Lepsius called the "historical plasticity" of the nation, which allows a great deal of ideological freedom to those who choose to embrace nationalist rhetoric and symbolism, precisely because nationalism is not tied to any particular ideology.[4] Second is the presence of truly nationalist statements in Fichte's writing that were easy to remove from context and portray as a justification for rabid nationalism. Though by no means a militant nationalist, Fichte's passionate oratory style, wide range of ideas, dedication to the cause of the German nation, and fiery language made him malleable to a wide range of political movements that did not necessarily reflect his own convictions. Many other pieces of Fichte's writings thus wholly contradicted the positions of the monarchists, conservatives, and nationalists who invoked him. Through careful emphasis and omission, a wide range of authors invoked a single figure to justify almost every major political ideology in the Weimar Republic.

The present study serves a dual purpose: to analyze the way in which conservative writers distorted Fichte's life and work to support their own political agenda and to study the development of the anti-democratic, anti-socialist, proto-fascist portrayal of Fichte in the Weimar Republic. This will be achieved through a close reading of three of the most important writings by conservative, anti-democratic authors that focused on Fichte during this era. These documents come from a range of time within the interwar years, which makes it possible to trace the evolution of this particular portrayal of Fichte during the Weimar Republic.

Inspiration Amidst the Ashes

Amidst the crisis of the immediate postwar era in Germany, a so-called "Fichte Renaissance" began when concerned Germans looked back to their heroes to find stability, hope, and guidance for the future.[5] A sense of political crisis lingered over much of Europe after the war, particularly in Germany, where the legitimacy of the new government remained contested. In the chaotic aftermath that immediately followed the war, more editions of Fichte's Addresses were published, while a flood of pamphlets and newspaper articles demanded that Fichte's proposals be enacted to support a moral and national rebirth.[6]

The best example of writings from this era was undoubtedly Fichte und Deutschlands Not ("Fichte and Germany's Suffering"). Published under the name "From One Who Loves Germany" in 1919, the author's choice to publish anonymously reflected his political ideas and self-image as a humble servant of the German nation.[7] The author served with the Imperial German Army on the Eastern Front during the First World War, where he witnessed the collapse of the army in Riga, present-day Latvia, in 1918, an experience which clearly made a lasting impact on him. The author belonged to a young generation of German nationalists known as the Conservative Revolutionaries who experienced their formative years in the trenches of the First World War, felt humiliated by Germany's defeat, but saw opportunity in the end of the old Wilhelmine monarchy.[8] The challenge for the young generation, they believed, was to create a "new nationalism."[9] This political movement was characterized by an aggressive style, military organization, and strict hierarchy and found political unity in its vehement opposition to democracy, socialism, and communism. In many ways, the Conservative Revolutionaries can be considered predecessors of fascism.

The author of Fichte und Deutschlands Not focused almost entirely on Fichte's Addresses, elevating them to the status of a near-divine text for Germans to engulf themselves in. The thrust of the book's argument rests in its first two sentences:

> Fichte's Addresses to the German Nation deserve, now more than ever, the attention of all Germans. Not as a curiosity that attracts us for days or hours, but as a book of life, as our German Bible, in which we cannot deepen ourselves too often or too seriously.[10]

The author went so far to directly compare the Addresses to Christianity's sacred text, revealing the deep influence of religion on nationalist rhetoric and an active effort to elevate the nation to holiness. While the author of Fichte und Deutschlands Not conceded that the Addresses were well known, he insisted that their truest and deepest meaning of national rebirth had been lost on the masses. The task fell to Germans to discover the true meaning of the Addresses, and awaken within themselves the spirit of the nation. The author's vision of the nation as an immortal object of dedication and sacrifice which places tremendous burden upon the citizen stands in marked contrast to Fichte's own vision for the nation as a beacon of truth and a vehicle for the moral and spiritual enlightenment of the individual and of humankind. Fichte's firm commitment to republican ideals made him difficult to invoke to support the author's anti-democratic conviction. To square the circle, the author asserted that Fichte's democratic spirit was "strong" and "proud," as opposed to Weimar's "flabby" and "bourgeoisie" nature.[11]

In light of his vehement opposition to liberalism, democracy, and socialism, the author's calls to become "German in Fichte's meaning of the word" meant wholeheartedly embracing Fichte as a prophet of nationalism.[12] By seeking to establish the Addresses to the German Nation as a kind of holy book of the nation that would uplift and inspire Germans in their daily lives, the anonymous author sought to glorify Fichte as a timeless and fervent hero to the German nation and as a progenitor to the fascist movement.

The Völkisch Movement, Religion, and Fichte

The powerful and conservative Völkisch movement in Germany also came to embrace Johann Gottlieb Fichte in criticizing the Weimar Republic and advocating for a traditional, religious, and intensely conservative society in Germany. With a romantic focus on folklore and ethnicity, the Völkisch movement shared elements in common with the American Populist movement. Finding strength in the rural areas, the Völkisch movement combined utopianism with a reactionary spirit. As historian Petteri Pietkäinen observed, "Völkisch ideology was not a coherent set of ideas and ideals but rather a cauldron of beliefs, fears and hopes that found expression in various movements and were often articulated in an emotional tone."[13] Originally a nationalist reaction against industrialization, urbanization, and Progressivism that flourished in the 1870s, the movement gained strength in the interwar years. Supporters of Völkisch movement did not seek political means to address its concerns, but instead focused on the spiritual needs of the Volk.[14] As Nicholas Goodrick-Clarke argued, the Völkisch movement drew its strength from "the delayed political unification of Germany" and "a widespread romantic reaction to modernity."[15] While diverse, the

various elements of the movement shared a patriotic interest in German folklore, local history, and an anti-urban revolt against modernity. Fichte was embraced by the Völkisch movement, particularly so in the interwar years, as a source of inspiration and a justification for their vision of German society.

An influential Völkisch leader by the name of Helmuth Johnsen connected Fichte to the movement, using him to promote a pious, Protestant, and populist vision of German society. After serving in the First World War, Johnsen became an evangelical bishop in a rural region north of Nuremberg, a leader in the Völkisch movement, a popular activist. An early supporter of Adolf Hitler, Johnsen volunteered as the local coordinator for Hitler's march to Berlin, a dramatic political feat modeled on Mussolini's March on Rome that was supposed to follow the Beer Hall Putsch in Munich 1923. Johnsen saw Fichte as a powerful figure who could justify his religious and anti-Semitic worldview, using him widely and regularly in his speeches, sermons, and writings. Fichte became so central to his thought that Johnsen dedicated his doctoral thesis in philosophy to the philosopher.[16] While the thesis was not all that widely circulated, it provides a thorough account of Johnsen's portrayal and interpretation of Fichte that was widely disseminated within the Völkisch movement.

Part philosophy, part history, part biography, Johnsen's thesis follows the interplay between Fichte's personal and philosophical development, with specific focus on Fichte's role in the development of German thought and political philosophy. Published in 1929, Johnsen's portrayal focuses on more than Fichte's Addresses to the German Nation alone, but places strong emphasis on Fichte's patriotism throughout, downplaying Fichte's liberal and socialist lines of thought.

Johnsen framed Fichte as the heroic answer to the moral and existential shock that faced Germans in the wake of defeat and crisis that followed the war.[17] Out of the depths and darkness of despair, Johnsen, like many other writers during this time, offered Fichte as the best solution. More than Fichte's ideas, philosophy, or worldview, Johnsen saw the man himself as a larger-than-life hero who could inspire and uplift the struggling souls of the nation:

> He was in a period that is similar to ours in many ways, and has been a truly guiding personality – upright and brave before the enemy, true to himself in action and inaction, thorough in his knowledge and research, humble before God, well aware of the contrast between humans values and action. His shining example of a heroic life encompassed insightful scholarship and unceasing joy of work.[18]

That passage exemplifies the intentional, conscious effort to cultivate the image of Fichte as a national hero. By enshrining Fichte's personality over his thought and his actions over his ideas, Johnsen pulled Fichte away from his role as thinker and philosopher and made him a tool in the contemporary political battle. In order to cultivate a shared sense of identity and belonging to the nation, Johnsen sought to raise the greats of German history far above ordinary Germans. Heroes were to shine from the heavens as guiding stars for soul of the nation.

Johnsen embraced the Führerprinzip, the principle that a strong, driving leader should guide the nation. This thought originated in conservative and proto-fascist circles as part of the total rejection of democratic systems. Johnsen argued that Fichte should become the moral guide for the German nation in much the same way that a powerful leader should lead the state. He wrote, "May Fichte be the Führer and teacher of our Volk in the necessary renewal of its life- and world-view to ease the despair of our fatherland."[19] This passage provides clear insight into Johnsen's vision of reform in Germany and reveals his obsession with power. One's politics were not simply an opinion in Johnsen's eyes, but a "moral obligation" to embrace his ideals of humanity. Throughout his work, Johnsen exaggerated Fichte's Christian belief, manipulating his portrayal of Fichte in order to promote his own vision for a conservative, Christian, Völkisch Germany.

Fichte and National Socialism

Perhaps no individual has more wrongly misappropriated Fichte than Ernst Bergmann, who saw in him a direct progenitor of National Socialism. As professor of philosophy at the University of Leipzig, Bergmann developed a philosophical system which connected religion, race, and national awakening. After joining the Nazi Party in 1930, Bergmann earned the distinction of becoming the party's chief philosopher. Convinced that Jewish influence in Christianity was "destroying the German soul," he argued that Germans should create a national church based on popular mysticism and Nordic mythology.[20] Ernst Bergmann embraced Johann Gottlieb Fichte as a national hero and claimed him as a forerunner to National Socialism.

Bergmann published his first book on Fichte as part of the campaign designed to rally support for the war effort during World War I. Bergmann delivered a series of lectures on Fichte in 1932 which he went on to publish as Fichte und der Nationalsozialismus ("Fichte and National Socialism"). Primarily a work of propaganda, Fichte und der Nationalsozialismus epitomized the historical misappropriation and abuse of Fichte and his writings for political motives and portrayed Fichte as the forefather of National Socialism.

In Fichte und der Nationalsozialismus, Bergmann harkened back to past moments of crisis, glorified Fichte as a great leader of German history, and called upon Germans to rise up to the standard set by the glories of national heroes past. Like his contemporaries on the right wing of the political spectrum, Bergmann portrayed the crisis of Weimar society as a transformational moment and "turning point" in history which required Germans to look back to German heroes of the past.[21] According to Bergmann, the political and economic crises that engulfed German society after World War I demonstrated the corruption of liberal democracy, which could only be solved through the strong leadership of National Socialism. Invoking national heroes and describing previous times of trial for the nation added gravity and a seeming stoic legitimacy to Bergmann's blatantly political claims. Bergmann recounted and exaggerated Fichte's achievements, going so far as to say that the realization of Fichte's dream of a national education system helped to "resurrect the nation's soul."[22] As Bergmann argued, Fichte's greatness afforded him a "unique position" in German history:

> Should the claim of "native born," which National Socialism applied to the cultural manifestations of a people also be carried over to philosophy? One may and should try, but one will find hardly any

> other German thinker whose system so unambiguously bears the stamp of a German national philosophy on its face like that of our Fichte.[23]

According to Bergmann, both Fichte and National Socialism tapped into the "unfathomable depths of the German spirit," and thus shared the same purpose. Just as the National Socialist movement measured other aspects of culture and politics by the degree to which they were "native-born," Bergmann applied the same metric to philosophy, finding in Fichte a convenient figure to justify his own ideas and vision. Bergmann went even further in arguing that Fichte played a pivotal role in the development of National Socialism, writing

"Regarding Fichte as the first major precursor [to National Socialism] is not unjustified." [24]

Such a reading, however, ignores Fichte's other writings and aspirations. Fichte's version of nationalism was inherently tied to cosmopolitanism and support for a League of Nations more than a century before its creation. Fichte did not claim that the Germans were superior, but that they possessed a unique spiritual and historical position. Fichte's nation was not one intended for competition and combat with others, but rather the vehicle for moral, spiritual, and philosophical improvement of those within the nation. Fichte's cosmopolitan desire for the improvement of mankind began with a national desire for the improvement of the German nation. Bergmann engaged only briefly with Fichte's liberal and socialist thought, which largely contradicted his own beliefs, and argued that Fichte's nationalism and socialism came to a natural fruition in the National Socialist movement.[25] This substantial omission reveals Ernst Bergmann's attempt to claim Fichte as a forerunner to National Socialism as a concerted effort at cherry-picking passages from Fichte. In so doing, Bergmann created a caricature of a purely nationalist Fichte that was detached from the man and his thinking.

Bergmann believed Nazism and Marxism were engaged in a titanic battle to control the future. Bergmann saw Fichte as a central figure in this ideological battle, going far enough to call him "the moral conqueror of Marxism."[26] Bergmann missed the profound irony of his claim. Just as Fichte was venerated by nationalists of the right as a national hero and committed patriot, so too was he celebrated by socialists for fighting for the people and advancing socialist thought. Bergmann would have been shocked to learn that one of the greatest collections of Fichte's writings resided in the Marx-Engels Institute in Moscow, where, since before Bergmann's time, he has been celebrated as a socialist and revolutionary.[27] Seeking to strengthen support for the Nazi party and its ideological vision, Bergmann omitted those aspects of Fichte which did not conform to his preferred portrayal of the national hero.

Ernst Bergmann's relentless effort to portray Fichte as a forerunner to the National Socialism exemplified the misappropriation of Fichte's political and historical role to support a radically conservative agenda. Bergmann advanced a familiar, though deeply flawed, image of Fichte as an ardent patriot, early nationalist, and timeless hero of the nation. Bergmann added the novel and profound claims that Fichte was a direct forerunner to National Socialism and that Fichte was the moral conqueror of Marxism. Both of these claims followed in the line of thinking promoted by the previous two writers, but they represent the ultimate fruition of the effort to claim Fichte, a cosmopolitan and republican philosopher, as a forerunner to and hero of the most xenophobic, destructive, nationalist parties and regimes the world has ever known.

Visions of the National Hero

Though the details vary among the authors on the right, a shared image emerges from their writings that is vehemently nationalist, staunchly against the Weimar democratic system, fiercely anti-socialist, and that portrays Fichte as a forerunner to fascism. All three authors intentionally contributed to the mythos around Fichte as a national hero, and invoked him to support their own political visions. Just as fascism came to dominance, so too did the portrayal of Fichte as prophet of an extreme and aggressive version of nationalism. Through the curtailing of political and academic freedom, that portrayal became the only one.

A substantial debate emerged in the scholarship on the strength of the connection between Fichte and National Socialism that has not to this day been wholly resolved. On one side, scholars such as Timothy Ryback have argued that Fichte was in fact the philosopher closest to National Socialism in tone and spirit. In Hitler's Private Library, Ryback drew parallels between Fichte and Hitler in calling for an overthrow of the political elite, drumming up support for a people's war, and dreaming for the unity of the German people. [28] Some truth can certainly be found in this reading, and one can find powerful and haunting echoes of Fichte's philosophy and writings in the National Socialist movement. In addition, Fichte continued in a line of anti-Semitic German and, for that matter, European, thinking by arguing against Jewish emancipation in 1793 and claiming that Jews made up a state within a state that would undermine the German nation.[29] One must be careful not to give Fichte too kind a treatment, but on the other hand, one also risks conflating Fichte with the deeply flawed portrayal advanced by vehement nationalists in the late nineteenth and early twentieth centuries. Ryback's portrayal of Fichte as the philosophical forerunner to National Socialism is a treatment not dissimilar to that offered by Ernst Bergmann.

By using his name and cherry-picking from his writings, Nazi leaders bent Fichte to their image and, in so doing, gave authority to their own movement. In "Fichte's Contribution to German Nationalism," Thomas Zingales argued that Hitler's movement was in no way based on Fichte's philosophy. Hitler overlooked Fichte's cosmopolitan and transcendental spirit, making only a shallow attempt to legitimize the Nazi movement with the veneer of a philosophical system.[30]

Fichte fiercely criticized the Blutgemeinschaft idea that was so central to Hitler's ideas of race and nation, arguing instead that language formed a much stronger bond of national identity than blood ever could. Fichte also criticized nationalism of the land or soil, which he saw as the basest and most hollow form of patriotism.[31] Further, Fichte did not at any point consider National Socialism's "racial purity" in his work. Fichte did not consider it possible, necessary, nor even productive for German national unity.[32] As scholar F.W. Kaufmann argued in his article, "Fichte and National Socialism," even Fichte's most nationalistic writings were "not a prophecy of German racial preeminence, but a challenge to take the lead in responsible world-citizenship." [33] We must also remember that, according to Fichte, the nation was merely a vehicle

for spiritual and moral progress, one that may at some point lose its usefulness. This stands in sharp contrast to the dreams of fierce nationalists, for whom the realization of the nation was the end in and of itself. The spirit of Fichte's work, if not its tone, was antithetical to National Socialism.

Certain aspects of Fichte's writings were particularly susceptible to misappropriation. In addition to his nationalism and passionate rhetoric, Fichte's elevation of Germanness to a metaphysical essence and his nationally oriented vision of socialism appealed to these leaders. Less conservative than the Hegelians and more passionate than the Kantians, Fichte provided just the right combination of philosophical justification, populist appeal, and ideological malleability to make a compelling patron saint of National Socialism. Despite being susceptible to misappropriation, Fichte was not a forerunner to National Socialism; Nazi leaders deeply misrepresented him and his ideas to make him appear so.

Perhaps the clearest lesson from this study of Fichte is the profound malleability of nationalism as a political ideology.[34] Through careful emphasis and omission, writers actively manipulated portrayals of Fichte to justify an extremist and destructive ideology. Since nationalism is not bound to any particular doctrine or model of organizing politics, society, or the economy, it can be harnessed by a wide range of interests. Nationalism does not require sound logic, but only passionate intensity. Even the most dangerous ideology can gain credibility by bending national symbols, heroes, and rhetoric to its will. Movements that seek to legitimize themselves through the use of national rhetoric therefore deserve the deepest skepticism

References

1. Meinecke F. Cosmopolitanism and the National State. Kimber RB, translator. Princeton: Princeton University Press, 1970:73.
2. Reden an die deutsche Nation in the Original German.
3. Historical Exhibition of the German Bundestag. The Political Parties in the Weimar Republic. March 2006.
4. Lepsius MR, Campbell JA. The Nation and Nationalism in Germany. Social Research, 2004; 71:481-500.
5. Johnsen H. Das Staatsideal J.G. Fichtes. Erlangen: Friedrich-Alexander-Universität,1928:6.
6. Fichte JG. Addresses to the German Nation. Moore G, ed. New York: Cambridge University Press, 2008: xxxv.
7. "Von einem, der Deutschland lieb hat" in the original German.
8. Woods R. The Conservative Revolution in the Weimar Republic. New York: St. Martin's Press, 1996:2.
9. Ibid.
10. Anonymous. Fichte und Deutschlands Not. Berlin: Verlag von Martin Warneck, 1919:3.
11. Ibid.
12. Ibid., 6.
13. Pietikainen P. The Volk and its Unconscious. Journal of Contemporary History 2000; 35:524.
14. Ibid.
15. Goodrick-Clarke N. The Occult Roots of Nazism: Secret Aryan Cults and their Influence on Nazi Ideology. New York: New York University Press, 2004:3.
16. Johnsen H. Das Staatsideal J.G. Fichtes. Erlangen: Friedrich-Alexander-Universität,1928.
17. Ibid., 6.
18. Ibid., 5.
19. Ibid., 6.
20. Bergmann E. Nordisch-Germanischer Glaube oder Christentum. Leipzig 1934:6.
21. Bergmann E. Fichte und der Nationalsozialismus. Breslau: Walther Gehl und Georg Wolff, 1933:8.
22. Ibid., Vorwort.
23. Ibid., 6.
24. Ibid., 7.
25. Ibid.
26. Ibid., 8.
27. Engelbrecht HC. Johann Gottlieb Fichte: A Study of his Political Writings with special reference to his Nationalism. New York: Columbia University Press, 1933:190.
28. Ryback R. Hitler's Private Library. New York: Alfred A Knopf, 2008:107.
29. Dawidowicz LS. The War Against the Jews: 1933-1945. New York: Bantam Books, 1986: 34.
30. Zingales T. Fichte's Contribution to German Nationalism. Masters Thesis, Stanford University, 1955:38.
31. Meinecke, Cosmopolitanism and the National State, 73.
32. Engelbrecht, Fichte: A Study of his Political Writings, 153.
33. Kaufmann FW. Fichte and National Socialism. The American Political Science Review 1942; 36:470.
34. Sheehan JJ. Nation und Staat: Deutschland als "imaginierte Gemeinschaft." In: Hettling M, Nolte P, eds. Nation und Gesellschaft in Deutschland: Historische Essays. Munich: Verlag C.H. Beck, 1996.

Michael Albada graduated from Stanford in June 2011 with a degree in history. Originally from a small town north of San Francisco, Michael is currently serving as an AmeriCorps member at a charter school in East Palo Alto. Next year, he will be leaving California to teach high school math in Philadelphia and try to close the achievement gap with Teach for America. In his free time, Michael enjoys skiing, hiking, playing bass, and exploring the Bay Area.

Terrible lizards and the wrath of God: How 19th century Christianity and Romanticism affected visual representations of dinosaurs and our perceptions of the ancient world

Aaron Peterson

Contemporary popular understanding of prehistoric life contains a number of persistent falsehoods: namely, the association of chaos, violence, death, and intense geologic activity with the ancient world as a whole and dinosaurs in particular. These beliefs are deeply intertwined with a long historic tradition of artistic representation of dinosaurs, beginning with the paintings and mezzotint frontispieces of Romantic artist John Martin. While the first ten years of prehistoric reptile artwork exhibited a strong tendency towards a naturalistic, Enlightenment style of representation, Martin's experience as a painter of grandiose biblical and historical epics helped establish a clear break from that tradition towards one that emphasized the primordial chaos and violence of the antediluvian world. Equally important were the early fossilists and geologists who, in their attempt to reconcile their findings with biblical Scripture, gave artists like Martin the creative freedom to envision and create a world dramatically darker and less orderly than either contemporary or modern interpretations of nature. The remarkable power of visual representation in the public imagination has allowed these misconceptions to flourish for almost 200 years.

In the 1993 film *Jurassic Park,* chaos theoretician and mathematics professor Ian Malcolm gives an account of creation: "God creates dinosaurs. God destroys dinosaurs. God creates man. Man destroys God. Man creates dinosaurs." At which point, paleobotanist and would-be love interest Dr. Ellie Sattler jumps in to finish, "Dinosaurs eat man… woman inherits the earth."[1]

While Dr. Sattler's vision has yet to materialize, Dr. Malcolm should have given more credit to God's staying power. Viewing a film like *Jurassic Park* with a critical eye, it becomes clear that there are many elements we associate with ancient life – namely both biological and geological chaos, violence, and exaggerated capacities for destruction – that lack any firm scientific grounding. These associations have been around for much longer than many people realize. For the past 200 years, much of what we envision when someone says the word "dinosaur" comes directly from the attempts of late 18th and early 19th century geologists, a fractious and fascinating set of characters, to reconcile their new fossil discoveries with Scripture. The first wave of visual representations of prehistoric reptiles – in particular the mezzotint frontispieces of Romantic painter John Martin – drew inspiration from both Scriptural text, especially the first three lines of Genesis, and from the long tradition of Biblical artwork. These influences, operating at the peak of the Romantic Movement, had a profound effect in determining how the 19th century public imagined prehistoric reptiles and the antediluvian world they called home. Over the past two centuries, while our knowledge of dinosaurs has increased tremendously, remnants of this religious influence still linger on.

This paper will trace the origins of the scientific understanding of extinction and Deep Time that made visual representations of dinosaurs possible. I will present a brief historical retrospective of early representations of ancient life, including artwork by Henry de la Beche, Emile Boblaye, and John Samuelson Templeton. I will explore the development of Christian ideas in early geological thought, and then discuss the reasons behind John Martin's entrance into artistically representing the ancient world, his association with the Romantic movement and impact on his painting, his background as a historical-biblical and landscape painter, his commissioners (Gideon Mantell, Thomas Hawkins, and George Richardson), and his lasting contribution to our interpretations of the ancient past.

In Search of Deep Time

At the dawn of the eighteenth century, much of the educated Western world believed the earth to have begun on or around October 23rd, 4004 BC; fittingly, a Sunday.[2] This chronology was developed by Bishop James Ussher through the compilation and careful selection of a massive collection of historical sources dating back to Nebuchadnezzar of Babylon, and the Masoretic text of the Tanakh for everything older.[3] Enlightenment thinkers were delighted by Newton's achievements in mathematics and physics, and many a young geologically minded empiricist hoped to explain which natural laws resulted in the transformation from a flat, uniform Edenic Earth to the "gigantic and hideous ruin"[4] of mountains and channels and valleys seen in the present day.

New fossil discoveries forced the reality of extinction upon the scientific community by the late 18th century, but the mechanism behind it was still not agreed upon. Georges Cuvier, then studying at the Museum of Natural History in Paris, was a catastrophist: he believed that present-day geologic observations could be explained by a series of sudden natural disasters. Cuvier's primary rival was Charles Lyell, of the uniformitarian school of thought, which maintained that the same geological processes that can be observed today – wind and water erosion, gradual uplifting of land, deposition of sediments – are responsible for the formation all geological features found on the planet. This process by definition must happen slowly, over huge timescales. Suddenly the distant past was not just thousands of years ago but hundreds of thousands or even millions. Cuvier wrote angrily that his opponents were arbitrarily adding thousands of years "with the stroke of a pen."[5] While Cuvier felt concern that his side was losing the scientific war, he would later be vindicated – artistic rep-

resentations of the prehistoric world from Cuvier's time to the present have retained a distinctly catastrophic flavor.

In 1811, Mary Anning, then twelve years old, and her brother Joseph discovered a remarkable find. Thirty feet above the shoreline in the Blue Liassic cliffs of Lyme Regis, in Dorset, UK, erosion had exposed a strange reptilian fossilized skull. The Annings dug out the surrounding rock and removed the skeleton. The mystery animal had the fins of a dolphin, the pointed snout of a swordfish, the teeth of a crocodile, the spine of a fish, and the chest of a lizard. The find was reported in *The Times* of London, and a local landowner, Lord Henry Henley, bought the fossil from the Annings for £23 and had it sent to the London Museum of Natural History.[6] Charles Konig named the creature Ichthyosaurus, or "fish-lizard." Henry de la Beche adopted this name, and Cuvier determined it was indeed a reptile. The first large extinct marine reptile was officially scientifically described. A new era had begun.

After the Anning Ichthyosaur, a fossil frenzy was launched and new finds became more common. Mary Anning discovered the world's first plesiosaur skeleton (a totally different, long necked marine reptile) in 1821, which the Rev. William Conybeare described.[7] The eccentric and irrepressible William Buckland described the first dinosaur, Megalosaurus, in 1824, to a packed audience at the Geological Society of London.[8] Physician turned geologist Gideon Mantell named Iguanodon in 1825 from a set of large teeth that Cuvier said resembled those of a modern iguana.[9] In 1828 Buckland credited Anning with the first pterodactyl found outside of Germany, which caused quite a commotion at the Geological Society.[10] In 1841, with only three species positively identified,[11] Richard Owen coined the word, "dinosaur."

Imagining The Past: Spy Holes in Time

Initially, there was significant reluctance among fossilist circles to publish any sort of recreation of deep time, for fear of seeming fanciful or speculative and losing credibility. If any fossilists attempted to artistically represent any sort of prehistoric reptile before 1830, that record has been lost to history. Publications were limited to detailed sketches of the actual skeletons. The Rev. William Conybeare even felt uncomfortable publishing skeletal reconstructions of plesiosaurus for three years after his initial naming and description of the animal, until Mary Anning found a more complete skeleton.[12]

Remarkably, this reluctance to publish professional scientific illustrations of deep time resulted in some of the first depictions of deep time fauna taking the form of satirical cartoons, gently poking fun at some or other member of the geological community. One of the most important of these is Henry de la Beche's cartoon entitled *Awful Changes* (Figure 1), which poked fun at a statement by Charles Lyell. Lyell had recently advanced a theory that took uniformitarianism a step perhaps too far, considering its total lack of evidence; that natural history in the long term is cyclical. Not only will geologic features of the past return, but so will biological features, and "The huge iguanodon might reappear in the woods, and the ichthyosaur in the sea, while the pteordactyle might flit again through umbrageous groves of tree-ferns."[13] De la Beche, an Englishman, determined to make sure this idea was laughed out of court, and drew up a cartoon featuring a "Professor Ichthyosaurus" lecturing at a podium to an audience of ichthyosaurs, plesiosaurs, crocodiles and pterodactyls about the weak teeth and useless jaws of an extinct human skull. While the cartoon is appealingly clever, it also may show some insight into what geologists presumed their discoveries actually looked like.

While some artistic license obviously must be made (some of the animals are wearing spectacles, after all), it is interesting to note the beaklike snouts, the depiction of eyes and nostrils, and the lack of a dorsal fin, not to mention the overall cheeriness of de la Beche's ancient tree-fern-ringed lecture hall. The overall message is that the ancient world was a very civilized place – and even though that message is being gently parodied, it is still the primary message of the cartoon. Clearly, we can see that the connection between dinosaurs and violence/chaos does not originate from within the geological community, who, with a few key exceptions, tended to think of their discoveries as not altogether unpleasant creatures.

De la Beche was able to lampoon members of the Geological Society with such impunity in part because earlier that year, he had drawn the first true scene from deep time. Entitled "*Duria Antiquor*: A More Ancient Dorset" (Figure 2), the cartoon depicts an aquarium-style view of the ocean, showing both above and underwater images of plesiosaurs and ichthyosaurs, crocodiles and fishes and turtles and pterodactyls and ammonites. While several of the animals are engaged in consuming one another, this was not an altogether uncommon mode of depicting wildlife, and in fact the mood of the scene is generally relaxed and humorous. A crocodile smiles from the corner; several ichthyosaurs are lounging in the shallows and spouting cheerily.

Fig. 1 De la Beche H. "Awful Changes." 1830. Reproduced in Scenes from Deep Time, Rudwick MJS.

Some of the animals even appear to be defecating, perhaps a winking nod to Buckland, who studied coprolites. Once again, we can see a much lighter style than what was to come. The scene is filled with color, and the expression on many of the animals is playful. We can see everything in the frame; nothing is obscured by water or shadow. The overall impression is that the ancient world was not altogether frightening or threatening or even mysterious, but, following classical Enlightenment thought, perfectly orderly and understandable.[14] Although it was never included in a scientific publication, Buckland used *Duria Antiquor* during several lectures, and lithograph prints sold briskly. The world was undeniably ready for more glimpses into deep time.

While *Duria Antiquor* never reached mainstream saturation, it was widely enough distributed to spawn numerous copy-cat attempts. In 1834, across the channel, a steel engraving entitled *Extinct Animals* appeared next to an article of the same name by Emile Boblaye in the multivolume *Illustrated Dictionary of Natural History* (Figure 3).

Boblaye's work was increasingly speculative; it, like the other engravings in the collection, was brightly colored, with plesiosaurs and ichthyosaurs painted in dark grey, much like a modern whale. By presenting the ancient world in the same artistic style as the other paintings in the collection, UC San Diego Professor Martin Rudwick argues, Boblaye allowed readers' imaginations to soar as they perceived the fauna of deep time to be just as real as the fauna of, say, Patagonia or Borneo. Boblaye pushed the envelope even further; in the accompanying article, he expresses support for Lamarck's theory of transmutation,[15] which, though considered anathema in England, was steadily growing in popularity across the English Channel.

This growing possibility that fossils could be used to express dangerously "incorrect" or heretical viewpoints lay behind the next wave of published material and was vital to the eventual integration of Christianity-influenced elements into depictions of prehistoric life. William Buckland's 1836 treatise *Geology and Mineralogy: considered with reference to natural theology*, commissioned under the Earl of Bridgewater, consisted primarily of neither geology nor mineralogy but of anatomical analyses of extinct animal fossils, with the intent of showing that they were at least as well adapted to their environments as modern animals.[16] This would supposedly undercut the insidious "French" notion that organisms in their present state have evolved from lower, less well-adapted forms.

Fig. 2 De la Beche, Henry. Duria Antiquor. 1830. Original lithograph by George Scharf, reproduced in Scenes From Deep Time, Rudwick MJS.

Fig. 3 Pfitzer, Johann. 1834. Extinct Animals. For Emile Boblaye. Reproduced in Scenes From Deep Time, Rudwick MJS.

And then there was Thomas Hawkins. Hawkins was a young man of considerable financial means who had acquired one of the most impressive fossil collections in England. He was also a flamboyant eccentric who would later be accused of fabricating large portions of said fossil collection after selling it to the British Museum for £1250.[17] Gideon Mantell, concerned that his name had been tarnished by association with Hawkins' duplicitous transactions, said of Hawkins, "The fact is Mr. H is a very young man who had more money than wit and happened to take a fancy to buy fossils."[18] The scandal led to a special investigation by the House of Commons, and was widely reported in popular press.[19] In spite of his character flaws, Hawkins managed to acquire friends and defenders in the upper echelons of the Geological Society, including William Buckland, who threatened to prosecute Hawkins' critics for libel.[20] This effectively ended the matter, and allowed Hawkins to get away with fraud and a mostly intact reputation with the general public, while generating significant publicity. Always the self-promoter, Hawkins would not let the opportunity go to waste.

Hawkins was a flood geologist – of a sort. He held to his own peculiar brand of Biblical literalism, which allowed for significant spans of time in between the second and third lines of Genesis, in which "the earth was formless and empty, darkness was over the surface of the deep, and the Spirit of God was hovering over the waters."[21] When interpreting fossil remains, Hawkins stuck to this paradigm, which featured successive stages of antediluvian life, all destined for watery annihilation in the Great Deluge. Who better than Hawkins to narrate his reconciled natural history in his own uniquely florid prose: "In the Beginning, chaotic, void, nascent, shadowy, nerveless negative…. Is the interpretation of Nature and Science and Reason and History and Revelation…then a Vision of Abysmal Waters, swarming with all wondrous creatures of Life, and gelid Swamps with amphibious things, and Dragon Pterodactyles flitting in the hot air with Vampire Wing."[22] The ichthyosaurs and plesiosaurs in his collection came from this ancient world of darkness and chaos.

Hawkins wanted to portray this vision to the educated people of the world through a series of books featuring detailed sketches of his fossil collection accompanied by text outlining this natural history and denouncing the twin evils of Lyellian Uniformitarianism and Lamarckian transmutation. Ideally the frontispiece for this masterwork would be a lithograph depicting the pre-Creation world of the ancient reptiles in all its dark, violent, chaotic splendor. Unfortunately for Hawkins, the artist he picked for his first book,[23] John Samuelson Templeton, either did not understand his commission or utterly failed to realize it. The result, *Extinct Monsters of the Ancient Earth* (Figure 4), was a well-lit and listless scene, featuring two miniature plesiosaurs with unfortunately kinked necks like coat hangers, one ichthyosaur beached glumly under a palm tree, and another barely poking its head above the surface, jaws slightly relaxed, perhaps from boredom. Hawkins was evidently displeased, and decided to seek an illustrator who could more accurately depict his vision of the deep past. He would eventually find his man: John Martin.

John Martin and the Great Shift

Although seldom remembered among the great Romantic landscape painters as more

Fig. 4 Templeton, John Samuelson. 1834. Extinct Monsters of the Ancient Earth. Reproduced in Scenes From Deep Time, Rudwick MJS.

than a footnote, during his lifetime John Martin was considered a rival and equal of Joseph Turner and hailed by the Magazine of the Fine Arts in 1833 as "among the greatest geniuses of all time."[24] Born in 1789 in Northumberland, he was the twelfth child of William Fenwick Martin and Isabella Thompson, and his family barely subsisted above the poverty line.[25] When John was fourteen, the elder Martin became aware of his son's extraordinary aptitude for drawing, particularly his fascination with the interplay of light and shadow. In 1806 Martin was apprenticed to the Italian artist Boniface Musso, and in 1810 first exhibited his paintings before the Royal Academy. His break came in 1816, after selling a painting called *Sadak in Search of The Waters of Oblivion,* based on Ridley's *Tales of the Genii,* to William Manning, a governor of the Bank of England. Shortly after he found himself painting historical scenes and landscapes for Prince Leopold and Princess Charlotte of Belgium.

John Martin soon became a household name, synonymous with dark and elaborate grandiosity of subject, theme, and style. A devotee of the Romantic Movement in every sense, Martin's artwork is characterized by the dominance of an exuberance of raw emotion over the finer points of artistic accuracy and logic. For instance, in *The Great Day of His Wrath* (Figure 5), enormous boulders descend from the heavens towards the judgment of the wicked, while *The Deluge* (Figure 6) depicts a tremendous volume of water ascending and descending from all sides with little regard for gravity or normal wave dynamics. Martin's visual style can perhaps best be described as Post-Enlightenment: the power of Nature and her God are capable of overwheliming..This made him the perfect artist for Hawkins' "execrable and dreary things in the abounding chaos." It would also have serious repercussions for how ancient life was viewed for the next 200 years.

Martin was at the height of his popularity in 1838 when Gideon Mantell commissioned him to execute a mezzotint for the frontispiece of his book, *Wonders of Geology,* based on his popular paleontology and geology lectures at Brighton. Martin had some experience with geology at this point; he had met Cuvier, after all, and Martin was familiar with Mantell, having frequented his fossil museum. Mantell, unlike most of his fellow Geological Society members, was not wealthy. A country physician by trade, he gave up more and more of his profession to devote time to his fossil collection and museum. This paleontological obsession, while presumably rewarding, ruined his marriage and meant that he was constantly experiencing financial peril.[26] Mantell desperately needed this book to be a popular success, so he turned to Martin, with the hope and expectation that Martin's popular "Gothick" style would generate publicity and entice readers to purchase the text.

The result (Figure 7) was a radical departure from the de la Beche influenced tradition of brightly colored, labeled, educational scenes filled with cheerfully cartoonish animals. Martin's unearthly antediluvian vision, *The Country of the Iguanodon*, is utterly nightmarish. Martin's Iguanodons are true monsters, based only cursorily on Mantell's fossil reconstructions and drawing significant anatomical inspiration from the tradition of art based on St. George and the Dragon.[27] Apart from placing the Iguanodon's famed thumbspike on the beast's nose, as was customary for another half-century or so, Martin's creatures are grotesquely obese, with fearsome neck-wattles and thick rolls of flesh. Their skin is neither uniformly scaled nor smooth, in contrast to earlier works, but lumpy and irregular. Their eyes are dark and beady, ensconced by enormous (and anatomically non-existent) sockets. One animal, most likely a Megalosaurus, is attempting to devour an Iguanodon, which in turn is piled over another unidentifiable reptilian. The plants are suitably tropical; evidently, Mantell had shown Martin enough fossils for him to at least get the foliage correct. Tonally, Martin is the first artist to fully embrace shadow as a motif of the ancient world, shrouding most of the background in murky darkness, punctuated only by the disembodied rolling of sinuous shapes. The divisions between water and land appear unclear and rudimentary. Though Martin did not yet know Thomas Hawkins, the scene calls to mind the eccentric fossilist's oft-quoted Bible passage: "The earth was without form and void, and darkness was over the face of the deep." The Spirit of God does not exactly hover over the waters, but there is a mysterious, sunless bursting of light in the upper right hand corner, and one can't help but wonder. Mantell was evidently pleased with Martin's work, proclaiming in his book, "Iguanodon country – a country, which language can but feebly portray, but which the magic pen of a John Martin, by the aid of geological research has rescued from the oblivion of countless ages, and placed before us in all its hues of nature…" [28] *Wonders of Geology* went on to become highly successful with popular and technical audiences, and Martin's vision of deep time became deeply engrained in the mind of the general public.[29]

Hawkins must have been impressed with Martin's vision of the ancient world, because he commissioned the artist to create a frontispiece for his own 1840 book, titled in typical Hawkins fashion, *The Book of the Great Sea Dragons, Ichthyosauri and Plesiosauri, Gedolim Taninim, of Moses. Extinct Monsters of the Ancient Earth.* The frontispiece that emerged, entitled *The Sea-Dragons as They Lived* (Figure 8), is even darker and more "Gothick" than the scene Martin painted for Mantell. The dark waters roil and churn. There is a pale flicker of a sun, covered almost entirely by thick blackness. A dead ichthyosaur rots in the shallow water while a pterodactyl picks at its eye socket; this world before man was an unkind time ruled only by death and despair, Martin shows us. Three of the titular sea-dragons are engaged in battle. They are clearly intended to represent ichthyosaurs and plesiosaurs, yet Hawkins

Fig. 5 Martin, John. 1851-53. The Great Day of His Wrath. Reproduced in John Martin, Painter, Pendered M.

appears to have been even less diligent in providing the artist with instructions than Mantell, for both reptiles have claws instead of flippers, bulging binocular eyes, and the plesiosaur has a forked tongue. Every effort is made to intensify their ferocity, from their gaping mouths filled with sharp teeth to their posture, leaping in fury or else coiled and waiting for an opening to attack. The plesiosaurs in particular look especially snakelike, with a serpentine curve in the neck entirely unseen in previous artists' work. Perhaps Hawkins' title inspired him; Biblical scholars have alternately interpreted Gedolim Tanninim to mean "giant reptile" or "giant snake." Hawkins, always playing up his flair for the dramatic, wrote of the scene, "Wide over the desolate Seas warring Dragons innumerable and hideous, enacting Perdition… John Martin has barely attained, with all his stupendous Powers, the utter hideousness their own."[30] One is hard pressed to imagine how much additional hideousness any artist could extract from bones preserved in rock.

Martin created one final artwork on the pre-Adamite world, a scene entitled *The Age of Reptiles* (Figure 9), for George Richardson's 1842 book *Geology For Beginners.* While reducing the dramatic darkness of his previous work, Martin also abandons any pretext at anatomical accuracy, instead depicting two animals that appear to borrow features from both the lizard and the crocodile. His plesiosaur is entirely out of the water, which would have been physically unfeasible for such a massive aquatic animal. More importantly, though, Martin continues to partially submerge his creations in shadow as a representation of the unknown and mysterious elements of deep time, the exaggerated sharp teeth and forked tongues, the massive ungainliness, the constant associations with death. Interestingly, alongside all the previously discussed elements of his scene for Richardson, the title *The Age of Reptiles* has had one of the most lasting effects. It has become a widely accepted popular phrase for what we now consider the Mesozoic era, stretching from 250 million years ago to 65 million years ago, and has lent itself to everything from Rudolph Zallinger's 1947 mural at the Yale Peabody Museum[31] to a dinosaur-themed comic book series by Ricardo Delgado.[32]

Martin's Legacy

Martin's impact was gradual. Richardson switched back to his previous illustrator for his next edition of *Geology for Beginners*. Franz Unger's Primitive World sequences in 1851 – the first sequence of scenes representing different eras of prehistoric life – may have been based on Martin's tradition of showing sequences of events (e.g., The Deluge sequence), but stylistically owes little to Martin.

However, in 1855 W.F. Volliner, under the pseudonym W.F.A. Zimmermann, published *Wonder of the Primitive World*, which contains a clearly Martinesque frontispiece.[33] Titled *The Primitive World* (Figure 10), it contains every Martin-inspired detail from an obscured sun revealing crawling monsters obscured by heavy shadow, to clawed, serpentine, forked tongued plesiosaurs engaged in battle with an ambiguously drawn reptile that could be anything from an iguanodon to an ichthyosaur, to a catastrophic volcano in the background and what appears to be a lightning strike *setting the water on fire.* Violence, chaos, excessive geological activity; check, check, and check. This is clearly not the world we inhabit today, but a world that operates by different laws. God has yet to bring any semblance of order or structure to this truly primitive world. However, the meaning behind this frontispiece is considerably obscured when critical analysis is applied to Volliner's text. Volliner was far more scientifically rigorous than Hawkins was, and by the 1850s the concept of deep time was accepted by the growing scientific elite.[34] Volliner's artwork contradicts his goals for scientific accuracy because by 1855 such artwork had become expected, almost obligatory.

Fig. 6 Martin, John. The Deluge. 1828. Reproduced in John Martin, Painter, Pendered M.

Even as knowledge of the ancient past grew and the physical look of dinosaurs changed, certain elements of Martin's style remained present. Louis Figuier's extremely popular 1862 sequence of scenes, *The Earth Before the Deluge,* depicts for the Lower Cretaceous a scene entitled *The Iguanodon and the Megalosaur* (Figure 11). While the animals themselves are more closely modeled on Waterhouse Hawkins' famed Crystal Palace statues, elements of Martin remain prominent – the gaping teeth, the massive bulk, the bestial ferocity. In another Figuier picture, from the Lower Oolite, a dead ichthyosaur has washed up on the beach, while numerous others are filled with chaotic elements: massive waves, creeping shadows, dramatic storms, predation events, and the now-obligatory dueling ichthyosaur and plesiosaur. Martin's style is alive and well.

Conclusion

During the 170 years between Martin and the present day, our knowledge of the ancient world has grown exponentially. We

Fig. 7 Martin, John. The Country of the Iguanodon, 1838. Reproduced in Scenes From Deep Time, Rudwick MJS.

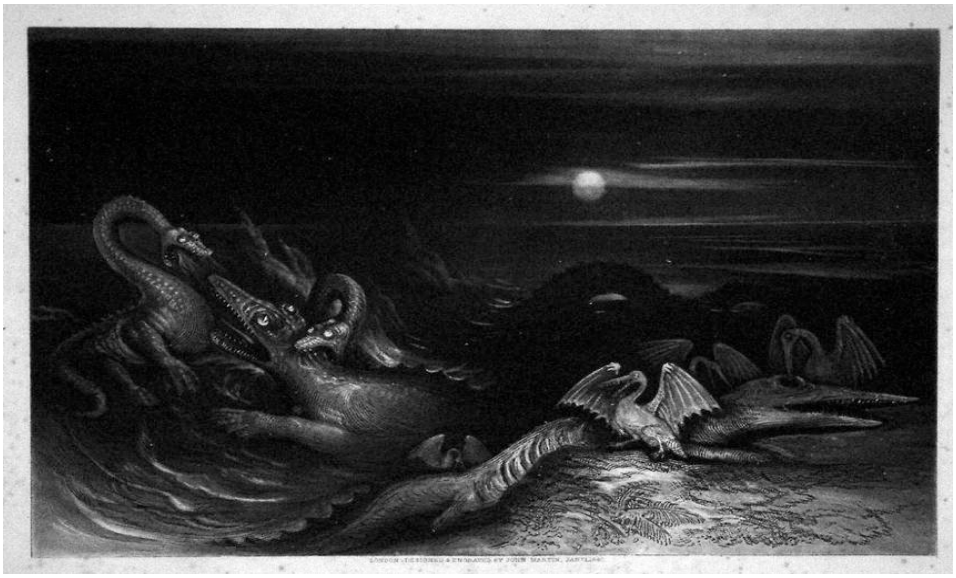


Fig. 8 Martin, John. 1840. The Sea Dragons As They Lived. Reproduced in Scenes From Deep Time, Rudwick MJS.

Fig. 9 Martin, John. 1842. The Age of Reptiles. Reproduced in Scenes From Deep Time, Rudwick MJS.

Fig. 10 Volliner, W.F. 1855. The Primitive World. Reproduced in Scenes From Deep Time, Rudwick MJS.

know now that life has existed on planet Earth for at least 3.8 billion years. We know now that ichthyosaurs and plesiosaurs were fully aquatic and viviparous, giving live birth to their offspring. We know that Iguanodon's thumbspike does not belong on its nose, that many dinosaurs walked upright on two legs, and far from the three species identified in the 1840s, Earth was once stalked by over 700 scientifically identified species of dinosaur.[35] Each successive discovery alters our interpretation of how this world looked. In the early 1990s a series of fossil discoveries in China proved conclusively that many non-avian dinosaurs had genetically avian feathers.[36] Only in the past eight years or so has this discovery become well-known to the mainstream public, but within time it is highly probable that a new, feathered visual paradigm for dinosaurs will emerge.

In spite of this rapid rate of discovery, many of our most prominent, almost instinctual associations with dinosaurs are traceable to the first 15 years of dinosaur artwork, and the work of John Martin in particular. Dinosaurs are big. Dinosaurs are lumbering. Dinosaurs are fearsome, with large teeth and claws. Dinosaurs inhabited a world very different from the one we are familiar with today. Dinosaurs are extinct now. All of this is true and obvious, yet the fact that these characteristics are still our go-to reaction upon hearing the word dinosaur is a testament to Martin's ability to form lasting images emphasizing all of these characteristics.

More interestingly, many of our common misconceptions or exaggerations regarding the prehistoric era are traceable to Martin and his immediate influences and antecedents. For instance, Lyell's uniformitarianism was eventually proven mostly correct – the same geological processes that shaped the earth 225 million years ago are still present today in about the same proportion. Volcanoes, earthquakes, catastrophic drought, continental drift – none of this was significantly more present during the Age of Reptiles than it is today. Why, then, do innumerable illustrations and film depictions ranging from Disney's *Fantasia* to Don Bluth's *The Land Before Time* and Disney's *Dinosaur* portray near constant geological and environmental disaster? Equally, why are death and violence still a constant presence in representations of the ancient past, when modern visual representations of nature's living biota only rarely include such topics? Perhaps the answer is because both Cuvier's catastrophism and Hawkins' interpretation of Scripture implicitly and explicitly support greater levels of geological chaos in the ancient world. The artistic tradition started by Martin entrenched these associations too firmly in our common psyche for mere scientific facts to shake them off. We still feel, on some instinctual level, that the ancient past operated by different natural laws than the present.

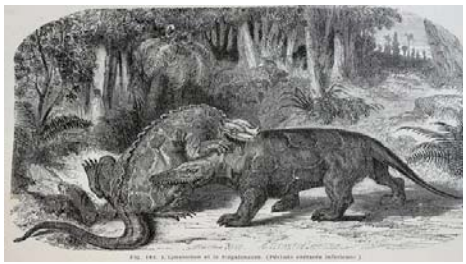


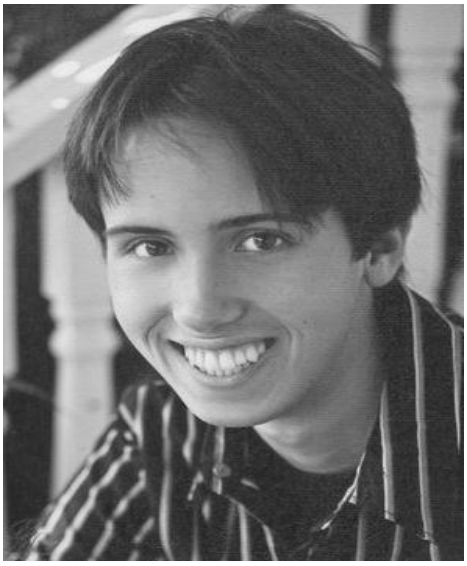
Fig. 11 Figuier, Louis. 1862. The Iguanodon and the Megalosaurus. Reproduced in Scenes From Deep Time, Rudwick MJS.

When a film like *Jurassic Park* uses dinosaurs as a manifestation and a motif for chaos theory, Spielberg isn't broaching new ground. Rather, he's relying upon a connection as old as our knowledge of dinosaurs themselves. John Martin never set out to change our view of ancient life, and he certainly never predicted that science would become excised from religion to the degree that it has. Yet by committing to paper a vision of the ancient world so profoundly influenced by non-empirical sources, he changed our interpretation of dinosaurs forever.

References

1. Jurassic Park. DVD. Dir. Steven Spielberg. 1993; Universal City, Ca. Universal Home Entertainment. 2000.
2. Gould SJ. "Fall in the House of Ussher", Eight Little Piggies. New York: W. W. Norton, 1996.
3. Gould SJ. "Fall in the House of Ussher", Eight Little Piggies. New York: W. W. Norton, 1996.
4. Burnet T. The Sacred Theory of the Earth. London: R. Norton, 1691.
5. Rudwick MJS. Georges Cuvier: Fossil Bones and Geological Catastrophes. Chicago:University of Chicago Press, 1997:228-229
6. Goodhue T. Fossil Hunter: The Life and Times of Mary Anning. Bethesda, Md:Academica Press, 2004:19-22.

7. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001: 69-71.
8. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001: 77.
9. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001:110-112.
10. Goodhue T. Fossil Hunter: The Life and Times of Mary Anning. Bethesda, MD: Academica Press, 2004: 57-58.
11. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001:176.
12. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992:42.
13. Lyell C. Principles of Geology. London: J. Murray, 1830.
14. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992:44-45.
15. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992:64.
16. Buckland W. Geology and Mineralogy, Vol. 1. London: William Pickering, 1836:292-303.
17. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001:144
18. Dean DR. Gideon Mantell and the Discovery of Dinosaurs. Cambridge: Cambridge University Press, 199:139.
19. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001:142-143.
20. McGowan C. The Dragon Seekers. Cambridge: Perseus Publishing, 2001:142-143.
21. Genesis 1:2.
22. Hawkins T. Memoirs of Ichthyosaurii and Plesiosaurii, extinct monsters of the ancient earth. London: Rolfe and Fletcher, 1834:5.
23. Hawkins T. Memoirs of Ichthyosaurii and Plesiosaurii, extinct monsters of the ancient earth. London: Rolfe and Fletcher, 1834: Frontispiece.
24. Pendered ML. John Martin, Painter. His Life and Times. New York: E.P. Dutton & Company, 1924:14.
25. Pendered ML. John Martin, Painter. His Life and Times. New York: E.P. Dutton & Company, 1924: 36
26. Mantell G. The Journal of Gideon Mantell, Surgeon and Geologist, covering the years 1818-1852. London: Oxford University Press, 1940: 139-142.
27. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992: 82
28. Mantell G. The Wonders of Geology. 2 vols. London: Relfe and Fletcher, 1839:
29. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992: 80-84.
30. Hawkins T. The Book of the Great Sea Dragons, Icthyosaurii and Plesiosaurii, Gedolim Taninim of Moses, Extinct Monsters of the Ancient Earth, With Thirty Plates copied from specimens in the authors collection of fossil organic remains, Deposited in the British Museum by Thomas Hawkins, ESQ, F.G.S. etc. London: William Pickering, 1840: Frontispiece text.
31. Zallinger R. “Age of Reptiles.” Yale: Peabody Museum, 1943-1947.
32. Delgado R. Age of Reptiles: Tribal Warfare. Dark Horse Comics. 1993-1997.
33. Volliner W.F. The Primitive World. 1855. Reproduced in MJS Rudwick. Scenes From Deep Time. Chicago: University of Chicago Press, 1992.
34. Rudwick MJS. Scenes From Deep Time. Chicago: University of Chicago Press, 1992:135-137.
35. United States Geological Survey. 2001. “How Many Types of Dinosaurs Are Known?” December 5, 2011.
36. Schweitzer MH, Watt JA, Avci R et al. Beta-keratin specific immunological reactivity in feather-like structures of the Cretaceous Alvarezsaurid, Shuvuuia deserti. J Exp Zoo 1999; 285-2:146-157.



Aaron Peterson is a junior from Santa Rosa, California. A History of Science and Medicine major and Creative Writing minor, he is interested in exploring the intersections between history, art, religion, ecology, and evolution.

The Bulb Vacuum System (BVS) – Negative Pressure Therapy for Healing Acute and Chronic Wounds in Underdeveloped Areas

Damir Ljuboja

Negative Pressure Wound Therapy (NPWT) is a vacuum-based, electric form of treating acute and chronic wounds, primarily implemented in the United States and other developed nations. Current NPWT devices are limited by cost, dependence on electrical input, large size, the need for a trained operator, and diminished user mobility during use. This research was focused on producing a simplified NPWT device which addressed these issues in conjunction with nonconventional sealant and bandaging substitutes for use in underdeveloped areas, relief efforts, and domestic first-aid kits. Through sealant and dressing, stress, comparative vacuum establishment, reverse flow/air-leakage, and bacterial growth tests, it was determined that a bulb-pump vacuum system with grocery bag or polyethylene dressing and Scotch tape sealant was the most effective. This manual system had no limitation on the amount of air it was able to displace, was able to endure temperatures ranging from -80 °C to 40 °C, reduced bacterial growth by 76.6%, and had a cost of $6.50, a decrease of 99.86% from high-end NPWT pumps. The functional and economic success of this system supports its incorporation in target areas and also has important implications for the development of similar acute and chronic wound treatments globally.

1 INTRODUCTION

Negative Pressure Wound Therapy is a treatment that aids in the healing of wound injuries by increasing blood flow to the target area and promoting the formation of granulation tissue[1,2]. It also functions in removing and draining fluid and leakage, diminishing both the amount and rate of infection and bacterial growth, assisting in closure, and providing protection from the surrounding environment[3,4,5,6]. Such a device typically is comprised of: 1) a gauze or wound dressing to fill the cavity, 2) a drainage tube held near the area of injury, sometimes inside of the dressing, 3) an adhesive sealant placed atop the dressing to create an air-tight seal, 4) a container or outlet for the fluid drained from the injured area, and 5) a low-pressure (around 40 to 200 mmHg, depending on severity and depth) vacuum to create negative pressure[2]. The dressing of the wound is likely to last no more than three days and thus requires frequent replacement[2]. According to the Food and Drug Administration, NPWT is proven to provide aid to chronic wounds, burns, diabetic and pressure ulcers, acute wounds, traumatic wounds, and dehisced wounds. In terms of clinical proof, NPWT has been shown to increase the rate of healing by 300% and, in some cases, even more[7,8]. This study aimed to simplify current negative pressure wound therapy systems and to significantly improve upon simplified versions that already exist for potential application in underdeveloped countries, disaster relief efforts, and in domestic first-aid kits.

1.1 Scientific Theory

The forces of macrostrain and microstrain resulting from negative pressure are the foundation of vacuum wound treatments[2,3].

Macrostrain is the stretching that occurs when negative pressure pulls the dressing as a vacuum. Physically, this pulls the edges of the wound together, causing a decrease in wound volume, and enables quicker healing. The vacuum distributes pressure evenly across the wound and prevents bacterial infection by bringing the edges of skin closer together under an air-tight environment.

Microstrain occurs on the cellular level, causing stretching, or microdeformations on the cell surface. This incites a response in the cells that increases metabolic activity, fibroblast migration, and extracellular matrix production. By this mechanism, the negative pressure wound therapy system reduces edema and promotes perfusion and the formation of granulation (connective) tissue that replaces fibrin clots in healing wounds. The granulation tissue engenderment is facilitated by the microstrain's promotion of cell migration and proliferation.

1.2 Problems with Current Designs

Classical electric NPWT pumps require a consistent power source to supply a vacuum. System inefficiencies, such as air-leaks, require a greater input of power for the desired function (200 watt maximum, 50 – 60 Hz electrical supply)[2,9]. Also, current battery life in the portable units is well below 24 hours[3,4]. Common complaints reference the fact that most NPWT systems today are big, bulky, and noisy with a plug-in system for take-home applications. The foremost dilemma is cost. Most of the time, the current system is difficult to cover with insurance[2].

This high cost may be partially attributed to digital readout screens, microprocessors to regulate vacuum strength in compensation of air leaks, and the need for a trained medical professional to apply the system. However, even after purchase, the actual use of such therapy devices by patients costs upwards of $100 per 2-3 hours of use due to their energy requirements[2,3,4]. Even in the developed world, NPWT devices are widely considered to be too expensive for the treatment of minor wounds. Prior studies indicate that a pertinent NPWT issue is the availability of electricity in foreign countries[2,11]. Power is not guaranteed in local hospitals, so a non-electrical device is ideal in such cases. Survey results also emphasize the difficulties associated with changing dressings and patient mobility. Thus a simpler lightweight design focusing on patient mobility and low cost is preferred (US$25 was deemed to be a very reasonable cost for pumps in free clinics) [2].

There are also problems associated with the widespread use of the only simplified NPWT device currently on the market. The design functions via limited suction provided by a bellows

pump applied to the wound area via plastic tubing (Fig. 1). This system is 252 mL in volume and 114 cm^2 in area and is difficult to transport[2]. Although it provides a strong negative pressure vacuum, some of this power could be sacrificed for a decrease in size and an increase in portability. There is no mechanism for transport while in use, such as an attachment to the body or a strap[2]. Bulkiness and poor portability must be addressed for simplified negative pressure wound therapy to be a practical choice in underdeveloped areas where the patient needs to continue working while being treated[12]. Furthermore, the simplified NPWT design in use does not draw fluid out at the wound site to keep it clean and sanitary but requires disassembly of the wrapping in order to clean the wound[2]. This problem may be solved by the incorporation of a one-way valve system that can be continuously pumped, which would pull fluid out of the wound area to prevent bacterial infection and other problematic situations.

1.3 Potential Markets

Potential markets are comprised of homes in the developed world, the armed forces, and countries in the third world, particularly parts of Africa and South America. Domestically, the use of simplified negative pressure therapy devices allows for instant access to treatment in retirement homes and work areas without considerable wait times and availability issues. On a global scale, simplified NPWT devices could aid disaster relief efforts[2]. In the third world, there are over 4 million cases of chronic wounds each year, of which 2.5% (classified as acute or chronic wounds) are good targets for NPWT[2,3,5,12]. The device developed in this study provides a cheap form of alternative treatment that could be used for these situations.

Fig. 1 Zurovcik's sNPWT Device, derived from a bellows plunger head.

2 METHODOLOGY

Tests conducted between September 15, 2010 and December 9, 2010 experiment with sealant and dressing combinations, stressful conditions, vacuum creation, reverse flow (air-leaks), and bacterial growth. Stress tests and bacterial growth tests were performed in an on-site laboratory.

2.1 Sealant and Dressing

Sealant tests were two-fold – potential wrappings were tested in conjunction with several sealants to determine which combination was the most effective at maintaining an air-tight seal. The first round of tests involved applying a bandaging to a flat, granite surface and applying sealants to keep it in place while a fan blew air across them (at medium setting). The air was smoothed out of the bandaging prior to fan application, and changes in compression, indicative of air-leaks, and tightness were recorded over time. The same experiment was then performed on the human body using the three best combinations of nonconventional wrappings and sealants as previously determined.

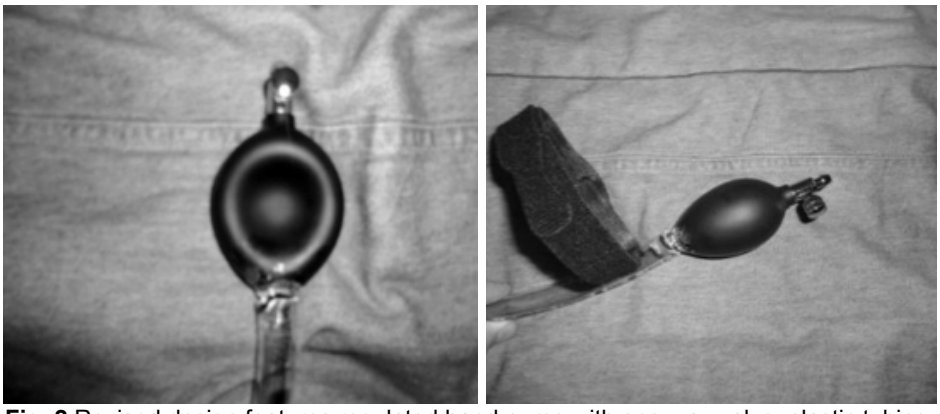


Fig. 2 Revised design features regulated hand pump with one-way valve, plastic tubing, and Velcro strap shown in both relaxed (left) and compressed (right) positions.

2.2 Stress

Stress tests were conducted using the top three combinations of dressings and sealants. These were two-step and involved exposure to cool (4 °C, -20 °C, and -80 °C) and warm (20 °C, 40 °C, and 80 °C) temperatures in a set of three refrigerators and a water bath, respectively. Temperature of water baths and the -80 °C refrigerator was regulated via external thermometers. Likewise, the 4 °C and -20 °C refrigerators had internal thermometers. Each combination was sealed at the tubing base of the bulb pump, and the air was removed manually to create a vacuum within the dressing. For cold trials, this set-up was placed in each refrigerator for 30 minutes, and an evaluation of the durability of the seal was made (if it continued to be air-tight or lost vacuum strength) every 10 minutes. In warm trials, only the end of the pump that had the sealant-dressing combination was submerged into the water bath for 30 minutes, and the durability evaluation was done. The overall evaluation was made by the level of compression of the bulb – "maintained seal" was deemed to be at 70-100% compression and "loss of seal" was at anything below ~70% compression for each set of conditions.

2.3 Comparative Vacuum Establishment

A comparative test between this design (Table 3) and the existing simplified NPWT model was conducted by measuring the amount of air able to be vacuumed out of and pumped into a rubber balloon (Wal-Mart Best Occasions Decorative Balloons). Both devices were used to draw out as much air as possible from a balloon uniformly blown up to a radius of 6 cm and a volume of ~905 cm^3. A similar test was conducted

Average for n = 3		**Dressing Used**		
Sealant	Plastic (grocery bag)	Polyethylene	Saran Wrap	Parafilm
Cake Frosting (cf)	> 6 hours (more than 7)	2.75 hours	2.30 hours	4.2 hours
Neosporin (petroleum jelly-base)	> 6 (but less than 7 hours)	2.25 hours	.50 hours	> 6 (but less than 7 hours)
Scotch Tape	> 6 hours (more than 7)	> 6 hours (more than 7)	4.5 hours	> 6 hours
Superglue	> 6 hours (more than 7)	> 6 hours (more than 7)	> 6 hours (more than 7)	> 6 hours (more than 7)

Key: selected for further testing

NOTE: Hold is on the basis of an **air-tight** seal.

Table 2 Sealant and Dressing Compatibility Test Results (n=3)

when pushing air into an empty balloon (initial radius of 0 cm). Measurements were made via tape measure and calculations of volume were determined using the formula for the volume of a sphere.

2.4 Reverse Flow/Air-Leakage

Reverse flow tests were conducted on a time-trial basis – the top three combinations from the initial sealant test were used in conjunction with the bulb pump and the ability for the wrapping to remain air-tight was recorded. The combination was applied to the right calf in all trials. The basis for this was observation of the bulb's compression/deflated state. Once the bulb was decompressed to below 70%, the system was said to no longer be air-tight. During these tests, each combination was subjected to a mile-long walk and .5 mile jog to simulate strained conditions.

2.5 Bacterial Growth

Bacterial tests were developed as a means of determining whether or not negative pressure wound therapy prevents infection by slowing bacterial growth. A plate of 20 mL LB agar (5 g bacto-tryptone, 2.5 g bacto-yeast extract, 5 g NaCl, 7.5 g bacto-agar, and 500 mL dH2O stock, autoclaved for 30 minutes on "liquid" cycle) was used for each trial. A final concentration of 50 µg/mL of Kanamycin was added to each plate to slow bacterial growth to an observable level. The plates, when solidified, were treated with 5 mL (10 mL in second trial) of Novagen BL21 (DE3) cells transformed with Carbonic Anhydrase VII plasmids using sterile technique near an open flame. The plates were placed in separate plastic grocery bags and the bags were closed with the tape sealant in an air-tight manner at the bag opening. For one of the bags, the plastic tubing of the pump was inserted prior to the application of the sealant, and a vacuum was created by extraction of all the air. Both bags were placed on the lab benchtop at room temperature; growth was observed periodically in intervals of 6 hours for 12 hours and then in 8 hour intervals for an additional 16 hours by breaking and re-establishing the aforementioned seals, each time for less than 3 minutes.

3 RESULTS

3.1 Sealant and Dressing

The top three dressing-sealant combinations were tape with grocery bag, tape with polyethylene, and cake frosting with grocery bag, respectively. During the time-trial test, these combinations retained the tightest seals and continued to do so beyond the maximum recorded six-hour mark. Three trials were conducted and average results are rendered in Table 2.

3.2 Stress Tests

Refer to Table 5 and Fig. 6. All combinations retained greater than 70% bulb compression in the colder temperatures (-80 °C, -20 °C, and 4 °C) at all times. When placed in 20 °C temperature water, at 10 minutes all combinations had greater than 70% compression, but after 20 minutes (and at 30 minutes), the cake frosting–grocery bag combination fell rapidly below 70% compression. For the 40 °C bath, at 10 minutes combination 3 fell below 70% compression. At 20 minutes nothing new occurred, but at the 30 minute mark, combination 1 lost its seal. For the 80 °C bath, all combinations lost their seal/bulb when compressed less than 70% at all points in time – the seal was lost within moments of placement into the water bath. Overall, the tape with grocery bag and polyethylene maintained over 70% bulb compression in 77.78% and 83.33% of instances, respectively. The cake frosting and grocery bag combination only maintained its seal 55.56% percent of the time.

3.3 Comparative Vacuum Establishment

Refer to Table 3. The ability of the pump to create a vacuum was judged relative to the existing NPWT design by means of a comparative expansion and contraction of a balloon with a limited/maximum diameter. On average, the existing design was able to deflate the initial 6 cm radius by 4.79 cm (460.4 cm^3 volume) and inflate it (initial radius for this part of the test was 0 cm) to a radius of 4.81 cm (466.1 cm^3 volume). The proposed design was able to inflate the balloon until it exploded and deflate it completely – there was no limit on the amount of air which it could provide or remove.

3.4 Reverse Flow/Air-Leakage

Deflation

Trial	*BVS Design*	*Design in Use (sNPWT)*
1	all (no limit)	4.72, 440.5
2	all (no limit)	4.84, 474.9
3	all (no limit)	4.81, 466.1
4	all (no limit)	4.77, 454.6
5	all (no limit)	4.79, 460.4
Average	**all (no limit)**	**4.79, 460.4**

Radius (cm), Volume (cm3)

Inflation

Trial	*BVS Design*	*Design in Use (sNPWT)*
1	exploded (no limit)	4.93, 501.9
2	exploded (no limit)	4.69, 432.1
3	exploded (no limit)	4.85, 477.9
4	exploded (no limit)	4.77, 454.6
5	exploded (no limit)	4.83, 472.0
Average	**exploded (no limit)**	**4.81, 466.1**

Radius (cm), Volume (cm3)

Table 3 Comparative Vacuum Establishment Results

Fig. 3 Example of human testing done with top three nonconventional combinations of sealants and dressings (this particular example is of cake frosting and plastic grocery bag). The image features an example prior to the application of the device (left) and after (right).

Refer to Table 4. In the first trial, the combination that maintained an air-tight seal for the greatest period of time was the grocery bag with tape and the high-polymer polyethylene with tape (tie). The grocery bag with cake frosting broke its seal (bulb inflated more than 70%) at approximately the 14 minute mark. This occurred during the transition from the mile-long walk to the .5 mile jog. During the second trial, a similar situation occurred – the grocery bag and polyethylene with tape combinations maintained air-tight seals during the entire test, but the cake frosting-grocery bag combination failed once more, this time at the 16 minute mark.

3.5 Bacterial Growth

Refer to Figs 4 and 5. Bacterial growth tests showed a significant dependence upon the application of the bulb vacuum. In the first trial (5 mL of culture applied to each plate), the plate subjected to the bulb vacuum showed the growth of 1 colony mass whereas the plate without the negative pressure had a growth of 5 colony masses. In the second trial (10 mL of culture applied to each plate), the use of negative pressure on one of the plates resulted in a growth of 7 colony masses. The plate without the vacuum applied had a growth of 26 colony masses. There was an 80% and 73% reduction of bacterial growth in these trials, respectively, through the application of NPWT. An average growth reduction of 76.5% was observed.

4 DISCUSSION

In initial sealant-dressing tests, medical dressing and medical tape were not tested in the human trials due to their current commercial application (known to be effective). Their high cost, however, eliminated them for potential use in conjunction with the goals of this experiment. For example, Sterling Silverlon Wound Dressing 4"X4" Argentum Medical (Box of 10) costs $418.00 and an associated adhesive, Dermabond Adhesive Liquid Stitches Vial .5ml High Viscosity (Box of 12), costs $538.83. The most effective of the tested sealant materials were the cake frosting and tape. For cake frosting, it was presumed that the effective bond was caused by its ability to adhere to itself and to particular, plastic solids. Although the cake frosting was messy, its ability to provide a cheap and effective air-tight seal as a temporary solution moved it on to further testing. The polyethylene sheet, although made from the same plastic as the grocery bag, was not as effective because it was double layered and, accordingly, required a greater amount of sealant to create and maintain a leak-proof closure. The parafilm adhered to the skin with the cake frosting but was not able to cling on as effectively when exposed to other liquids, particularly sweat that naturally occurred during human testing. The plastic bag was not only cheap and easy to obtain in a large quantity, but it attached firmly to the skin and, when frosting was applied, conformed to the desired shape. This aspect is of great value because, if the dressing is able to completely adjust to a needed shape, it will prove extremely useful due to this versatility in applicability. The tape sealant

Model	Number of Pumps	Range of Prices	Average Price	Difference Between Average Price and Medicare Purchase Price
Model 1	8	$1,085–$2,825	$1,995	$15,170
Model 2	60	$3,448–$6,173	$4,970	$12,195
Model 3	103	$2,449–$3,950	$2,934	$14,231
Weighted average			$3,604	$13,561

* The prices include documented discounts but not shipping charges.

Source: OIG analysis of supplier documentation, 2008.

Table 5 Prices that Suppliers Paid for New NPWT Models

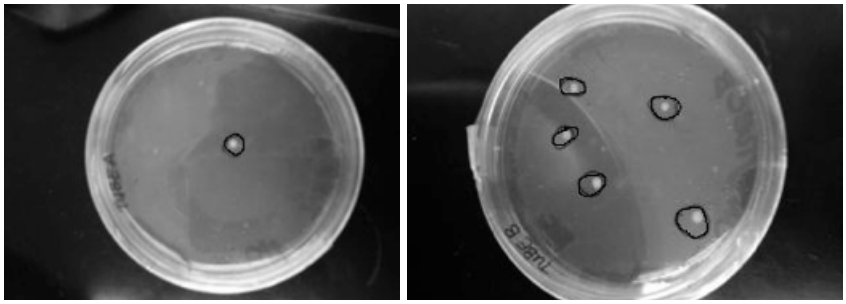


Fig. 4 Bacterial Test Trial 1 (5 mL cell culture) results from plates with (left) and without (right) the bulb pump applied.

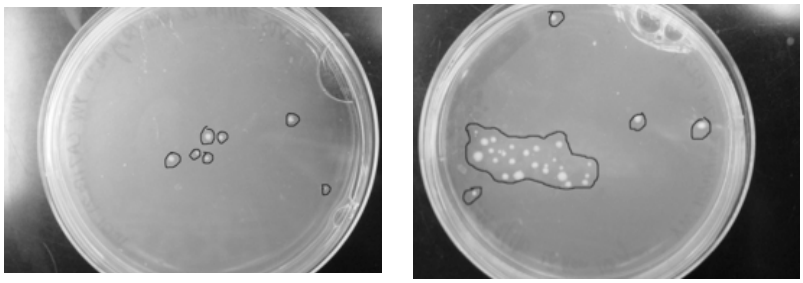


Fig. 5 Bacterial Test Trial 2 (10 mL cell culture) results from plates with (left) and without (right) the bulb pump applied.

Combination Used	Trial 1: Duration of Seal (minutes)	Trial 2: Duration of Seal (minutes)
Grocery Bag w/ tape	~17 minutes (entire time)	~17 minutes (entire time)
Grocery Bag w/ CF	14 minutes	16 minutes
Polyethylene w/ tape	~17 minutes (entire time)	~17 minutes (entire time)

Table 4 Reduced Flow / Air-Leakage Results

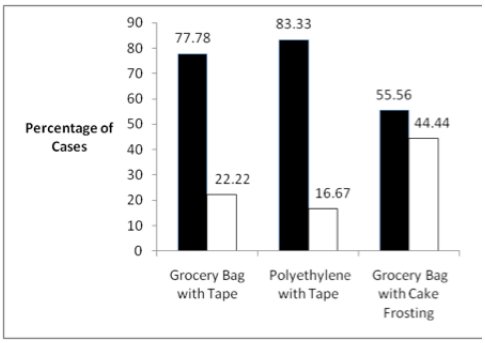


Fig. 6 Stress tests column representation of results for each combination with an overall visualization of the percentage of cases in which the seal was maintained versus those in which the seal was lost.

was just as effective as the cake frosting in these trials and was advanced to further testing in combination with the low-polymer grocery bag and the high-polymer polyethylene.

The stress test results further supported the use of the tape in conjunction with the grocery bag or the high-polymer polyethylene due to the long-lasting seal they created and maintained under stressful conditions. A general correlation was seen in all instances, indicating greater maintenance of the seal under cooler temperatures relative to warmer ones. Verified in all cases, this likely had to do with the melting point of the materials used, primarily that of the tape sealant, since plastics (polyethylene homopolymers) generally melt at temperatures upwards of 100 °C[13]. A concern in the conduction of this aspect of the testing was the influence of water on the higher-temperature region of the spectrum. Pressure from the water may have significantly reduced the ability of the bulb to remain compressed by leaking in through small holes in the dressing (mostly in the case of the grocery bag [low-polymer plastic]). Also, there was a concern that that hot water would dissolve the cake frosting after it was allowed to dry (prior to testing).

The vacuum establishment results were used to verify the implementation of the device developed over the existing design in underdeveloped areas. The ability of the small bulb to fully decompress a wound wrapping regardless of its size was a great advantage and allowed for flexible use. The existing design is suited only for wounds which contain less than ~450 cm^3 of air when wrapped. Given its large size relative to the bulb pump proposed in this report, a greater amount of possible air-flow was expected. However, this was not the case because the one-way valve in the bulb pump system constantly pushed air out as new air took its place. The existing design could remove a greater quantity of air but could do so only once. This prevents the design currently in use from adjusting to air-leaks that develop over time without completely dismantling the set-up.

In air-leakage tests, the performance of the grocery bag and the polyethylene dressings with tape was a success, defined as a trial in which the seal did not break. The cake frosting-grocery bag combination failed in both cases but was still effective throughout the majority of the time it was tested. The seal broke only when running and after at least .1 miles of distance had been traveled. This problem could have potentially been remedied by allowing the seal more time to dry before performing the trials (approximately 5 minutes were allocated for this). The tape and grocery bag and tape and polyethylene combinations were deemed to be effective candidates for bacterial tests given these results and those of prior tests.

Bacterial growth was successfully decreased via the application of the bulb pump in combination with the tape and grocery bag dressing. This was evident in the major decline of colonial growth on the plates.

In the construction of the bulb vacuum system, a $5 bulb and valve assembly (supplied by Mountainside Medical), 30 cent tubing, $1 Velcro strap, and less than 15 cents of glue (Gorilla Glue and Liquid Silicone) were used. The total cost of the aforementioned was ~$6.50 at Home Depot. This represented a decrease in cost by 99.4% - 99.9% (when compared with the $1,085 - $6,173 range of model costs in Table 1). This design was therefore 167 times cheaper than the former model and 950 times cheaper than the latter. This estimate did not include the energy cost of using each of the pumps (which is upwards of $100 per 2-3 hours)[2,3,4]. Wrapping/dressing and sealant costs were decreased from bundle packs of $418 and $539 to less than 10 cents and ~1-3 dollars for cake frosting and Scotch tape, respectively. This was a 99.98% decrease in dressing cost and a 99.6% decline in sealant cost. Overall, the grocery bag dressing is 4180 times cheaper than Sterling Silverlon Wound Dressing 4"X4" Argentum Medical, and the average between the cake frosting and Scotch tape cost is 270 times cheaper than a .5 mL vial of Dermabond Adhesive - Liquid Stitches.

This data supports the exploration of simplified NPWT devices such as this for implementation in target areas. The manual nature of this design, low manufacturing cost, ease of use, portability, and effective functionality address all of the problems associated with current designs and provide support for widespread use and development.

5 CONCLUSIONS AND FUTURE WORK

Through sealant and dressing, stress, comparative vacuum establishment, reverse flow/air-leakage, and bacterial growth tests, a bulb-pump NPWT made from grocery bag and high-polymer polyethylene dressings with tape sealant was determined to be effective. This manual system functioned at temperatures ranging from -80 °C to 40 °C, reduced bacterial growth by 76.6%, could displace an unlimited amount of air, and reduced manufacturing costs by 99.86%. Total price of development was ~$6.50, a 950-fold decrease compared to high-end

At 10 Minutes		Temperature (celsius)				
Combination	-80	-20	4	20	40	80
1) Grocery bag with tape	Y	Y	Y	Y	Y	N
2) Polyethylene (high polymer) with tape	Y	Y	Y	Y	Y	N
3) Grocery bag with cake frosting	Y	Y	Y	Y	N	N

At 20 Minutes		Temperature (celsius)				
Combination	-80	-20	4	20	40	80
1) Grocery bag with tape	Y	Y	Y	Y	Y	N
2) Polyethylene (high polymer) with tape	Y	Y	Y	Y	Y	N
3) Grocery bag with cake frosting	Y	Y	Y	N	N	N

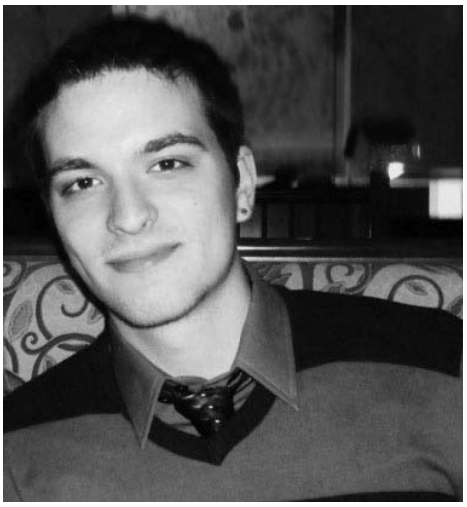
At 30 Minutes		Temperature (celsius)				
Combination	-80	-20	4	20	40	80
1) Grocery bag with tape	Y	Y	Y	Y	N	N
2) Polyethylene (high polymer) with tape	Y	Y	Y	Y	Y	N
3) Grocery bag with cake frosting	Y	Y	Y	N	N	N

Table 5 Stress Tests Results

Y - maintained seal
N - loss of seal

NPWT pumps[3,4]. Sealant price was reduced 270-fold , and dressing cost was decreased 4180-fold relative to common, medical-grade dressings and sealants[14]. This device could be used in underdeveloped countries in Africa and South America for relief efforts and as domestic first-aid kits.

Future research should further explore sealant-dressing combinations to optimize a simplified NPWT. Continued research will also focus on determining the most effective size of bulb to use. Also, bacterial tests with various vacuums will be explored to observe the impact of vacuum strength on bacterial infection rates and in conjunction with live test subjects. Additionally, expanded clinical trials focused on underdeveloped regions, such as Opsta Bolnica "Prim. Dr. Nakas" hospital in Sarajevo, Bosnia and local clinics across Nicaragua will take place.
gards to the upstream control of gene expression. At the clinical level, more research must be done linking ambient air exposure to chronic disease prevalence and severity, particularly with asthma. And at the molecular level, we need to further elucidate the genes and pathways that regulate Treg function. With a better understanding of the factors that can lead to asthma, we will be able to advance its pharmacological treatment as well as develop preventative measures for susceptible individuals.



References

1. Ennker IC, Malkoc A, Piertrowski D, et al. The Concept of Negative Pressure Wound Therapy (NPWT) after Poststernotomy Mediastinitis: a single center experience with 54 patients. J Cardiothorac Surg. 2009; 3:4-5.
2. Zurovcik DR. Development of a Simplified Negative Pressure Wound Device. Boston: Massachusetts Institute of Technology University Publications, 2007:1-188.
3. Hyun-Joo L, Joon-Woo K, Chang-Wug O et al. Negative pressure wound therapy for soft tissue injuries around the foot and ankle. J Orthop Surg Res. 2009; 4:1-4.
4. Pliakos I, Papavramidis TS, Mihalopoulos N et al. Vacuum-assisted closure in severe abdominal sepsis with or without retention sutured sequential fascial closure: a clinical trial. Surgery 2010; 148:947-953.
5. Beno M, Martin J, Hatzl J. Vacuum sealing in the treatment and prevention of prosthetic infection - A Case Review. Rozhl Chir 2010; 8:508-512.
6. Bloom DE, Sachs, JD. Geography, Demography, and Economic Growth in Africa. Brookings Papers on Economic Activity 1998; 2:207-273.
7. Moshiro C, Mswia R, Alberti KG et al. The importance of injury as a cause of death in sub-Saharan Africa: results of a community-based study in Tanzania. Public Health 2001; 2:96-102.
8. Mendonca DA, Cosker T, Makwana NK. Vacuum-assisted closure to aid wound healing in foot and ankle surgery. Foot Ankle Int. 2005; 26:761-766.
9. Levinson DR. Comparison of prices for Negative Pressure Wound Therapy Pumps. Washington D.C.: Department of Health and Human Services, 2009:23.
10. World Energy Outlook. Implications of Current Energy Policies. In: OECD/IEA, eds. Global Energy Trends Until 2030. World Energy Outlook, 2009:109-134.
11. Davies JB, Sandstrom S, Shorrocks A et al. The World Distribution of Household Wealth. Gothenburg: Gothenburg University Publications, 2008:1-24.
12. Ohene S, Tettey Y, Kumoji R. Injury-related mortality among adolescents: findings from a teaching hospital's post mortem data. BMC Res Notes 2010; 3:12-14.
13. Polyethylene Homopolymers-A-C-8. In: Honeywell, eds. Material Safety Data Sheet. Morristown: New Jersey Press, 2005:8.
14. Sullivan N, Snyde DL, Tipton K et al. Negative Pressure Wound Therapy Devices. In: Institute E, eds. Technology Assessment Report. Rockville, 2009:492.

Damir Ljuboja is a junior from Sarajevo, Bosnia, double majoring in Biology Honors and Humanities Honors at the University of Texas at Austin. Alongside exploring his scientific research interests, Damir enjoys rowing in the sweltering southern weather, producing electronic music, and writing poetry. He will apply to medical school this summer but remains open to the possibilities of the future, whether they lie in an academic hospital setting, a translational research lab, or a clinic abroad.

A thermal model for analysis and control of drilling in icy formations on Mars

Akul Aggarwal

The presence of ice poses a threat to drilling systems acquiring rock or soil samples on Mars. Testing has shown that heat generated by drilling produces water vapor or liquid which may refreeze on the surface of a drill bit. Further experience has shown that the adhesion of ice is sufficient to cause drill bits to stall, resulting in permanent loss of the bit. In addition, heat created during drilling can impact the chemical and physical composition of the sample even at temperatures below the melting point of ice, so restrictions may be imposed upon the allowable heating of a soil or rock formation during a mission. Developing a thermal model of the bit and formation is essential to prevent the loss of scientific data and drill bits, as well as to prevent complete failure of the mission in a spacecraft carrying only a single drill string. Using only properties of the formation and drilling system, a computer model has been created to predict temperatures throughout the bit and formation. This model has been shown to be in agreement with preliminary data acquired from tests performed under Mars-like conditions and using a prototype Mars Sample Return (MSR) drill. This model may be used to create schedules for drilling operations by determining the frequency and duration of pauses necessary to let the borehole cool before drilling may safely resume. This method has advantages over alternative scheduling methods such as monitoring drill parameters (e.g., motor torque) which may only change after ice has already frozen to the bit. The thermal model will be an essential tool in safely acquiring unaltered geologic samples, either for in situ analysis or for return to Earth.

I. INTRODUCTION

Low temperatures and pressures on Mars create ambient rock and soil conditions in some locations that are near the triple point of water[1]. The presence of water ice in the top meter of soil is well documented by the Mars Odyssey mission[2] and ice has been sampled in one location by the Phoenix lander[3].

When water ice is contained in rocks or soil, heat created by friction during coring or drilling may be sufficient to cause sublimation of the ice, causing gaseous water and rock cuttings to eject from the borehole[4]. Although this phenomenon may be useful in assisting or replacing the auger's role in removing cuttings from the borehole, it presents a serious problem in the form of vapor-deposited ice. Gas particles which contact cooler areas of the bit further from the cutting site may refreeze, creating deposits of ice which grow throughout the duration of a test. The adhesion of ice onto many metals is substantial: the shear strength of ice adhering to steel at temperatures below -7 °C is about 2 MPa[5]. As these deposits grow, they have the ability to freeze to the borehole or create a rough surface that may restrict the ability of the bit to rotate. Additionally, these deposits may clog the auger channels and inhibit the ability of the bit to remove cuttings from the borehole. Removal of the cuttings from the borehole is essential, otherwise existing cuttings would be reground indefinitely and further penetration would be impossible[4].

An additional consequence of heating while drilling or coring is the modification of the rock or soil content and subsequent loss of scientific validity of the samples. Evolution of volatile materials located within the geologic sample, due to substantial heating of the sample's contents, will be detrimental to analysis. Therefore, a heating system which can dislodge a frozen bit is not an acceptable means of acquiring samples, although such a system may be useful as a last resort on a mission with a single or limited quantity of bits[4,6].

For these reasons, regardless of the water content of the formation to be sampled or the ambient temperature and pressure, a temperature threshold may be placed as a requirement for sample acquisition operations. A thermal model will be vital in ensuring that this requirement is satisfied for a number of reasons. Using a set of sensors to measure temperatures of the rock while drilling is extremely challenging to implement from a hardware design point of view, considering that bits will likely be changed or a single string may be constructed from a stack of segments, and that a large number of slip rings would need to be incorporated into the drill. Additionally, an embedded sensor would not be in direct contact with the formation and might not necessarily measure the hottest location of the rock or soil during all sample acquisitions, as this location may vary depending on the composition of the formation.

A computer simulation with the sampling system's operating parameters and the thermal parameters of the rock, which may be estimated from inspection of the rock using cameras or other sensors, could provide a complete and accurate thermal profile for the duration of the test. This model could be used to create drilling schedules, composed of periods of drilling and pauses to allow the borehole to cool, which could keep all locations within a rock below a specified threshold temperature[4]. Such a schedule would aid in protection of the scientific validity of samples, prevention of lost bits, and acquisition of samples in minimal time.

II. EXPERIMENTAL SETUP

In order to acquire data to test the validity of a thermal model, a sequence of drilling tests was run on a block of 45 MPa Indiana limestone. These tests were conducted in a vacuum chamber at 6.0 torr to simulate Mars pressure and specifically to significantly reduce the convection of heat to the atmosphere. The Honeybee Robotics System for Automated Subsurface Sampling (SASSI) corer/drill, a prototype for an MSR mission, was equipped with a 2 cm diameter full face powder-acquisition bit used to conduct all tests. The SASSI can be seen in Figure 1, along with the arm and z-stage used to position and preload it against the rock.

Table 1 provides relevant parameters of the rock and drill, while Table 2 summarizes the tests performed. All of this information was input into the computer simulation to model each of the experiments.

Four 4-wire resistance temperature

Fig. 1 The Honeybee SASSI, a prototype for MSR, was used for all tests. It is equipped with the full face powder acquisition bit. The SASSI is a rotary-percussive drill, although percussion was disabled during these tests. The arm and the translation stage can be seen in the right of the image.

detectors (RTDs) were placed on and inside a block of Indiana Limestone at a variety of locations in order to provide surface and interior temperature values. These RTDs can be seen installed on and in the limestone block in Figure 2. The positions of the RTDs are displayed in Figure 3. Henceforth, RTDs 1 and 2 are referred to as exterior RTDs, since they are mounted to the top face of the rock, while RTDs 3 and 4 are referred to as the interior RTDs. The RTDs are attached or potted using thermally conductive epoxy to prevent additional thermal resistance and provide firm fixture. The entire experimental setup, including the chamber interior, limestone block, and support and positioning structures, is visible in Figure 4. The Omega F2020-100-B-100 RTDs used for testing have a maximum error in accuracy of about ±0.5 °C at the temperatures encountered during testing.

The goal of each test was to acquire thermal profiles at the locations of the RTDs as the rotating bit generated heat during drilling. Although the SASSI is capable of rotary percussive drilling, only rotary drilling was enabled for these tests. In test 1, a rest period allowed the drilling site to cool before drilling resumed. No active cooling system was employed; all cooling refers to equilibration of hot material with cooler areas of rock further from the drilling site. Test 2 was an uninterrupted test. Both tests were concluded with rest periods.

III. DISCUSSION OF THERMAL MODEL

The computer model, which was programmed in MATLAB, is based upon a cylindrical discretization of the formation and bit, as well as the temporal discretization of the duration of the cutting process of the bit. An axisymmetric assumption was used, allowing a two dimensional display along the r and z-axes to represent the entire system. The code is based upon the discretized heat equation. At present, the simulation allows for full-face drill bits of any size, although coring bits have not yet been incorporated into the code. Frictional heat is supplied at the interface of the bit cutting surface and rock at the bottom of the borehole. In accordance with results from research on frictional heating in drilling, 15% of this heat enters the bit, and the remaining enters the rock at the bottom of the borehole[7]. Friction is reduced when pressure is reduced to that of ambient conditions on Mars, so it is crucial that the correct frictional heating power levels be input into the model[8].

The outer edge of the bit and the inner surface of the bit are assumed to be in thermal contact at all times. Therefore, thermal equilibrium is enforced between each element of the bit along bit's exterior edge and each corresponding element on the innermost surface. No heat is assumed to be created at these sites. The simulation allows for penetration and power input to be paused in order to simulate a rest period, but heat transfer still occurs within the bit, within the rock, between the outer surface of the bit and inner surface of the borehole, as well as at the bottom of the hole where thermal equilibrium is enforced between the drill elements and the corresponding rock elements beneath.

The rock simulated in the tests was 160 mm in radius - extending about four times as far from the bit as the furthest location actually measured in the experimental testing, and 120 mm in height

Component	Material	Conduction Coefficient	Heat Capacity	Density
		[W/m-K]	[J/kg-K]	[kg/m^3]
Rock	Indiana Limestone	1.3	880	2760
Bit cutter	Carbon Steel	54	490	7850

Table 1 Rock and drill parameters

Test #	Borehole Center X-Pos	Borehole Center Y-Pos	Avg Auger Power	Average Rate or Penetration (ROP)	Duration/Rest Periods
	[mm]	[mm]	[W]	[m/min]	
1	0	18.58	6.75	2.40E-04	Drilled for 9 minutes, rested for 24 min. Drilled for 46 minutes, rested for 21 min.
2	22.46	-18.46	7.01	1.50E-04	Drilled for 90 minutes, rested for 28 min.

Table 2 Summary of tests.

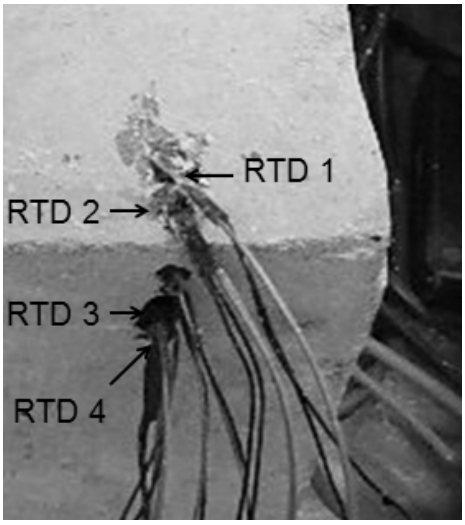


Fig. 2 The set of 4-wire RTDs inserted in and on the limestone block.

- nearly 10 times the depth of penetration of any test. Element spacing in MATLAB was 2 mm in the r-direction, and about 0.2 mm in the z-direction (exact spacing is calculated based on the penetration rate of the experiment to be simulated).

IV. COMPARISON OF SIMULATION AND EXPERIMENTAL DATA

The recorded experimental data and the predictions created by the computer simulation are shown in Figures 5 and Figures 6 for test 1 and test 2, respectively. Ambient temperatures were 25.4 °C and 24.2 °C for the tests. As can be seen in Figure 5 and Figure 6, the experimental and simulated results show good agreement, with average errors being 0.86 °C and 1.26 °C, repectively. These errors correspond to 4.8% and 9.0% of the maximum rise in RTD 1 temperature for test 1 and test 2. This agreement confirms the accurate modeling of thermal profiles in a bit and formation during a sample acquisition operation. Figure 7 shows the thermal profile of the sub-region of rock closest to the borehole at 480 seconds into test 1. Figure 8 shows the thermal profile of the bit at the same moment.

The maximum rock temperature at 480 seconds is 120 °C, corresponding to a rise of about 100 °C above the ambient temperature. If this temperature had exceeded a threshold established for an actual drilling operation on Mars, a rest period would need to be scheduled prior to the 480 second mark, or the power usage of the drill would need to be restricted.

There is a small amount of disagreement between the simulation and the experimental results, primarily in test 2, and in particular for the RTDs closest

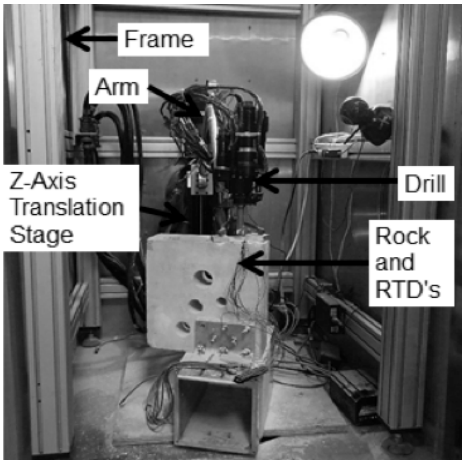


Fig. 4 The experimental setup consists of a vacuum chamber (1 m wide, 1 m deep, 3.5 m tall) containing a translation stage, an arm, a drill, a frame used to mount the translation stage to the chamber, and the limestone block with four embedded RTDs.

to the bit. After about 3,000 seconds, corresponding to a penetration depth of about 7.0 mm, there is an upward kink in the measured temperature histories. Although smaller in magnitude, test 1 shows this feature at the same penetration depth as test 2, occurring at about 3,700 seconds into the test. The kink is most visible in RTDs 1 and 2. No changes in power consumption of penetration rate occurred during this time. The consistent depth at which this feature occurs in the experimental data seems to suggest that the feature is an unmodeled consequence of the bit geometry, since no abnormalities were found in the limestone block.

Although the cutter is v-shaped and not flat as it is assumed to be in the simulation, the entire diameter of the bit reaches the rock quite quickly, so the effects of the v-shaped cutter are assumed to be negligible. Much of this small v-shaped region penetrates the surface during a short "hole start" routine designed to prevent the bit from wandering across the surface of the rock.

It is hypothesized that prior to penetration depths of about 7.0 mm, the bit is more effective at removing cuttings from the borehole than at later times. This may be due to the fact that although the cutter sweeps out the full cross-sectional area of the bit, there are large tapered sections in the bit which provide pathways for cuttings to escape the borehole. These pathways are viewable in Figure 9. An increased cutting removal rate would allow hot, freshly cut particles less time to conduct heat into the formation. The temperature of these cuttings is critical to determine, since these particles may contain quantities of volatile water.

After penetrating past 7.0 mm, cuttings removal is made more difficult as two of the four pathways become blocked (the left one in Figure 9), while the other two become increasingly blocked (the right one in Figure 9). When cuttings can no longer escape quickly, thermal power input to the rock reaches the expected value entered into the simulation. At this point, the simulated and experimental curves are seen to be nearly parallel, although an offset has been created during the first regime of penetration which provides error throughout the duration of the test.

Additionally, gaseous convection to the thin atmosphere may cause a non-negligible reduction in temperatures. At present, convection is not included in the simulation. During early phases of penetration, while heat is being generated by the cutter close to the surface of the rock, convection at the surface of the rock may play a more substantial role than in later phases when the borehole is deeper. This may also explain the reduced rate of temperature increase experienced at depths less than 7.0 mm. Convection could also play a role in the borehole where the bit causes rapid movement of air near the hottest regions of the rock and bit.

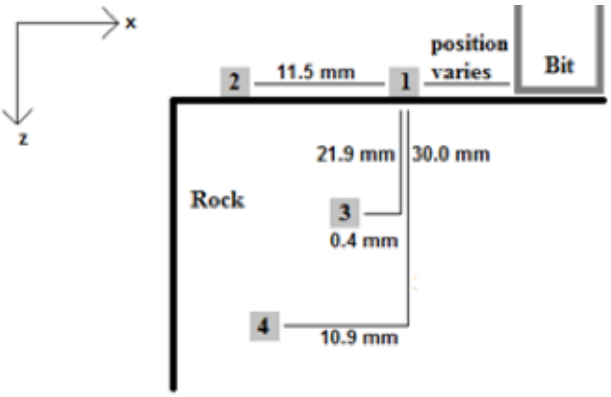


Fig. 3 Using the location of RTD 1 as the origin, the locations of the RTDs are displayed. RTD' are shown in the image as shaded, numbered squares. All RTDs are all placed within the x-z plane so there is no variation in the y-direction, but the position of the bit varies in both x and y between each test. The polar r-axis measures radial distance from the z-axis within the x-y planet.

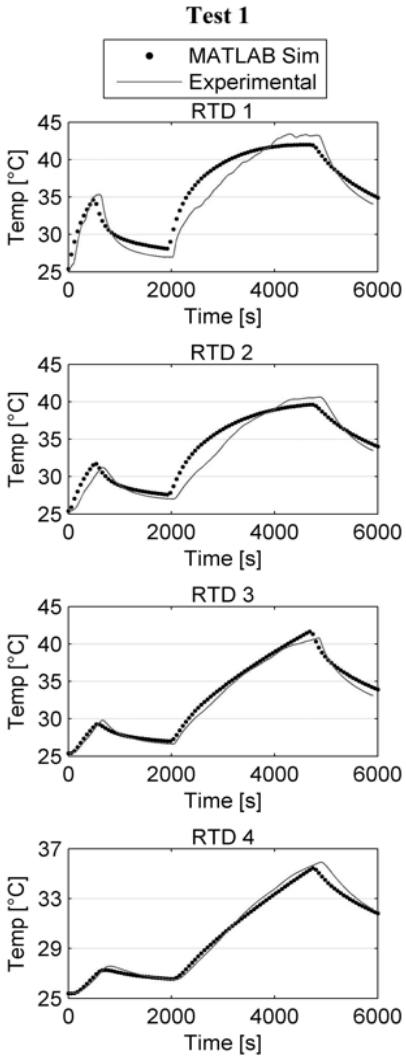


Fig. 5 The experimental data for each RTD in test 1 is plotted along with the model's predictions. Average absolute error is 0.86 °C (about 4.8% of the temperature rise of RTD 1).

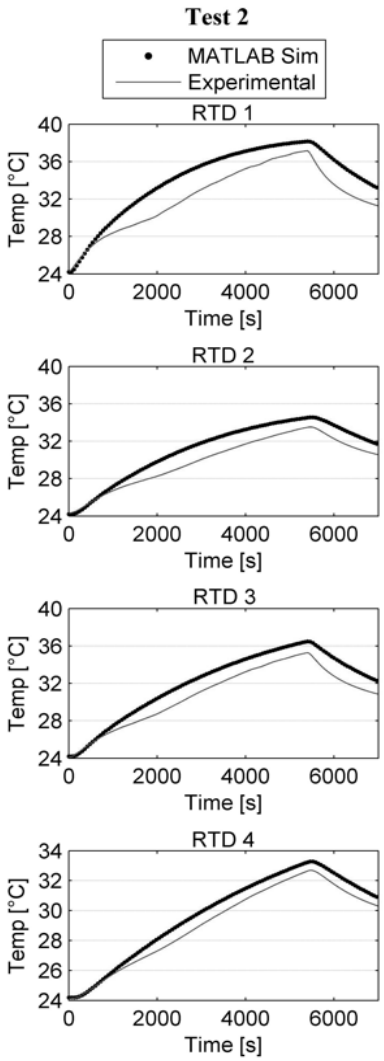


Fig. 6 The experimental data for each RTD in test 2 is plotted along with the model's predictions. Average absolute error is 1.26 °C (about 9.0% of the temperature rise of RTD 1).

V. CONCLUSIONS

The agreement of the results from simulated and experimentally conducted drill tests demonstrates that complete knowledge of the thermal environmental of a rock or soil formation during sample acquisition is possible and reliable, making it a powerful tool for Mars exploration. The ability of a thermal model to regulate power usage and create a schedule of periods of drilling and pausing provides a means for obtaining samples with minimal risk to hardware or scientific validity. Without such a model, achieving these gains would require observation of motor feedback parameters, such as auger current, at which point substantial and irreversible ice deposits may have already formed.

In addition, more recent versions of the simulation have been expanded to allow coring bits to be modeled, since core samples are of particular interest to MSR. Besides the obvious differences in cutter geometry, core sample simulations and experiments vary from full face simulations because the core samples inside the bit are subjected to different boundary conditions than rock elsewhere, so the thermal profile of core rock may differ from the rock studied thus far.

Further work on this research topic includes more advanced modeling on the motion of cuttings and heat transport as powder travels upward and piles on top of the rock surface. Currently, the model does not account for the effects of cuttings once they have left the borehole. In the experiment and in actual Martian drilling operations performed on surfaces level to the ground, hot cuttings are piled on the rock's surface. The model predicts these cuttings to be at high temperatures, as these cuttings are created from the hot rock cut just beneath the bit. These cuttings convect little heat to the atmosphere because of the reduced atmosphere on Mars, and may conduct significant amounts of heat back into the top surface of the rock. However, the cuttings may serve as a heat sink once they have cooled and heat continues to flow from beneath the surface where frictional heating is still continuing. Since tests on Mars may be performed by coring vertical surfaces such as crater walls, the piling characteristics of the cuttings should be allowed to vary as a function of surface orientation.

Future research should also include investigations into increased clearing efficiency at the early stages of penetration, and should bit geometry be the cause of the change in temperature rise as hypothesized, the simulation should be made to account for variations in clearing rate. Additionally, gaseous convection should be included in the model to attempt to more closely model losses of heat at the surface of the rock and in the borehole.

Lastly, future testing should incorporate creation and testing of acquisition schedules. These schedules will be created by the model to prevent threshold temperatures from being reached. These schedules can then be used to control experimental testing and verify that samples can be acquired without creating unsafe temperatures. Long-term test plans include performing schedule-controlled tests on icy formations at Mars ambient temperature and pressure to demonstrate the effectiveness of the simulation in realistic conditions and the necessity of using a thermal model to guide sample acquisition on future missions.

VI. ACKNOWLEDGEMENTS

This research was made possible by the financial support of JPL's Strategic University Research Partnership (SURP). The authors would like to thank Samad Hayati, Chad Edwards, Lori Shiraishi, and Kristo Kriechbaum of the Jet Propulsion Laboratory (JPL) for their input and discussions. The authors would also like to thank Gale Paulsen of Honeybee Robotics for aid in conducting the experimentation.

The author states co-authorship with Timothy Szwarc, Prof. S. Hubbard, Prof. R. Christensen, Prof. Brian Cantwell, Dr. K. Zacny.

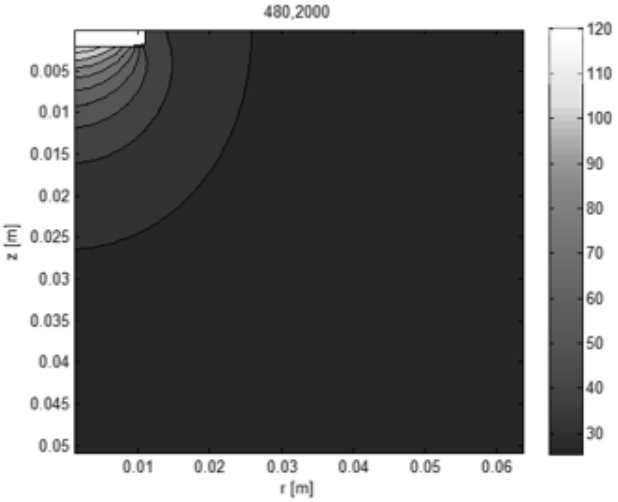


Fig. 7 Thermal profile of the rock in test 1 after 480 seconds (8 minutes) of drilling. Ambient temperature is 25.4 °C. The borehole can be seen at the top left corner of the image represented by a white rectangle. The system is axisymmetric, with the left edge of the image representing the axis of symmetry. Hottest temperatures are shown to be at the bottom center of the borehole in this image.

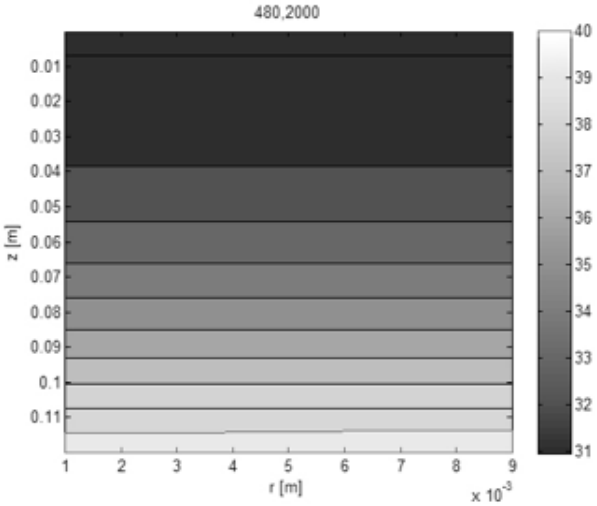


Fig. 8 Thermal profile of the bit in test 1 after 480 seconds (8 minutes) of drilling. Ambient temperature is 25.4 °C. Note that because MATLAB draws contour plots by connecting midpoints instead of drawing boxes for each element, the bit actually spans from 0 to 1 mm along the r-axis. Hottest temperatures are at the bottom of the bit during drilling.

References

1. Haberle RM, McKay CP, Schaeffer J et al. On the Possibility of Liquid Water on Present-day Mars. J Geophys Res, 2001; 106.
2. Feldman WC, Boynton WV, Tokar, RL et al. Global Distribution of Neutrons from Mars: Results from Mars Odyssey. Science 2002; 297.
3. Arvidson RE, Bonitz RG, Robinson ML, et al. Results from the Mars Phoenix Lander Robotic Arm Experiment. J Geophys Res 2009; 114.
4. Zacny K A, Quayle MC, Cooper GA et al. Laboratory drilling under Martian conditions yields unexpected results. J. Geophys. Res., 109, E07S16. 2004; doi: 10.1029/2003JE002203.
5. Raraty LE, Tabor D. The Adhesion and Strength Properties of Ice. Proc R Soc Lond A, Math Phys Sci 1958; pp. 184-201.
6. Backes P, Lindemann R, Collins C et al. An integrated coring and caching concept, Aerospace Conference, 2010 IEEE; 10.1109/AERO.2010.5446985
7. Uhlmann E, Worner E, Brucher M. More efficient cutting process due to the heat-spreading effect of CVD diamond. Industrial Diamond Review 2003; pp. 25-29.
8. Zacny KA, Cooper GA. Friction of drill bits under Martian pressure, J Geophys. Res., 112, E03003 2007; doi:10.1029/2005JE002538.

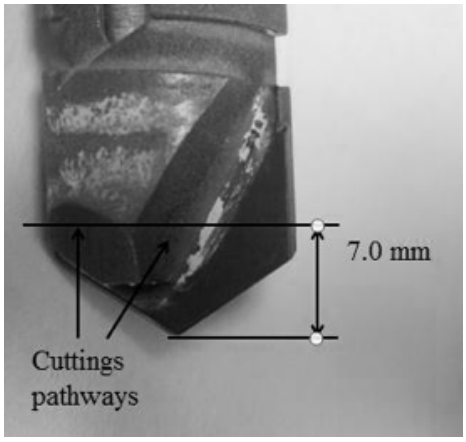


Fig. 9 A close-up image of the bit shows the pathways by which particles escape. Notice that the pathway on the left becomes nearly closed once penetration has reached a depth of 7.0 mm, while the pathway on the right becomes increasingly blocked as penetration increases. The reverse side of the bit has the same two pathways as the two seen here.

Akul Aggarwal is from Fremont, CA. He is majoring in Engineering Science with an emphasis in Mechanical Engineering at UC San Diego. Akul is interested most in the design and fabrication of mechanical parts, especially for cars, planes, submarines, and rockets (essentially, things that move, and move fast). In his spare time he loves to read, swim, and run. He is currently working on the Human Powered Submarine at UCSD. During the winter, he enjoys going to Tahoe to carve some snow.

Coarse head pose estimation using image abstraction

Anant Vidur Puri

We present an algorithm to estimate the pose of a human head from a single image. It builds on the fact that only a limited set of cues are required to estimate human head pose and that most images contain far too many details than what are required for this task. Thus, non-photorealistic rendering is first used to eliminate irrelevant details from the picture and accentuate facial features critical to estimating head pose. The maximum likelihood pose range is then estimated by training a classifier on scaled down abstracted images. This algorithm covers a wide range of head orientations, can be used at various image resolutions, does not need personalized initialization, and is also relatively insensitive to illumination. Moreover, the facts that it performs competitively when compared with other state of the art methods and that it is fast enough to be used in real time systems make it a promising method for coarse head pose estimation.

Fig. 1 Sample Results

I. INTRODUCTION

Head pose estimation is an intriguing and actively addressed problem in computer vision. The reason for this is the application potential of an accurate pose estimation system in Human Computer Interaction, which is one of the most upcoming research areas in recent times. Some of the applications in this field are emotion recognition, unobtrusive customer feedback, biological pose correction, and interactive gaze interfaces. Knowledge of head pose is also extremely useful in a host of other head and face related computer vision applications including surveillance and avatar animation. Highly accurate systems have been proposed, including 3D models,[1] fiducial point fitting,[2] and machine learning (ML)[3,4] techniques. However, most of these methods suffer from one or more drawbacks such as high computational complexity,[2] requirements for huge training sets,[4] tedious alignment issues,[3,4] sensitivity to illumination and facial expressions,[4,1] personalized initialization1, and non-scalability to estimate pose for multiple subjects.[2,1] It is thus with an aim to simplify the problem and to seek an easy, fast, and accurate solution that we propose our algorithm.

Our approach is essentially to seek appropriate features that can be leveraged by statistical learning techniques. The unique part is that it takes advantage of the fact that most images have a level of detail that is much more than what is required to estimate pose. So, we abstract the image using non-photorealistic rendering (NPR).[5] NPR reduces the complexity of analysis while sending the information across in a better and easier form. This approach leads to a computationally light system suitable for real time use and scalable to multiple subjects. Abstraction affords robustness to identity and facial expressions. Also, our approach is relatively insensitive to skin and lighting variations. The obtained features lead to considerably fewer alignment issues than other ML techniques for head pose estimation. The method's accuracy and computational lightness make it ideal as a first step to estimate finer head pose. Fig. 1 illustrates the excellent performance of our algorithm showing the results on certain random images from the Internet.

The rest of the paper is organized as follows. Section II describes image abstraction while section III describes the NPR algorithms used in our system. Sections IV and V focus on the training and testing pipelines, respectively. We present our results in section VI, where we also

Fig. 2 Various forms of image abstraction (left to right): original image, color segmentation image, contours image, skin detector output, cropped upper face.

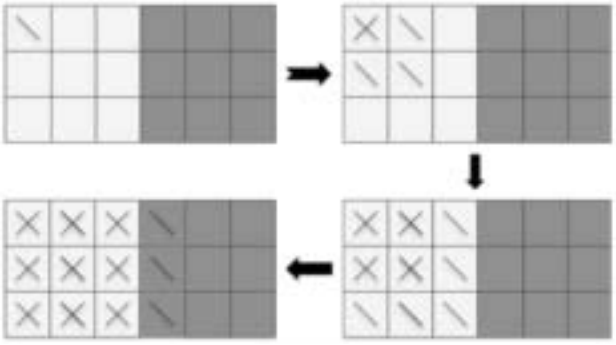


Fig. 3. Region growing using initial seed. The single line denotes pixels being processed in current step. Crossed pixels have been analyzed and are part of current segment.

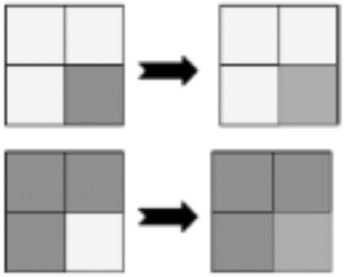


Fig. 4 Refining segments by smoothing out edges

Fig. 5 Output of NPR algorithms (left to right): original image, region segmented image, stroke rendered image.

describe a heuristic to estimate a usable output range of pose. The final section is the conclusion describing future scope and possible modifications to the algorithm.

II. IMAGE ABSTRACTION

The usefulness of abstraction can be best understoodthrough a series of examples shown in Fig. 2. It shows the actual image of a person followed by the same image after applying different forms of abstraction.

All these images contain far less detail than the original image but they all convey the pose of the head. The crucial fact is that our brain often does several tasks for example, shape recognition using much lesser features than what are present in the actual image. A recent study[6] shows how people can interpret shapes from line drawings just as well as from shaded images. Experiments have shown that humans take less time to interpret a scene when shown cartoon drawings rather than real pictures.[7] This suggests that our brain uses only those features which are critically required to perform the given task. Therefore, if a machine learning algorithm can abstract out the useful details of an image prior to analysis, it is possible to increase its efficiency and overall performance by more closely emulating actual brain function.

Abstraction, however, is a subjective term and can be taken to have many meanings. At one extreme it could mean fundamental features like contours, which are the building blocks of computer vision algorithms. At the other extreme it could mean an intelligent graphic rendering of the scene. There is a middle ground, where more useful features can be readily obtained from a scene if suitable image abstraction techniques can be applied. Non-photorealistic rendering (NPR) lies in this middle domain. In the next section, we describe NPR and how it leads to image abstraction. We also present the two techniques that we have tried for our goal of head pose estimation.

III. NON-PHOTOREALISTIC RENDERING (NPR)

Non-photorealistic rendering seeks to give an artistic impression to images. These algorithms automatically gauge information to stylize the image by simulating paintbrush strokes, creating segments of uniform color, introducing a pen and ink illustration effect, etc. Stylizing the image to make it look like a painting has an effect akin to that of abstraction (i.e., preserving necessary information while doing away with unnecessary details). This fact can also be inferred from a simple observation: *most real life paintings lack extensive detail yet manage to draw attention to key features*. Paintings like Van Gogh's *The Potato Eaters* have not been painted with intricate detail. Earlier works[8] have discussed the perceptual impact created by nonphotorealistic rendering and present experiments that illustrate the mutual beneficence of NPR and perception. NPR can thus reduce complexity of a computer vision algorithm by accentuating salient information. To investigate this prospect, we use two simple forms of NPR to simplify and abstract features of the head so as to make pose estimation easier. These are:

1) Regional Segmentation
2) Stroke Rendering

Both methods are explained in the following subsections.

A. Regional Segmentation

The method is essentially color based segmentation: seed growing followed by refinement steps. First, a seed is planted, and a region is grown by adding all neighborhood pixels with RGB values within a certain threshold. The algorithm then recursively calls itself on the neighbors of the last added pixel(s). Whenever the threshold is violated, a new seed is planted, and a new region is grown. This continues until every pixel is a part of some segment. This process is illustrated through Fig. 3.

Refinement of segments happens in two steps. The first step involves smoothing the segment boundaries. This is done by considering four pixels of the initially segmented image at a time. If three out of four pixels belong to one segment then all four are merged into the same segment by filling the fourth pixel with an intensity value that is intermediate between that pixel and the others, as shown in Fig. 4. The second refinement step is repeated segmentation, i.e., segmentation and refinement are done repeatedly until the number of segments becomes constant. This ensures that the image cannot now be divided into more segments with that threshold value. The results of applying our regional segmentation algorithm on four random images are shown in Fig. 5. Unnecessary irregularities in the image are smoothened out and important patches emerge, while unimportant patches are merged as one. Fig. 6 shows similar results for images from the Pointing Database[9] (which is used for training and testing our pose estimation system).

We undertake two steps to make the system robust to illumination and skin color. First, we have used histogram equalization on the original images. Further, we normalize the pixel values p_i of the NPR image, $p_i = (p_i - \mu)/\sigma$ where μ and σ are the mean and standard deviation of that NPR image. It is noticed that various parts of the head are visibly accentuated, to make detection and analysis easier and faster.

Hairlines show up in all cases; eyes also result in distinct blobs. Often these two alone are sufficient to determine pose. The algorithm also detects mouth and nose as distinct. Moreover, hair and skin can show up as different segments. This approach is adaptable to a wide range of people, environmental conditions, and facial expressions.

Fig. 6 NPR results for Pointing Database (top to bottom): original image, region segmented image, stroke rendered image.

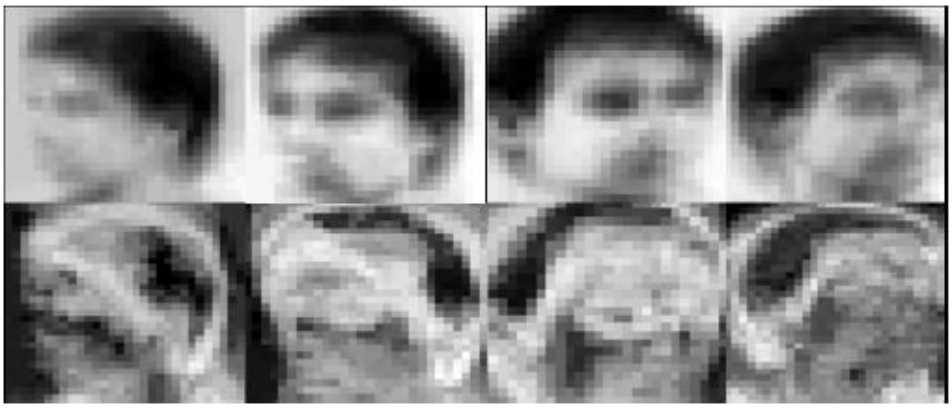


Fig. 7 Mean (top) and variance (bottom) images for different poses for regional segmentation.

These observations are strengthened by the mean and variance images, shown in Fig. 7 (derived during training). Note that the NPR images are converted to gray scale and used by the system during training and testing. We observe that the locally uniform regions in the mean images show that the images are smoothened to uniform color segments. We also get some regions of very low variance, which implies that critical features of the face are strongly captured. Often, shadows are visible on the face, and this may initially seem like a deterrent. On observing the learned model, however, we realize that the shadows and where they lie (which side below the chin, etc.) are themselves cues for pose and are learned by the system. Since only the basic features of the face, which remain more or less same throughout the dataset, are retained, the algorithm learns lesser but more relevant information and is thus able to respond better.

B. Stroke Rendering

The second method involves rendering the image using strokes like in line drawings. First, edges are detected using simple

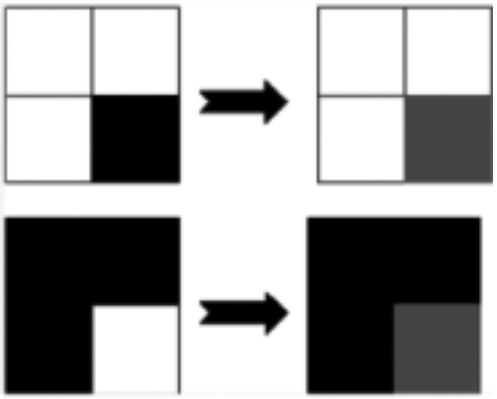


Fig. 8 Edge Refinement

Sobel masks of size 3X3. These edges are then thresholded to retain only those edges which define major segments in the image.These edges are then refined by considering four pixels at a time; if three of them are black (i.e., they form an edge), the fourth one is given half the maximum grayscale value. Similarly if only one is black, it is reassigned half the maximum grayscale value. This is illustrated in Fig. 8. The edges detected by using this method are, however, very thin. It could be difficult to train a system on such thin edges as the amount of data would be too little. So we do a thickening operation on the edges, similar to dilation. The mean and variance images for stroke rendering are shown in Fig. 9. While the mean images do show some clarity in terms of making out the face and pose, there are very few smooth, low variance regions. This could challenge a system to learn a suitable pattern. (Thus it is no surprise that regional segmentation works better, as will be seen in Section VI). As far as computational performance is concerned, these approaches perform remarkably well. A Python implementation is currently able to non-photorealistically render at 3 fps (for both region segmentation and strokes) with image size 320X480 pixels while running on a 2.4 GHz Intel 3 Processor. We note that our algorithms give sufficiently clear NPR results even for head image sizes as low as 96X96. A C implementation would run much faster and could be readily integrated into real time systems.

Fig. 10 Elliptical head tracker + GrabCut (left to right): Elliptical head tracker output, Automatically marking sure foreground and background region guided by the ellipse, GrabCut output.

IV. TRAINING A LEARNING ALGORITHM

A learning-based system is expected to perform well if trained on abstracted images of human heads. We have explored this idea using a naive Bayes classifier. As mentioned before, we have used images from the Pointing Database (30 different images are available for each pose) for training our system. The training procedure is explained below:

(i) Images are cropped according to the boundary of the head and the cropped images are histogram equalized to be robust to illumination effects.

(ii) A non-photorealistic rendering algorithm is applied to generate abstracted images. The same parameters are maintained during training and testing.

(iii) The result is then converted to a gray scale, 32X32 frame. This step helps in normalizing the locations of the abstract regions across various head shapes. It also saves computations during both training and testing.

(iv) A statistical model is learned (Section V) at each pixel i for a given pose j, by computing the mean ij and standard deviation σ_{ij} of pixel values x_{ij} , as shown in equation 1. Here, each pixel is assumed to be statistically independent, which makes the computation easy; at the same time, the learned model is accurate enough for pose estimation. Separate models are trained for both NPR techniques.

$$\mu_{ij} = \sum_i \frac{x_{ij}}{N} \quad \sigma_{ij}^2 = \sum_i \frac{\left(x_{ij} - \mu_{ij}\right)^2}{N} \qquad (1)$$

Currently, the system learns models from frontal, 0°, to 90° in steps of 15° for pan right and left. Also, for certain pan angle ranges, models for tilt up 30º and tilt down 30° are learned.

V. ESTIMATING POSE

In a test image, the head region is first cropped. To the best of our knowledge several machine learning techniques impose a strict requirement on the cropped face. For example, in Huang et al. (2011), the crop is according to the positions of fiducial points. Finding these facial features itself is difficult, leading to a chicken-or-egg problem. Our method is suited for images cropped along the head boundary, which is a much easier requirement on any system. Obtaining this cropped image requires a foreground/ background separation technique. We have used GrabCut[10], available as an OpenCV function, on the output of a head tracker. For tracking, we have implemented a variation of the elliptical head tracker[11], which uses image gradient and color histogram for tracking a head through different poses in a video. A typical result from our tracker is shown in Fig. 10. Definite foreground and background regions are automatically marked for GrabCut, and we get cropped images as shown. Once the image is cropped along the head boundary, the same NPR technique used during training is applied, with the same threshold setting.

Fig. 9 Mean (top) and variance (bottom) images for different poses for stroke rendering

The probability P_{ij} that any pixel i of value x_i belongs to a specific pose j can be found using equation 2.

$$P_{ij} = \frac{1}{\sigma_{ij}\sqrt{2\pi}} e^{\frac{-(x_{ij}-\mu_{ij})^2}{2\sigma_{ij}^2}} \quad (2)$$

$$P_{j*} = \max_j \prod_i P_{ij} \quad (3)$$

The estimated pose j* maximizes the product of probabilities across all pixels according to equation 3. The same result can be achieved by finding the minimum log likelihood sum and thus saving on computation, as shown in equation 4.

$$L_{j*} = \min_j \sum_i \frac{\left(x_{ij}-\mu_{ij}\right)^2}{\sigma_{ij}^2} \quad (4)$$

VI. RESULTS AND DISCUSSIONS

We feel that coarse pose is best expressed as lying in a window, rather than one angle with an unknown associated uncertainty. Thus, the results of the maximum likelihood estimation can be further interpreted as follows:

(i) First, the log likelihood sums for frontal, pan 30°, 60°, and 90° are considered, and the least among these fixes one end of the output pose range.

(ii) Next, the log likelihood sums for pan 15°, 45°, and 75° are considered, and the least among these fixes the other end.

For instance, if pan right 30º had the least log likelihood sum in step (i), and pan right 15° had the least log likelihood sum in step (ii), the pose is classified as between 15° to 30° turned right, i.e., 22.5° ±7.5° pan right.

Given the pan window, we compare the tilt-up and tilt-down log likelihood sums at that pan window to estimate tilt angle. For the window of 0° to 30°, tilt models are built at pan 15°; for the window of 30° to 60°, at pan 45°; and for the window of 60° to 90° at pan 75°. The results we obtain for five-fold cross validation on the Pointing Database are encouraging, as shown in Tables 1 and 2. We show results for both methods regional segmentation and stroke rendering. The results show that regional segmentation works much better, implying that rough regional information corresponding to hair and facial features (e.g., eyes, and mouth) are the more appropriate features for head pose estimation.

More results on random images obtained from the Internet are shown Fig. 1. Note that our system is successful with facial expressions and occlusions. We also present estimated pose on outputs from our tracking + segmentation system in Fig. 11, showing that our pose estimation method can be readily integrated into a simple automatic system. Wu & Toyama (2000)[12] have also used Naive Bayes for estimating head pose. However, their method uses features extracted by three Gabor filters and a Gaussian filter, which is computationally much more expensive than our regional segmentation method, whose complexity is only of the order of the number of image pixels. We present cross validation results for our implementation of their system in Table 3.

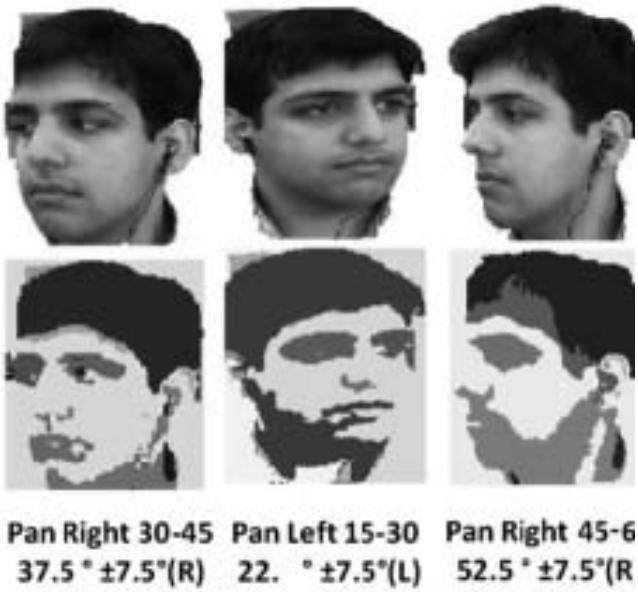


Fig. 11 Results on output of Tracker + GrabCut system

To further prove that an appropriate form of image abstraction is essential for a problem, we present cross-validation results for the simplest possible form of abstraction direct down sampling of the image to size 32X32. Comparing with these results, the better performance of the regional segmentation method is evident.

Table 1. Pan angle estimation.

Pan Results in Correct 15 ° Window							
	0 °-15 °	15 °-30 °	30 °-45 °	45 °-60 °	60 °-75 °	75 °-90 °	Avg
Color Segmentation	96%	96%	100%	92%	86%	82%	96 %
Strokes	76%	72%	84%	60%	92%	88%	79%
Pan Results with 22.5 ° Tolerance							
	0 °-15 °	15 °-30 °	30 °-45 °	45 °-60 °	60 °-75 °	75 °-90 °	Avg
Color Segmentation	100%	100%	100%	98%	94%	92%	99%
Strokes	84%	76%	88%	72%	98%	94%	85%

Table 2. Tilt estimation given pan window.

Tilt Results in Correct 15 ° Window given Pan Range				
	0 °-30 °	30 °-60 °	60 °-90 °	Avg
Color Segmentation	84%	91%	82%	87%
Strokes	71%	77%	78%	75%
Tilt Results with 22.5 ° Tolerance given Pan Range				
	0 °-30 °	30 °-60 °	60 °-90 °	Avg
Color Segmentation	89%	95%	90%	92%
Strokes	82%	86%	88%	85%

Table 3. Baseline Results.

Pan Result in Correct 15 ° Window						
	0 °-15 °	15 °-30 °	30 °-45 °	45 °-60 °	60 °-75 °	75 °-90 °
Filter Bank Approach	90%	85%	71%	52%	51%	45%
Downsampling	85%	80%	80%	71%	70%	65%

VII. CONCLUSION

We have demonstrated how image abstraction can be very useful for rough head pose estimation, as it reduces the level of detail in an image so that only the relevant information is learned by any learning algorithm. We have used simple NPR techniques and believe that a technique with better color segmentation or other suitable forms of stylization will further enhance the results. Sophisticated learning techniques could greatly improve recognition performance and at the same time, benefit from the features presented by image abstraction. Apart from this, we believe that this can benefit fine pose estimation algorithms by providing them with an initial rough pose window to work with. For instance, certain ML-based fine pose methods cannot be used practically because they take a lot of time to converge

to the correct (narrow) pose range. However, if a small correct window can be initially specified using our algorithm, then these methods can be used better.

ACKNOWLEDGMENTS

I would like to thank my mentor, Hariprasad Kannan who continuously inspired me to work harder, my parents who kept pushing me and my sister who always believed in me, Prof. Prem Kalra and Prof. Benjamin Kimia without whom I would have never ventured into this field and Abhinav Parihar for his continual advising. I thank God for giving me the strength to bring this to completion.

References

1. Cai Q, Sankaranarayanan A, Zhang Q et al. Real time head pose tracking from multiple cameras with a generic model. Proceedings of IEEE Computer Vision and Pattern Recognition Workshops (CVPRW), 2010: 25–32.
2. Saragih J, Lucey S, and Cohn J. Face alignment through subspace constrained mean-shifts. Proceedings of 12th IEEE International Conference on Computer Vision, 2009: 1034–1041.
3. Huang D, Storer M, Torre F, and Bischof H. Supervised local subspace learning for continuous head pose estimation. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2011.
4. Fu Y and Huang TS. Graph embedded analysis for head pose estimation. Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, 2006.
5. Strothotte T and Schlechtweg S. Non-photorealistic computer graphics: Modeling, rendering, and animation. Morgan Kaufmann, 2002.
6. Cole F, Sanik K, DeCarlo D et al. How well do line drawings depict shape? ACM Transactions on Graphics (Proc. SIGGRAPH), 2009: 183–188.
7. Ryan TA and Schwartz CB. Speed of perception as a function of mode of representation. American Journal of Psychology, 1956; 69:60–69.
8. Winnemoller H, Feng D, Gooch B et al. Using NPR to evaluate perceptual shape cues in dynamic environments. Proceedings of the 5th international symposium on Non-photorealistic animation and rendering, 2007.
9. Gourier N, Hall D, and Crowley JL. Estimating face orientation from robust detection of salient facial features. Proceedings of Pointing , ICPR, International Workshop on Visual Observation of Deictic Gestures, 2004.
10. Rother C, Kolmogorov V, and Blake A. Grabcut: interactive foreground extraction using iterated graph cuts. ACM Transactions on Graphics (Proc. SIGGRAPH), 2004; 23(3):309–314.
11. Birchfield S. Elliptical head tracking using intensity gradients and color histograms. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1998: 232 –237.
12. Wu Y and Toyama K. Wide-range, person and illumination insensitive head orientation estimation. Proceedings of the Fourth IEEE International Conference on Automatic Face and Gesture Recognition, 2000: 183–188.

Anant Vidur Puri, 21, a senior in Electrical Engineering at the Indian Institute of Technology, New Delhi, India, hails from the city of Chandigarh. Apart from being interested in Computer Vision and Image Processing, he is also a national champion in roller skating and an active debater and thespian, having participated in and won several national level parliamentary debates and stage plays. He also loves trying out different cuisines and reading fantasy fiction novels.

Politics, Twitter, and information discovery: Using content and link structures to cluster users based on issue framing

Moritz Sudhof

This paper explores the use of content and link structures to cluster Twitter users based upon how they frame an issue. We present a model that can be used to cluster users relative to a given topic. This model includes many free parameters whose tuning is explored in a case study of tweets relating to the 2011 Wisconsin union protests. In the case study, we find that content and different link types (hashtags and mentions) can be combined to cluster users based upon 1) their general attitude towards the topic and 2) the medium through which the attitudes are expressed.

1 Introduction

The barrier to becoming a content creator on the Web continues to decline, and the amount of content pertinent to any single issue continues to increase. Correspondingly, tools that can help us understand content are becoming increasingly crucial. Previous work has focused on understanding the temporal nature of content creation[1] and the manner in which this content spreads[2]. This research seeks to complement such temporal studies by endeavoring to understand the content based upon its approach to a topic.

Given a corpus of documents that all engage with a single topic, this research attempts to meaningfully cluster these documents based upon the style of their engagement with the topic. More specifically, this paper explores methods for leveraging the content and the link structures of tweets to build a model that clusters users.

2 Data

We begin by discussing the properties of Twitter data in general and the data used for this research in particular.

2.1 Twitter

Twitter is a microblogging service that is both a social network and a massive source of real-time information[3]. Twitter users communicate with their "followers" by posting messages, called tweets, that are 140 characters or less. Users can "mention" other users by including a user's username preceded by an "@," and users can tag their messages through hashtags, which are simply words preceded by a "#."

Several features make Twitter a convenient medium to focus on for this research. Twitter is a popular medium (the website serves about 140 million tweets per day), and it is relatively simple to download tweets, either through the API or through public archives. Hashtags simplify the task of bundling tweets that engage with the same topic, making it relatively easy to create a sufficiently large corpus of tweets that all discuss the same topic. Furthermore, Twitter is a personality-driven, informal source of information. This ensures that tweets are often opinionated and convey information using colorful language.

Finally, the mentions (@user) and hashtags (#hashtags) in Twitter data add internal structure. Mentions link tweets to users, and hashtags link tweets to general concepts. Both constructs are functionally important: mentions ensure that a specific user is included in or addressed by a commentary, and hashtags ensure that a tweet is grouped with all other tweets with that hashtag. More interestingly, though, they are also often highly performative – directed not at the person mentioned but at the audience instead, for purposes of, for example, "calling someone out." This usage makes them potentially useful in defining the topography of the dialogue.

A mention in a tweet is an indication that the tweet is relevant to the mentioned user, and multiple mentions in a tweet are an indication, or a vote, that each mentioned user is relevant to the tweet and thus to each other. Therefore, the link structure of all tweets is comparable to crowd-sourced votes (where the crowd is the corpus of tweets) regarding which users operate in similar corners of the topic space.

The purpose of hashtags, on the other hand, is ever-evolving and more difficult to pinpoint. Tags on Twitter are unique because of their conversational, not just organizational, nature. Studies have found that hashtags are generally used to filter and direct content so it appears in certain streams[4], and increasingly, those streams have been not purely topical but also coded with derision, irony, humor, and love, to name a few. In cases of rich and heavy hashtag usage, tags can be used to summarize the topical and emotive bent of the tweets.

2.2 The Corpus

The corpus used for this research consists of tweets discussing Gov. Walker's proposal to remove the collective bargaining rights of public employees and the protests that followed. The corpus was built by archiving tweets including a *#wiunion* hashtag and downloaded from TwapperKeeper, an online Twitter archiving service, for this research. Although this hashtag archive does not include all tweets related to the Wisconsin union controversy, the *#wiunion* hashtag was used by a wide array of Twitter users with varying opinions to indicate that a tweet was relevant to the discussion. The corpus consists of 219,528 tweets authored by 36,119 users between February 17, 2011 and March 18, 2011.

We chose a political controversy as the basis for our corpus because the political space is ideal for this research. First, it is a space reliably marked by a variety of different, conflicting perspectives, regardless of which topic filter is applied. Furthermore, opinions tend to be resolutely held and forcefully defended, resulting in strong language. Finally, in politics, the way something is said is often more important than what is being said. That is, political rhetoric emphasizes the presentation of information, and political actors often deliberately repeat certain words and phrases in order to define the framing of an issue. Politics is therefore the perfect space to first attempt to measure similarity in the approach to an issue.

3 The Model

Although certain aspects of the model are tailored to the Twitter space and the problem of identifying users to follow, the model was developed with consideration for its ability to be generalized to any set of documents with a link structure.

The model will cluster users relative to a given topic, which means a general corpus must be clustered by topic before the model can cluster users in a topic space. Effectively clustering a corpus by topic is not a problem addressed by this paper. Instead, topic-based corpuses were developed by choosing an issue, manually selecting the top descriptive hashtag(s) Twitter users employ when discussing the issue, and grouping all tweets with those hashtag(s).

3.1 User Extraction

Given a corpus of tweets that all address a similar topic (e.g., all tweets that share a certain hashtag), the model first identifies a set of users to compare. This is a necessary filter in the Twitter space: there are many users contributing to the dialogue on any given topic, but most users are sparse contributors to the discussion in terms of volume, not particularly reliable and influential actors in the space, or both. Extraction of a set of users seeks to ensure that users that are clustered contribute substantially to the dia logue in terms of both volume and influence.

Ensuring extracted users contribute substantially in terms of volume is a relatively simple problem, and to solve it, this model simply considers only those users that contribute more than a certain threshold of tweets to the topic corpus (40 tweets was deemed an appropriate barrier for consideration for this application). Measuring influence on Twitter is a more difficult problem that has been studied extensively by Weng and Leavitt et al., among others. The details and controversial aspects of different influence measures are beyond the scope of this paper. It is clear, however, that user influence, or "clout," generally holds across topic domains[5]. Furthermore, since we only seek to exclude from consideration those users who have little or no influence, exact influence rankings of the top users are not necessary.

We are therefore comfortable with leveraging measures of the relative influence of Twitter users available through numerous online services (this model utilizes Twitalyzer.com) to quantify a user's "clout."

3.2 Bundling

After a set of interesting users has been defined, the model bundles all tweets authored by a single user into a single document. Documents are then tokenized in a manner sensitive to the vocabulary of Twitter (paying attention to, for example, hashtags, mentions, and links) and transformed into "bags of words." In order to preserve the richness of the language, documents are tokenized in a manner sensitive to the Twitter medium (e.g., with consideration for the preservation of emoticons), and stopwords filtering and term stemming are not employed. After tokenization, each bag of words expresses, in sum, a user's engagement with the corpus topic.

3.3 Content Analysis

The next step is to measure similarity between documents by using vector space models (VSMs)[6].

Given a corpus vocabulary of size n, we can represent the bag of words of document i as an n-dimensional vector v_i, where each dimension represents one term in the vocabulary. The value of dimension j of document vector v_i, or $v_{i,j}$ then depends on the "importance" of term j in document i, where "importance" is quantified by a weighting mechanism. By representing documents as vectors in a Euclidean space, VSMs allow us to measure semantic similarity by leveraging Euclidean or non-Euclidean distance measures.

The most common method for weighting term vectors for information retrieval is by *tf-idf*[7]. *tf-idf* weighting formalizes the intuition that a rare term has higher information content than an expected term. More precisely, it scales a term frequency (*tf*) by the log of the term's inverse document frequency (*idf*). Under this weighting mechanism, given a document i and a term j,

$$v_{i,j} = \frac{y_{i,j}}{\sum_k y_{i,k}} \cdot \log \frac{|D|}{|i : j \in d_i|}$$

where $y_{i,j}$ is the number of occurrences of term j in document i, $|D|$ is the number of documents in the corpus, and $|i : j \epsilon d_i|$ is the number of documents in which the term i appears.

Although the model weights terms by *tf-idf*, we also weight terms by their odds, which is a measure for how significant a term is for a particular author. Given a document i, the odds of term j appearing is simply

$$O_{i,j} = \frac{y_{i,j}}{n_i - y_{i,j}}$$

where $y_{i,j}$ is as above and $n_i = \Sigma_k\, y_{i,k}$. This measure has produced mixed results in emphasizing distinctive, emotive terms[8], and we improve its performance by first filtering statistically insignificant terms by applying the G-test (also known as the loglikelihood ratio test) and filtering using the test's p-values.

The G-test compares observed frequencies to an expected frequency to determine whether the difference in observed frequencies is significant[9]. In general, given term j and a set of documents D, the G-value is computed as follows:

$$G = 2 \sum_{i \in D} O_{ij} \cdot ln(\frac{O_{ij}}{E_{ij}})$$

where O_{ij} is the observed frequency of term j in document i and E_{ij} is the expected frequency. For this application,

$$E_{ij} = \frac{\sum_{i \in D} O_{ij}}{|D|}$$

If the G-value is above a threshold value (which depends on the degrees of freedom of the data, or the number of documents being compared), the term is statistically significant. The G-test has been shown to be effective in several computational linguistics applications[10], and the model applies it to document vector pairs to filter insignificant terms before measuring similarity. After document vectors are weighted, the cosine similarity measure, which has proven a reliable similarity metric for textual analysis[11], is used to quantify similarity.

Given document vectors $\mathbf{x}$ and $\mathbf{y}$ with n dimensions, the cosine similarity is defined as

$$\begin{aligned} cos(\mathbf{x}, \mathbf{y}) &= \frac{\mathbf{x} \cdot \mathbf{y}}{\| \mathbf{x} \| \| \mathbf{y} \|} \\ &= \frac{\sum_{i=1}^{n} \mathbf{x}_i \cdot \mathbf{y}_i}{\sqrt{\sum_{i=1}^{n} \mathbf{x}_i^2 \cdot \sum_{i=1}^{n} \mathbf{y}_i^2}} \end{aligned}$$

The intuition behind the cosine similarity measure is that it is not the Euclidean distance between two vectors but the angle between them that defines

their similarity. This makes it a fitting similarity measure for documents. The difference between similarity measures that normalize for vector magnitude, though, is commonly thought to be insignificant in information retrieval applications[12], and we chose cosine similarity as it is a familiar and popular metric for natural language applications.

The fact that all documents address the same general content and are authored by often colorful personalities yields two interesting results: there are many content terms that are shared by all documents, and each document has idiosyncrasies that are more indicative of the author's writing style than the author's approach to the issue. In the case study in section 4, we will examine whether *tf-idf* or odds weighting performs better for this application.

3.4 Link Analysis

Parallel to the VSM analysis, the model leverages the link structures to measure link-based similarity between users.

3.4.1 Mentions

The mentions link structure is essentially a bipartite graph with tweets as nodes on one side linking to nodes of users on the other side. We use the SimRank algorithm, a measure of structural-context similarity, to measure similarity between nodes. Consider two distinct nodes a and b where *I(a)* and *I(b)* are the set of in-neighbors of *a* and *b*. SimRank is a recursive algorithm that defines the similarity *s(a, b)* as

$$s(a,b) = \frac{C}{|I(a)||I(b)|} \sum_{i=1}^{|I(a)|} \sum_{i=1}^{|I(b)|} s(I_i(a), I_j(b))$$

where C is a decay constant between 0 and 1 [13]. Similarities are measured by initializing selfsimilarities to 1 and all other similarities to 0, then iteratively computing similarities until they converge.

Consistent with our bundling of all of a user's tweets into a single document, we also condense the graph by grouping the source tweets by author and weighting an edge from source user *s* to destination user *d* by the number of times *s* mentions *d* divided by the total number of mentions by user *s*. To account for edge weighting, we now define similarity as

$$s(a,b) = \frac{C}{W(a)W(b)} \sum_{i=1}^{|I(a)|} \sum_{j=1}^{|I(b)|} W_{(I_i(a),a)} W_{(I_j(b),b)} s(I_i(a), I_j(b)) \quad (4)$$

where $W_{(I_i(a),a)}$ is the weight of the edge from node $I_i(a)$ to node a.

We are only interested in the similarity scores of a subset of all users in the topic space, but the recursive structure of the algorithm requires us to measure every relevant pair-wise similarity in order to measure the similarity between just one pair of nodes in a strongly connected graph. Furthermore, the space needed to store the similarities between all pairs of authors for large corpuses can quickly exceed the amount of main memory on most machines. Therefore, we use the Random Surfer Pair model to measure similarity between users[14]. Intuitively, random walks measure similarity between nodes *a* and *b* by measuring how long it takes for two random walkers, one starting at *a* and one at *b*, to meet if they walk in lock-step. Experimentally, as few as 1000 walks have been shown to reasonably approximate similarity between two nodes. This algorithm is generalizable to any directed graph.

3.4.2 Hashtags

Link analysis of the hashtag link structure is directly comparable to the mentions link structure analysis except that with the hashtag bipartite graph, we are interested in similarity between source nodes, not destination nodes. However, since hashtags are not links as much as indicators of topical and emotive alignment, modeling them as links may not be the most appropriate analysis. Conceptualizing hashtags as the set of all of the topical and emotive spaces a user inhabits suggests using a different similarity metric: the Jaccard similarity coefficient.

Intuitively, the Jaccard similarity coefficient measures the similarity between two sets based upon how much the two sets *do* overlap compared to how much the two sets *could* overlap. Formally, the Jaccard similarity of sets *A* and *B* is measured as

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

Applied to two sets of hashtags, this similarity measure quantifies the emotive and topical overlap of two users.

As a set-based similarity measure, the Jaccard similarity coefficient does not discriminate between set elements; all hashtags are created equal. Users use some hashtags much more than others, however, and this difference in use is significant. We alter the above formula slightly to account for the relative importance of hashtags. Given a hashtag *i* and a user *a*, let us define the weight w_{ai} as

$$w_{ai} = \frac{f_{ai}}{\sum_j f_{aj}}$$

where f_{aj} is the frequency of hashtag *j* in all tweets authored by user *a*. Weights are therefore simply normalized frequencies (and the weights for all tags for any given user sum to 1). We can therefore define weighted Jaccard similarity coefficient as

$$J_w(A,B) = \frac{1}{2} \sum_{i \in |A \cap B|} min(w_{ai}, w_{bi})$$

The case study below will analyze the difference between clustering based upon SimRank similarities and weighted Jaccard similarities in an attempt to understand the significance of hashtags.

3.5 Clustering

Finally, hierarchical agglomerative clustering (HAC) is used to cluster the documents. HAC views each element as a singleton cluster at the outset and clusters elements by successively merging clusters until all elements belong to the same cluster. At each iteration, the clusters to merge are chosen by identifying clusters that are "closest," where distance between clusters can be defined as the shortest distance, the longest distance, or the average distance between the clusters, among others. Experimentally, we did not find significant differences in clustering based upon the type of distance measure chosen, and this paper will not analyze their tradeoffs.

HAC provides several attractive features for this application. The dendrograms produced by HAC are ideal for visual evaluation, as the distance between clusters and cluster quality are visually apparent, as is the sequence of cluster mergers. Furthermore, HAC allows us to combine similarity measures. Since HAC does not require the ability to compute centroids of document sets in order to cluster documents, we can easily combine content-based similarities and link-based similarities before clustering, where link-based clustering can be based on either mentions or hashtags.

We combine similarities derived from different analyses by assigning a

Approach	Example tweet
front-line support, solidarity	"We are ONE! Show solidarity with teachers fire fighters nurses & all under attack in the states. Text ONE to 235246. #StateSOS #WIunion" –aflcio
protestors as heroes	"'The test of a true good leader is putting yourself on the front line. Thank you.' Thank the #wiunion 14 senators http://on.fb.me/eH0RtV" –aflcio
criticism of Gov. Walker	"RT @Nanbp: @GovWalker #fail MRT @evale72:While telling working families need 2"sacrifice"Walker outsourced jobs 2 India #WIunion" –ivanroberson
focus on Koch brothers	"Hey Walker How Much Money Do The Koch Bros have to pay you to stomp on the good People of America? HOW MUCH? #WIUnion" –supermanhotmale
criticism of media	"Nominate for Best Actor in a Supporting Role: #WIunion coverage by USTream and http://qik.com/Brandzel Brandzel #Oscars" –gottalaff
class struggle	"Republicans dont hate Entitlements They Just hate When Poor people Get Them #KOCH #CorporateWelfare #WIunion #RecallWalker #P2 #sgp #tlot" –sunshineejc
teachers as victims	"Dear #WI Kids Your Teachers Got Fired b/c They Didn't Want a Pay Cut to Give Tax Cuts to the Corporation That Laid Off Your Dad #WIunion" –thenewdeal
teachers as villains	"The protesting teachers of #WI are teaching kids to LIE abandon responsibility & hold others HOSTAGE #wi #wiunion #FIRETHEMALL" –conservativeind
demonization of unions	"#BigLabor 's Legacy of #VIOLENCE http://bit.ly/brDuEr #twisters #tcot #wi #oh #wiunion #ohunion #ocra #p2" –conservativeind
protesters as socialists	"it's hilarious how these socialists think they have a right to an unlimited supply of *your* money. #tcot #wiunion #p2" –brooksbayne
unions as irresponsible, greedy	"RT @lheal: I'm showing #solidarity w/ #wiunion teachers by ignoring my kids until someone pays for my health care & retirement. #p2 #wi" –conservativeind

Table 1 Examples of approaches to the controversy with accompanying example tweets. All tweets are either clearly pro-union or pro-Gov. Walker, but there is a rich variety of ways this support is expressed.

weight w to one of the similarity measures, s_1, to reflect its "value" relative to the other similarity measure, s_2. The combined similarity is then simply the weighted sum of both similarity measures, or $ws_1 + (1 - w)s_2$. Since we do not have labeled data for this task, we will determine the weight w, or relative "value" of a similarity measure, heuristically. That is, we let w be a range of values between 0 and 1 and choose the value that maximizes the clustering coefficient of the resulting combined similarity. Intuitively, the clustering coefficient measures the degree to which nodes in a graph are highly interconnected. A high clustering coefficient indicates that there are more clusters and groups than in a purely random network. We measure a clustering coefficient by creating a graph with all documents as nodes and only the greater half of pair-wise similarities as edges, and measuring the fraction of all possible triangles of the graph that exist. We evaluate the effect of combining content- and link-based similarities in the case study.

4 Case Study

We now evaluate the model by applying it to the Wisconsin union corpus.

4.1 Political Background

On Feb. 11, 2011, Wisconsin Governor Scott Walker proposed removing most public employee collective bargaining rights and tightening the requirements for union certification as part of an effort to close a projected $3.6 billion budget deficit. The legislation also proposed requiring state employees to contribute 5.8% of their salaries to pension costs and to pay 12% of their own health insurance premiums. Although Democrats and unions agreed to require state employees to contribute more towards pensions and premiums, they ferociously opposed the new union restrictions, especially the removal of collective bargaining rights. For overviews of the timeline of events, see references[15,16,17].

4.1.1 The Debate

The controversy surrounding the bill soon transcended both budget politics and Wisconsin state borders. By Democrats and union supporters the bill was seen as a union-busting measure meant to strip workers of basic rights, and by Republicans and supporters of Gov. Walker, it was viewed as a costsaving measure that proposed only modest and fair changes for public employees.

To Democrats, Scott Walker was using the budget deficit as an excuse to push an anti-labor agenda. They viewed the bill as a dangerous step towards seriously curtailing the rights of workers in favor of businesses. Supporters of Gov. Walker, conversely, argued that union supporters were overreacting and represented a dangerous culture of entitlement often referred to as indicative of a "nanny state."

4.1.2 The Role of Social Media

Although it is difficult to quantify the effect of social media on the debate, it is clear that social media played an integral role in shaping, not just commenting on, the controversy. After leaving the state to block a vote, Democratic state senators communicated with supporters at home primarily via Twitter, and Gov. Walker used his personal Twitter account (@scottwalker) to defend his proposals. Finally, Twitter and Facebook were the primary tools used to organize protests and disseminate information rapidly[18].

4.1.3 Note on the Corpus

On Twitter, many hashtags were used to tag tweets dedicated to the Wisconsin protests, but none were used as prevalently as *#wiunion*. This hashtag was introduced on Feb. 11 by Kristian Knutsen of Isthmus: The Daily Page, an alt-weekly based in Madison, Wisconsin. According to The Daily Page, the tag was meant to accompany any tweets referencing the showdown[19]. During the protests, *#wiunion* was consistently a trending topic nationally and even internationally for a brief time.

Since the corpus consists only of those tweets tagged with *#wiunion*, it does not encompass the entirety of the conversation regarding the protests in Wisconsin. Furthermore, by volume of tweets, the corpus is biased in favor union supporters. This dynamic is understandable

User	Alignment	Notes
aflcio	unions	Relays a lot of "on the ground" information about the protests. Frames unions as heroes. Relays a sense of worker solidarity, excitement regarding the protests, and defiance of Gov. Walker.
brooksbayne	Gov. Walker	Frames unions as greedy and their demands as unfounded. Tone is often caustic and mocking. Uses Marxist language when describing union supporters.
conservativeind	Gov. Walker	Relays similar sentiments to brooksbayne, but does so in a combative, not mocking, way, and without Marxist language. Makes the case that unions do not deserve the bargaining rights.
diggrbiii	Gov. Walker	More focus on Gov. Walker than other union opponents, rhetoric more toned-down.
gottalaff	unions	More commentary on media behavior than other pro-union users, mostly to argue unions are not receiving fair coverage.
ivanroberson	unions	Exclusively retweets. Amalgamation of many styles of union support rhetoric, focusing on relaying front-line comments and criticizing Gov. Walker.
sunshineejc	unions	Frames controversy as class conflict, pitting rich republicans vs. middle and lower class workers.
supermanhotmale	unions	Similar approach to sunshineejc, but anger directed at Koch brothers and Gov. Walker specifically, interspersed with exclamations of support for union protesters.

Table 2 Qualitative assessment of user's tweets

given that the legislation threatened to curtail union rights, and the defiance and accompanying energy began with prounion voices, with anti-union voices arising only in response.

This corpus is well-suited for this case study, however, because tweets in the corpus are at once binary and multi-dimensional. The issue is very polarizing, and users using the *#wiunion* hashtag come out either firmly in favor of the unions or firmly opposed. This rigid polarity makes a qualitative assessment easier as it reduces the dimensionality of the discussion: on a basic level, users can be determined to be either in support of the unions or Gov. Walker.

The corpus is still multi-dimensional, however, in terms of the many different ways users convey their support or opposition. By highlighting the nobility of the protesters, for example, a user voices support for the unions differently from the user that primarily operates by comparing Gov. Walker to a fascist. (For examples of different approaches to the controversy represented in the corpus, see Table 1.) Therefore, this corpus allows the case study to test the model's effectiveness on a corpus with a rich variety of approaches to an issue while also simplifying
a qualitative assessment.

For this case study, we will focus on a set of eight users. This limited focus ensures that we can retain oversight during the case study, aiding our qualitative assessment. After narrowing the user base of the corpus to only those users that authored at least 40 tweets in the corpus, we extract a set of eight users to analyze more closely by choosing the users with the highest influence ratings. We also manually ensure that this set of users includes supporters of both the unions and Gov. Walker and expresses that support in a variety of ways. (For a qualitative summary of the users' tweets in the corpus, see Table 2.)

4.2 Term Vector Weighting and the G-test

By definition, *tf-idf* weighting emphasizes the terms that are rarest across the corpus, and odds weighting emphasizes the terms that are used most often by a user. *tf-idf* weighting takes the whole corpus into account, whereas odds weighting only considers a single document. Intuitively, both try to capture a user's distinct approach: *tf-idf* emphasizes the unusual terms a user uses, thus highlighting the terms that distinguish a user from the rest of the corpus (note, of course, that a high frequency term with a low inverse document frequency could still be emphasized by *tf-idf*), and odds weighting emphasizes the terms used most often by the user.

To see which weighting mechanism best expresses a user's engagement with an issue in practice, we compare weighted vectors qualitatively. Given authors a and b and weighted vectors v_a and v_b, the difference vector $v_d = v_a - v_b$ can be used to identify the features that distinguish the two authors given the weighting chosen. Terms with high v_d values are those that most distinguish a from b, and terms with low vd values are those that most distinguish b from a.

Consider users *aflcio* and *diggrbiii*. We will compare these two users because they are on opposite ends of the political spectrum. Comparing them therefore highlights the general distinctive characteristics of *aflcio* and *diggrbiii*, not just those characteristics that distinguish them (this intuition was confirmed by

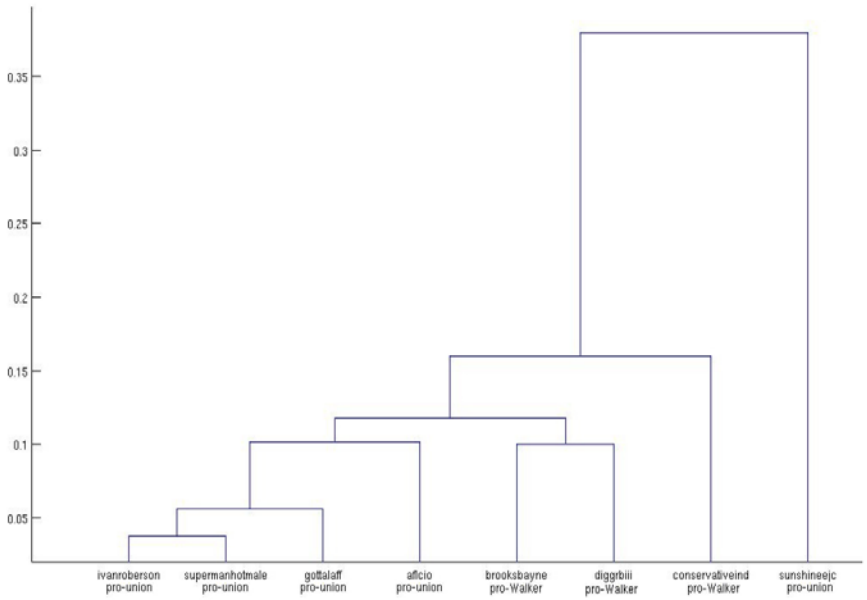


Fig. 1 Odds-weighted cosine similarity-based clustering, with G-test

examining other comparison vectors). Table 3 displays the most distinctive features for both *aflcio* and *diggrbiii* for both weighting mechanisms.

A quick perusal of the table shows, however, that neither weighting mechanism is very effective. Odds weighting, with its emphasis on words used frequently by an author, results in the elevation of many function words such as *on*, *from*, *with*, *the*, and *that*, which, though relevant to authorship attribution, tell us very little about the author's engagement with an issue. In contrast, *tf-idf* weighting, with its emphasis on very rare words, results in the elevation of terms that are used only once or twice and cannot be used as reliable indicators of a user's approach, such as *eaglery* and *chyron*. This problem is the motivation for the application of the G-test. Applying the G-test to a pair of users filters all features such that only those with a G-value greater than 3.8, which corresponds to a p-value of less than 0.05, remain. Table 4 displays the most distinctive features after application of the G-test. (The number of distinctive features is smaller as many terms are removed from consideration by the G-test.)

The quality of both sets of distinctive terms is improved by application of the G-test, but odds weighting seems more effective at emphasizing the terms that express the author's approach to the issue. Qualitatively, tweets authored by *aflcio* generally fall into three categories: those that express defiance at Gov. Walker's actions, those that express support for protesters and laborers, and those that relay information about on-the-ground events. All of these categories are represented by sets of the distinctive terms: *#statesos*, *#notmywi*, and *#standupoh express the defiance*, *#wearewi*, *solidarity*, *workers*, *rally*, and *labor* express the support of the protesters and workers, and *now*, *today*, and *capitol* are terms used when relaying information about current on-the-ground events. *diggrbiii*'s distinctive odds-weighted terms match the general sentiment of the tweets less facilely, but some differences between *aflcio* and *diggrbiii* are notable. Instead of *now* and *today*, *diggrbiii*'s term list includes was, which, in conjunction with *#tcot* (top conservatives on Twitter), *#tlot* (top liberals on Twitter), and *#debunked*, accurately reflects the more after-the-fact, armchair commentary *diggrbiii* engages in. (Note: *#wiunion* is included as a distinctive feature of *diggrbiii* because each tweet includes the hashtag and *aflcio* authored so many more tweets than *diggrbiii* to make the difference in occurrences statistically significant.)

The distinct terms resulting from a *tf-idf* weighting do not reflect the general body of tweets as well. For example, the terms *vp* and *biden* are included by aflcio as quote-attribution, but it is the sentiment of the quote, not who said it, that tells us about *aflcio*'s approach to the discussion. Since very few other Twitter users mentioned Vice President Biden in tweets relating to Wisconsin, though, *tf-idf* weighting makes these terms "characteristic" to *aflcio*'s content. Furthermore, although the G-test filtered many of the rare hashtags and mentions included by *aflcio*, some were not filtered and were again elevated to being very "characteristic" by their rarity.

4.3 Content-based Clustering

We are confident that the combination of the G-test and odds weighting outperforms *tf-idf* in terms of capturing a user's emotive approach to an issue, but the clustering based upon the cosine similarity of odds weighted vectors does not exactly match our intuition (see Figure 5). It is, however, promising. Ignoring the clear outlier, *sunshineejc*, the clustering successively merges first the pro-union users before adding the pro-Walker duo *brooksbayne* and *diggrbiii* and merging with the final pro-Walker user, *conservativeind*. There are not two distinct clusters, but the successive agglomerations generally matches our intuition.

The severity of *sunshineejc*'s outlier status, however, is disturbing. Qualitatively, *sunshineejc*'s approach to the debate is not as radically different from the approaches of all other users as this clustering suggests. *sunshineejc*'s tweets frame Republicans as defenders of rich people aiming to decimate the working class. The rhetoric is much more vitriolic than that of, for example, *supermanhotmale*, but the two users are certainly similar. An examination of the odds-weighted term vectors reveals that the terms that most distinguish *sunshineejc* from all other users include terms such as *hate*, *poor*, *#corporatewelfare*, *entitlements*, *#koch*, and *broke*. These terms suggest that *sunshineejc*'s singular focus on class distinctions and class anger does not fit into the framework provided by the other users. To human readers, *sunshineejc*'s attitude is clearly similar to that of, for example, *supermanhotmale*, but *sunshineejc*'s vitriolic rhetoric and singular class focus radically differentiate it from all other users simply because no other users use comparable rhetoric. Although this case reflects a limitation of this content-based approach, the limitation is exacerbated dramatically by the fact that none of the other seven users focus on the class dynamics of the issue at a level comparable to *sunshineejc*, which naturally means *sunshineejc* is an outlier.

4.4 Hashtags

Hashtag-based clustering yields dramatically different results depending upon whether SimRank or the weighted

Jaccard similarity coefficient is used. Using our notions of support for or opposition to unions as the reference for successful clustering, weighted Jaccard similarity clustering outperforms SimRank clustering. In fact, the Jaccard clustering matches the expected clustering exactly except for *sunshineejc*, which is again a clear outlier. This outlier status is attributable to *sunshineejc*'s liberal use of hashtags. Each tweet from *sunshineejc* includes at least four hashtags, and some include up to eight. The other users, who use at most half that number of hashtags, can thus overlap with at most half of *sunshineejc*'s hashtags.

SimRank clustering should not be dismissed, however. Although the clustering does not match the expected clustering exactly, clustering quality is greater. The two clusters have greater internal similarities and a lesser external similarity. It is interesting to note that *sunshineejc* switched from being an outlier to becoming a member of the closest pair of users. This switch is understandable, however, given that SimRank, unlike the weighted Jaccard similarity coefficient, does not "penalize" a user (reduce similarities between the user and all other users) for using hashtags no other users use, as SimRank random walkers always randomly teleport when reaching a dead-end. In fact, *sunshineejc*'s liberal use of hashtags opens many paths between *sunshineejc* and all other users, improving its structural-context similarity with other users. Conversely, *supermanhotmale*'s appearance as an outlier is attributable to the relatively scarce use of hashtags. The reasons for *sunshineejc*'s entry into the pro-Walker cluster and the corresponding entry of *diggrbiii* into the prounion cluster are not immediately obvious, however.

Although hashtag-based clustering produces intuitive results, we do not feel that hashtags by themselves should be used to cluster users for several reasons. First, hashtags are performative – they represent conscious decisions on the part of the user to include a tweet in a particular conversation. Theoretically, this makes them less reliable as indicators of a user's actual emotive alignment than links (which come from other users) and raw content (which is the actual currency of the user's communication). Relying excessively on hashtags also critically limits the potential of this model to be generalized to other domains. Furthermore, we have seen how a user's particular hashtag usage can dramatically influence that user's similarity to other users. Since hashtag usage is particular to a user, we cannot trust hashtag analysis to be consistent.

4.5 Mentions

Mentions-based clustering does not produce intuitive results. In fact, even with our knowledge of the users, it is difficult to substantiate an intuition behind any aspect of the clustering (see Figure 2).

It is possible that this seemingly non-meaningful clustering results from the binary use of mentions. We have observed that mentions encode much greater polarity than hashtags, especially when the dialogue is characterized more by attacking others and defending one's self than by the exchange of information. Users direct tweets at other users they either emphatically agree or forcefully disagree with, and little in between. In the aggregate, this mixture may confuse the attempt to measure similarity in approach as it is impossible to determine whether a user is mentioned in agreement or anger.

Analyzing the mentions-based clustering is complicated by the nature of the similarity measure. The effectiveness of this qualitative assessment relies on our ability to understand what we're clustering and how we're clustering it. That is, when we have an intuitive sense for how users *should* cluster, and when we have the ability to determine why our methods cluster users the way they do, we can evaluate both the performance of our model and the reasons behind this performance. Since we focus our analysis on a clustering of only eight users, understanding these users' approaches to the debate has been manageable. Furthermore, since comparing users has relied solely on the content produced by those users, understanding why the model clusters the way it does has also been possible.

Cultivating an intuitive sense for how users should cluster based upon mentions is not a manageable task, though, as it requires understanding not the eight users we have chosen to focus on but rather the manner in which all other users in the corpus engage with these eight users. Similarly, SimRank's recursive structure makes it exceedingly difficult to determine why users cluster as they do. On a superficial level, we understand that SimRank finds users *a* and *b* similar if users that mention *a* and *b* are similar. Concretely understanding why certain users are similar

	odds	tf-idf
aflcio	#statesos	#statesos
	#wearewi	#notmywi
	#notmywi	afl-cio
	to	#standupoh
	solidarity	blog
	capitol	biden
	afl-cio	vp
	#standupoh	#wearewi
	of	@melissaryan
	now	#humanrights
	#wiunion	statehouse
	workers	@wisaflcio
	rally	unbelievable
	on	pres
	blog	@tulaconnel
	one	#inunion
	from	officers
	with	text
	new	@defendwisconsin
	today	date
diggrbiii	#tcot	@diggrbiii
	the	@karoli
	?	#debunked
	@karoli	#tcot
	that	switch
	i	bait
	what	@derekahunter
	was	doctors
	@diggrbiii	@brentteichman
	#tlot	proving
	right	clinton
	would	classy
	and	#kochspiracy
	union	#tlot
	teachers	democrat
	#debunked	@mmfa
	doctors	eaglery
	democrats	@readyaimshoot
	this	chyron
	you	loony

Table 3 Effect of odds and tf-idf weighting

	odds	tf-idf
aflcio	#statesos	#statesos
	#wearewi	#notmywi
	#notmywi	#standupoh
	solidarity	afl-cio
	capitol	blog
	#standupoh	vp
	afl-cio	biden
	workers	#wearewi
	rally	@melissaryan
	now	#humanrights
	today	statehouse
	labor	@wisaflcio
	#solidaritywi	@tulaconnell
	#solidarity	officers
diggrbiii	#wiunion	@diggrbiii
	#tcot	@karoli
	@karoli	#debunked
	in	#tcot
	#tlot	doctors
	@diggrbiii	#tlot
	was	democrats
	democrats	was
	#debunked	have

Table 4 Effect of odds and tf-idf weighting, with G-test

and how that similarity cascades to increase other similarities, however, is difficult.

Although we cannot find meaning in the clustering produced by mentions-based similarity measures, we do not believe that the mentions-based clustering is absent of meaning – we simply do not know what it is yet. In the next section, we evaluate whether combining mentions-based similarity measures with content-based similarity measures can improve the clustering and shed light on the meaning encoded in a mentions-based clustering.

4.6 Combining Similarity Measures

We examine two dendrograms based on combined similarity measures, starting with the contenthashtag similarity measure. The optimal clustering coefficient was achieved at weighting hashtagbased Jaccard similarities by 0.7 and content-based cosine similarities by 0.3, then summing them. In this clustering (Figure 5), *sunshineejc* is no longer a complete outlier but is instead included in the wrong cluster. It is unclear whether this is an improvement, but the change does demonstrate the potential of using similarity measures that leverage different aspects of the content to help contextualize a user whose content is considerably different from all other content. More promisingly, the content-hashtag clustering does cluster users into two distinct clusters. This improvement is unsurprising given the quality and structure of the clustering based purely on hashtags.

For the content-mentions clustering (Figure 6), the optimal cluster coefficient was again achieved with a weight of 0.7 on the mentions-based SimRank similarities and a weight of 0.3 on the contentbased similarities. At first glance, the clustering does not seem meaningful. A closer look at the users clustered, however, reveals that the users in the first cluster, *ivanroberson*, *supermanhotmale*, and *aflcio*, are the only users who dedicate a significant proportion of their tweets to on-the-ground reporting and commentary, while all other users engage in the conversation in an abstract, armchair-commentary style. This is significant because while previous clusterings identify the users that have similar base attitudes towards the conflict, they do not distinguish users based upon *how* they expressed these attitudes. That is, no previous clustering distinguished between attitudes that were expressed via a discussion of the actual mechanics of the protests or via a more general commentary on the theoretical underpinnings of the controversy.

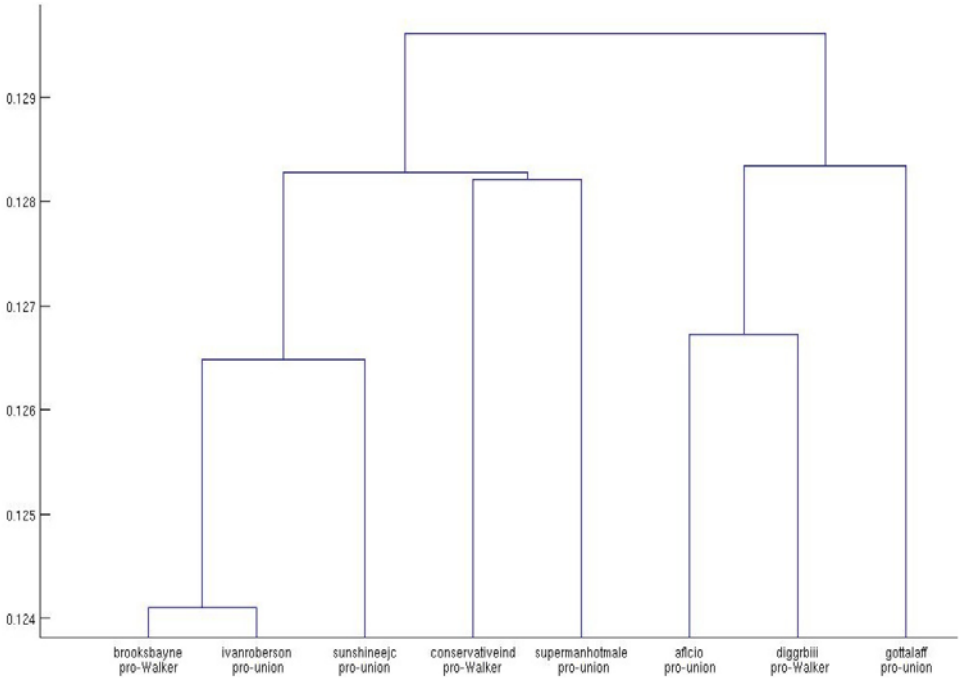


Fig. 2 SimRank-based mentions clustering

Note that since all of the users in the first cluster (*ivanroberson*, *supermanhotmale*, and *aflcio*) support the unions, it is unclear whether the clustering would have been as effective had some on-theground commentary users supported Gov. Walker. Furthermore, it is unclear whether this performance generalizes to other topics or domains or whether this is a pecuiliarity of the Wisconsin dataset that we have discovered.

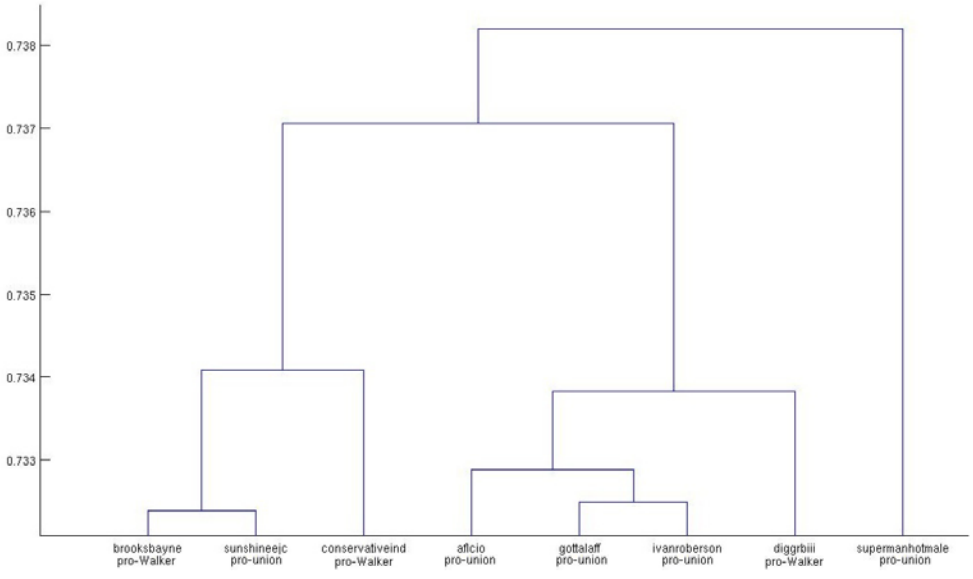


Fig. 3 SimRank-based hashtag clustering

5 Looking Ahead

We chose the Wisconsin corpus for our case study for a reason. The incredible polarity of the issue – the fact that a basic "a vs. b" narrative structure holds across the corpus – allows us to use "pro-union or pro-Walker" as a reasonable approximation of a user's frame. This quality provides us with a reference for evaluating the results of our model.

The intuitive meaning captured by the contentmentions clustering reminds us, however, that this is a simplification of the actual topography of the space. It reminds us of the core difficulty of this problem: it is hard to define, let alone quantify, a user's "approach" to an issue. Basic political bias

is certainly a part of it, but there are many more increasingly subtle aspects of a user's tweets that define how that user "frames" an issue. Whether the user engages with the controversy from the front lines or from an armchair is just another valid aspect of a "frame."

Given that a user's approach to an issue is multifaceted, we cannot summarize it with one similarity measure. Instead, we propose developing a system that clusters users hierarchically based upon different similarity measures corresponding to different aspects of the user's approach. The current model's ability to achieve this goal is rudimentary but demonstrates the concept well. Leveraging both content-hashtag clustering and content-mentions clustering, we could potentially engage in two-staged clustering in order to express the topography of the space as meaningfully as possible: content-hashtag clustering can be used to divide users into to clusters based upon their attitudes towards the topic, and the clusters produced by this clustering can be further subdivided by how the attitude is conveyed by using the results of a contentmentions clustering. With further research, we can identify other features that define a user's "frame" and leverage other aspects of the data to tell us how the users relate to one another with regard to that feature.

The narrowness of the scope required for a meaningful qualitative analysis will always limit the effectiveness of such an analysis and skew our conception of the model's performance. Therefore, it is important to develop a way to apply a comprehensive quantitative analysis to the model to verify and augment the qualitative analysis. There are two strategies that could be used to qualitatively assess the model. First, we could use hashtag-based clustering as a pseudo "ground truth" on which to evaluate the more general content-based clustering results. Second, we could leverage crowd-sourcing platforms such as MechanicalTurk to compile assessments of how meaningful different, more comprehensive, clusterings are.

Furthermore, additional research is needed to determine whether this model is generalizable. Twitter is a unique medium, and the model currently leverages much of that uniqueness, particularly the interesting link structures that result from the mix of performative and functional roles that links play. Research evaluating the model's performance in clustering content that is

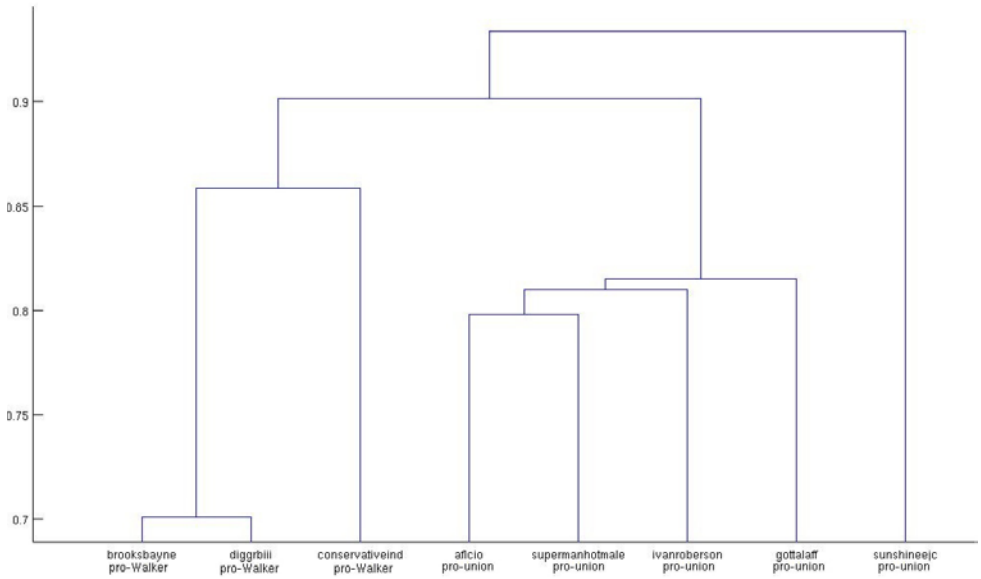


Fig. 4 Weighted Jaccard similarity coefficient-based hashtag clustering

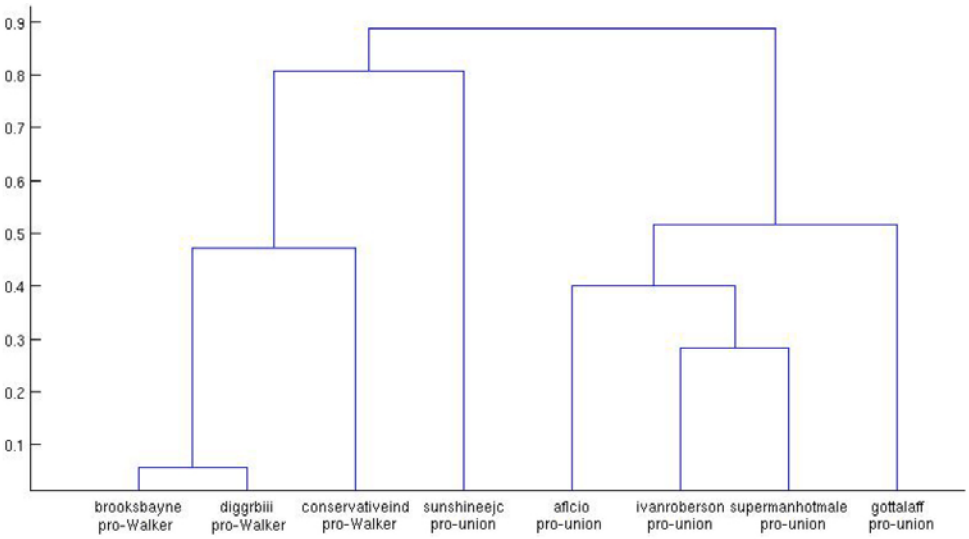


Fig. 5 Clustering based on a combination of hashtag- and content-based similarity

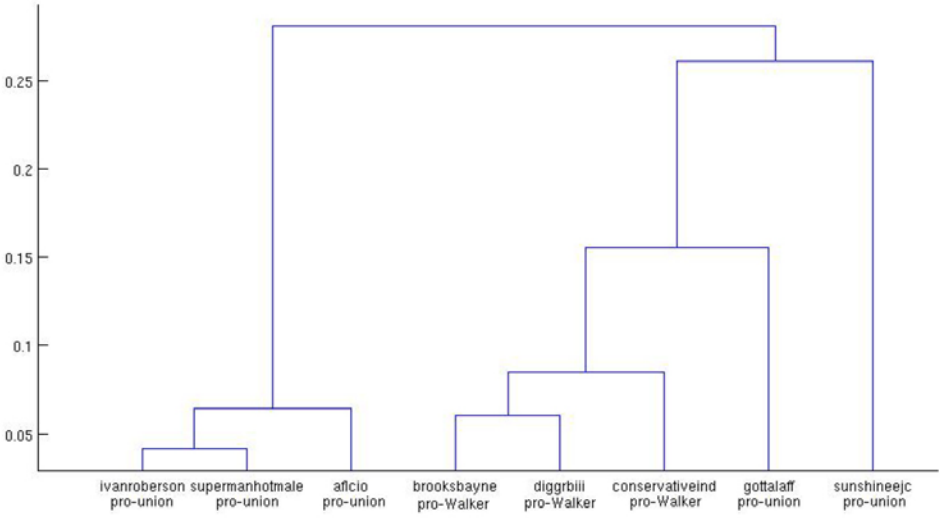


Fig. 6 Clustering based on a combination of mentions- and content-based similarity

less polarized with link structures that are more functional would be valuable.

Much work is still required, but this research already demonstrates the powerful symbiosis of quantitative and qualitative analyses. The development of an effective model and the understanding of social data are co-dependent, iterative processes: as we refine our understanding of the data, we can improve our model, and vice-versa. We began the case study with an understanding of how eight users qualitatively engage with the Wisconsin protest controversy, and we ended with a general model that we feel effectively clusters users based upon the user's engagement with the issue. We can now, in turn, leverage the model to help us understand how the 36,111 users we *didn't* analyze in the case study relate to each other.

Social phenomena, though always characterized by rich complexities, are now accompanied by ever greater mountains of data. To understand the complex dynamics of an issue, it is no longer enough to simply code up an intuitive model or study the raw data. Only an understanding of the data can fuel the development of an effective model, and only with the help of a model can we identify structure and meaning in mountains of data.

References

1. Gtz M, Leskovec J, McGlohon M, and Faloutsos C. Modeling blog dynamics. Proceedings of the Third International ICWSM Conference 2009, pages 2633, San Jose, CA, May. Association for the Advancement of Artificial Intelligence.
2. Leskovec J, Backstrom L, and Kleinberg J. Memetracking and the dynamics of the news cycle. Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2009; pages 497 506, New York. ACM.
3. Teevan J, Ramage D, and Morris MR. #twittersearch: A comparison of microblog search and web search. Proceedings of the Fourth ACM International Conference on Web Search and Data Mining, 2011, WSDM 11, pages 3544, New York. ACM.
4. Huang J, Thornton KM, and Efthimiadis EN. Conversational tagging in twitter. Proceedings of the 21st ACM conference on Hypertext and hypermedia, 2010, HT 10, pages 173178, New York, NY, USA. ACM.
5. Cha M, Haddadi H, Benevenuto F, and Gummadi KP. Measuring User Influence in Twitter: The Million Follower Fallacy. Proceedings of the 4th International AAAI Conference on Weblogs and Social Media (ICWSM), Washington DC, USA, May 2010.
6. Turney PD and Pantel P. From frequency to meaning: Vector space models of semantics. Journal of Artificial Intelligence Research, 2010; 37:141188.
7. Jones KS. A statistical interpretation of term specificity and its application in retrieval. Journal of Documentation, 1972; 28:1121.
8. Monroe BL, Colaresi MP, and Quinn KM. Fightin words: Lexical feature selection and evaluation for identifying the content of political conflict. Political Analysis, 2008; 16(4):372403.
9. Woolf B. The log likelihood ratio test (the g-test). Annals of Human Genetics, 1957; 21(4):397409.
10. Dunning T. Accurate methods for the statstics of surprise and coincidence. Computational Linguistics, 1993; 19(1):6174.
11. Turney PD and Pantel P. From frequency to meaning: Vector space models of semantics. Journal of Artificial Intelligence Research, 2010; 37:141188.
12. Rijsbergen CJ. Information Retrieval. Butterworth, 1979.
13. Jeh G and Widom J. Simrank: A measure of structural-context similarity. Technical Report 2001-41, Stanford InfoLab.
14. Jeh G and Widom J. Simrank: A measure of structural-context similarity. Technical Report 2001-41, Stanford InfoLab.
15. Hinton E and Kleefeld E. The wisconsin union struggle timeline. Talking Points Memo, 2011.
16. Isthmus: The Daily Page. Wisconsin capital protests: A 13-day timeline. Isthmus: The Daily Page, 2011.
17. Wisconsin State Journal. Timeline of budget bill protests. Wisconsin State Journal, 2011.
18. Sebastian M. 3 ways social media is fueling the protests in Wisconsin. Ragans PR Daily, February 2011.
19. Knutsen K. A guide to social media campaigns against scott walkers agenda for wisconsin public unions. Isthmus: The Daily Page, February 2011.

Moritz Sudhof is a recent graduate from Dallas, Texas. He majored in Computer Science. He has a passion for language and coding and is currently working on developing a first-of-its-kind emotional analysis engine at Kanjoya.

Masdar City: A model of urban environmental sustainability

Arthur Lau

In recent years, various proposals have emerged for the construction of "eco-cities," specially designed communities with an explicit environmental focus that often incorporate technologically sophisticated building design, comprehensive master-planning, renewable energy sources, and efforts to achieve resource self-sufficiency. This trend, however, is subject to intense criticism among some proponents of sustainable urbanism, on grounds ranging from the purported social exclusivity of eco-cities to the claimed incompatibility between a sustainable mode of living and continued economic growth in the contemporary Western model. Masdar City, an eco-city presently being built in the United Arab Emirates, serves as a useful case study in this debate, with such unconventional features as an underground network of tunnels for electric cars and an aim of developing a zero-carbon electricity supply. After examining the Masdar project in particular, and the scholarly context surrounding sustainable urbanism in general, this essay offers a new method for evaluating the environmental and social effects of eco-cities. Using the dual concepts of an "experimental" approach that favors a diverse collection of initiatives for ecologically conscious urban planning and a "human-benefits" perspective that emphasizes inhabitants' quality of life, the paper concludes that, despite its disadvantages, Masdar City represents an important advance in the field of sustainable urban design.

Located in the desert near Abu Dhabi and under construction since 2007, Masdar City is planned as one of the world's first completely sustainable communities, combining renewable energy sources and efficient resource usage with traditional Arabian design and spectacular architectural elements[1]. As Nicolai Ouroussoff reports in the New York Times, the entire city, developed jointly by the architectural firm Foster & Partners and the Abu Dhabi government, is raised on a 23 foot-high concrete base to maximize its exposure to cooling winds and decrease the need for air conditioning[1]. Gasoline-powered vehicles will not be allowed on the narrow streets of the city, about one square mile in area, but a fleet of computer-driven electric cars will navigate a complex of tunnels under the concrete base[1]. With an expected completion date of 2016, Masdar will require $22 billion for its construction, furnished by the government and independent investors [2]. Simon Joss of the University of Westminster includes Masdar among a collection of global projects, appropriately named "eco-cities," that represent the culmination of several decades of theoretical research into sustainable development[3]. Among advocates of sustainable urbanism, however, eco-cities built on the Masdar model remain controversial, with numerous scholars criticizing them as socially exclusive and overly dependent on technological improvements in place of broader societal change[1,4,5].

Following a review of Masdar City's master plan, particularly its environmental and social implications, and the current scholarship on sustainable urban design, this paper presents a general framework of analysis for eco-city development, as well as a specific assessment of the benefits and disadvantages of the Masdar project. Substantial progress in lessening the global ecological impact of cities requires widespread public support, and therefore depends on whether sustainable communities will also offer clear benefits for their residents' quality of life. Although it may perpetuate certain less desirable social effects, Masdar City could serve as the first clear public demonstration that environmental protection can be firmly and tangibly integrated with the development of a more attractive and livable urban community.

From any analytical perspective that might be applied to the city, the most striking feature of Masdar is undoubtedly the city's ever-present focus on efficiency and optimization, epitomized by a minutely detailed master plan that specifies everything from the type of electricity meters to the facades for the buildings[1,6]. For a community that aims to accommodate 40,000 residents, 50,000 commuters, and more than 1,500 businesses, this master-planning approach represents an immense and ambitious[7] undertaking. All buildings at Masdar, designed to maximize the use of natural light, must adhere to strict regulations concerning the use of insulation, low-energy lighting, and energy-efficient appliances[6]. Through these measures, Masdar City is projected to need only a quarter of the energy supply required by a normal city with the same population[8]. Water consumption will be reduced by installing high-efficiency fixtures and appliances and incorporating an advanced network of meters that not only notifies consumers of their usage levels but also determines the location of water leaks throughout the city[5]. In addition, the city's landscaping, which includes plants selected for their low water requirements, will be irrigated with treated wastewater, allowing Masdar's total water demand to be less than half that of a regular community[5].

Despite the fact that few buildings have been completed, Masdar's developers state that the city is already operating the largest solar photovoltaic plant in the Middle East[1,5]. Solar panels mounted on the rooftops and projecting over the streets will provide an additional source of energy, along with shade for the pedestrians below[7]. Furthermore, a planned geothermal energy project will pump water into the earth's crust to produce steam for electricity generation[9]. Some of the city's waste will be recycled, while some will be incinerated in an electricity-generating process that releases significantly less carbon dioxide than a landfill does[1,8]. With few sources of fresh water nearby, the city's water supply will be provided by a desalination facility that uses electricity from a massive hydrogen plant. This hydrogen plant is planned as the world's largest, and is expected to cost $2.2 billion[10]. Given the sheer size and diversity of its projects, Masdar serves as a showcase for unconventional planning methods and renewable energy technologies that other communities might find difficult to implement without Abu Dhabi's vast oil wealth. At the same time, the intense concentration on efficiency and renewable resources encapsulates Masdar's logical appeal: the notion that sustainable urbanism will become a broadly accepted and feasible goal only after someone demonstrates that a city with almost zero carbon emissions is possible on

a large scale.

In its architectural features, Masdar's developers show the same dedication to innovative urban planning, capitalizing on the environmental advantages of traditional Arabian architecture but not hesitating to employ expensive technological solutions. Centuries before the modern era, the people of the region designed their settlements to moderate the desert heat, building on high ground to obtain both defensive benefits and stronger winds and constructing tall wind towers to channel air currents onto the streets[1]. Such techniques inspired Foster & Partners to elevate Masdar above the surrounding area, employ wind towers of their own, and incorporate narrow streets oriented at an angle that maximizes shade at the ground level[1]. At the center of the developers' vision for the city lies the Masdar Institute, a graduate university affiliated with the Massachusetts Institute of Technology that opened in 2010 and focuses on research and engineering in sustainable technology[11]. To satisfy Middle Eastern norms for personal privacy, the residential buildings of the campus are enclosed in a wavy facade of concrete latticework, displayed in Figure 1, similar to the traditional screens known as mashrabiya[1]. While the latticework shields the interior from direct sunlight, the curves provide an angled view that prevents the inhabitants from seeing into the windows of buildings across the street[1]. In accordance with local custom, the campus is also segregated by gender, placing the living quarters for single men on one side and those for women and families on the other[1]. These aspects of the city's design contrast sharply with the distinctively modern style of the main buildings in the Institute, including the Knowledge Center, which houses study spaces under an iconic spherical roof covered with solar panels and zinc cladding[12], as seen in Figure 2.

Masdar's master plan expresses a sense of order and authority, founded in its developers' confidence that they can remake the whole structure of urban life through technological improvements, careful design, and active civic management. This observation is well illustrated by the "Personal Rapid Transit" system of automated electric cars, envisioned as a more efficient and attractive version of a public transportation network[6]. By eliminating not only gasoline usage but also the entire concept of the private automobile, the city's architects are transforming social standards and creating a new pedestrian community,

Fig. 1 The concrete facade of a building at the Masdar Institute.

easing the transition by offering almost exactly the same degree of comfort and convenience. From its developers' standpoint, Masdar City strengthens its emotional resonance with the promise of environmental protection and social revitalization that rise organically from the established local culture, essentially fusing modernity and tradition. In its larger structure, depicted in the artist's impression shown in Figure 3, the city exudes an aura of geometrical design, with the buildings arranged in neat squares, areas of solar panels and processing facilities placed around the outskirts, and parks occupying the remaining spaces. Of course, the city's architects could have the freedom to realize this ideal only in a sparsely populated desert region with extensive government funding.

In parallel with its promotion of sustainability, Masdar also furnishes a classic example of a development motivated by the financial potential of new technologies used to combat emerging social concerns. Over the next twenty years, the emirate of Abu Dhabi intends to use the Masdar project to encourage private entrepreneurship, facilitate scientific research in renewable energy, and generate economic growth in areas other than the oil industry[6]. As a special economic zone within Abu Dhabi,

Fig. 2 A view of the Knowledge Center at the Masdar Institute.

the city offers a zero-percent tax rate on income and imports, and permits the establishment of businesses that are wholly foreign-owned[13]. The developers argue that sustainable building practices will be adopted on a sufficiently broad scale to combat climate change only if environmentally conscious urban planning proves economically appealing[6]. Consequently, the meticulous master-planning approach that Masdar exemplifies might impose a level of social change that would be unacceptable for many people, and at the same time conform excessively to the conventional economic norms underlying Abu Dhabi's carbon-intensive culture. Currently, the United Arab Emirates possesses the world's highest per capita ecological footprint, a measure of the biological capacity required to fulfill resource demands and absorb waste emissions[14].

Indeed, the fundamental question of the extent to which sustainability requires profound societal and economic change applies not only to Masdar, but also to every other proposal in the literature on sustainable development. Many different models of sustainable urbanism have been devised, as Yosef Rafeq Jabareen, an urban studies lecturer at the Massachusetts Institute of Technology, explains in his article, "Sustainable Urban Forms: Their Typologies, Models, and Concepts."[15] "Neotraditional development," also known as the "new urbanism," seeks a more vibrant sense of community and more architecturally appealing development patterns, concentrating on mixed land use, walkable public spaces, diverse housing types, and somewhat higher population densities[15]. In another approach, "urban containment" efforts intend to lessen suburban sprawl by employing greenbelts, mass transit, and various regulatory policies[15]. In comparison with urban containment, as Jabareen notes, other initiatives promoting the idea of the "compact city" support a far higher level of building density and mixed land use, dramatically minimizing the demand for automobile travel[15]. Lastly, proposals for the "eco-city" emphasize urban greening and "passive solar design," which involves the efficient use of sunlight and absorbed heat in buildings, with a notable focus on the active management of social and economic life in the community[15].

In his survey of eco-cities, Joss notes that about three-fourths of current eco-city projects rely primarily on technological innovation rather than a more

holistic platform, including social and cultural aspects such as social justice and local democracy[4]. On this basis, Mike Hodson, a research fellow at the Centre for Sustainable Urban and Regional Futures (SURF) at the University of Salford, and Simon Marvin, a professor at Salford and the co-director of SURF, criticize recent eco-city proposals as socially regressive[16]. As their article "Urbanism in the Anthropocene: Ecological Urbanism or Premium Ecological Enclaves?" argues, cities must now guarantee their "ecological security" by safeguarding resources such as water, energy, and food supplies amid the uncertainties of climate change and a globalized economy[4]. Hodson and Marvin describe the emerging pattern of "integrated eco-urbanism," in which new developments "build ecological security by internally producing their own food, energy and other critical resources, reusing wastes as resources and reducing reliance on external infrastructures."[4] This approach, typified by Masdar City, emphasizes technological solutions, especially renewable energy sources, to overcome resource constraints in independent, self-sufficient enclaves[4]. From their perspective, eco-cities are offering ecological security as a "mobile financial product" that only the wealthy can afford, while excluding other social groups and failing to introduce any truly substantial behavioral or economic changes[4].

While Hodson and Marvin outline a socially oriented analysis, Susanna Myllylä of the University of Tampere and Kristiina Kuvaja of the University of Jyväskylä offer a critique of the eco-city movement from the perspective of the global "South," their term for developing countries[5]. In "Societal Premises for Sustainable Development in Large Southern Cities," they assert that the current concepts of an "eco-city" or an "ecological city" embody essentially "Northern" values and methods, prioritizing technocratic, managerial, and commercial strategies to combat environmental problems[5]. Eco-cities presuppose certain societal and governmental structures, particularly a democratic political system and an established civil society, when the formation of these structures itself presents a central challenge for communities in the South[5]. Indeed, as Myllylä and Kuvaja contend, efforts to implement sustainable urbanism in the South through Northern technological methods usually result in exclusive "ecological islets" surrounded by poor and environmentally degraded neighborhoods[5]. Instead, Myllylä and Kuvaja recommend a concentration on effective governance and civil society, as well as social justice and equity, noting that the pursuit of merely environmental sustainability as a goal in itself could worsen unjust social conditions[5].

In some ways, the diversity of perspectives on urban sustainability seems to arise from fundamental disagreements about sustainable development as a concept, particularly about whether planners should attempt to ameliorate specific environmental problems, or strive for a more comprehensive transformation of human society. According to Mark Roseland of Simon Fraser University, the very notion of sustainable development assumes that environmental protection can be reconciled with economic growth in a capitalist system[17]. While poverty is commonly viewed as a cause of environmental degradation, Roseland argues that both poverty and damage to the environment are actually produced by wealth[17]. To achieve the objectives of sustainability, he maintains that "we must . . . shift our economic development emphasis from the traditional concern with increasing growth to reducing social dependence on economic growth."[17] Likewise, in his article "Synergy City: Planning for a High Density, Super-Symbiotic Society," John Wood presents the case that the economic success of the developed world is based on the availability of inexpensive fossil fuels18. A researcher from the department of design at the University of London, Wood asserts that an ecologically conscious form of society is directly opposed to one that promotes consumerism, convenience, and profit, calling instead for people to adopt more "synergistic," collaborative lifestyles[18].

Many scholars believe, however, that modern society and economic prosperity are indeed compatible with vigorous efforts to protect and preserve the environment. For instance, in The Human Sustainable City, a compilation of scholarly articles on sustainability, Stefano Zamagni of the University of Bologna discusses an interesting statistical finding that could undermine Roseland's premise that environmental damage results from wealth. Although environmental decay increases with the growth in average income when the latter remains low, it decreases with income growth once the average annual income has exceeded a threshold of approximately $8,000 per capita[19]. Another paper in The Human Sustainable City, by Peter Hall of University College London, contends that ecological degradation harms economic development, and that the protection of the environment can promote economic growth[20]. Finally, Michael Breheny, a professor of Applied Geography at the University of Reading, advises caution in instituting drastic social changes for environmental aims. He produces evidence from surveys in Britain that the extensive social changes entailed by urban compaction, a frequently promoted method of reducing cities' environmental impact, would be strongly resisted by most residents[21]. As Breheny argues in another article, proposals to shift communities to the model of the compact city, with higher densities and greater provision of public transportation, would demand draconian regulatory regimes but yield, at maximum, about a 30% decrease in automobile energy consumption[22].

Practical examples of the implications of these different theoretical approaches to sustainable development can be found by comparing Masdar City to some of the various envisioned, or actually constructed, eco-cities around the world. As the New York Times reported in 2007, for example, the Chinese government, in partnership with the British engineering firm Arup, was planning to establish the Dongtan eco-city on an island in the Yangtze River[23]. In the city, designed to accommodate 500,000 inhabitants, electricity would be supplied by solar panels, wind turbines, and the burning of discarded rice husks, while gasoline-powered cars would be prohibited[23]. Farms and parks would occupy most of the island's land area, and almost all waste generated by the city would be recycled through an automatic sorting system integrated into the streets[23]. Nevertheless, by 2010, when the first phase was supposed to be finished, the completion of Dongtan was indefinitely postponed and most construction work was suspended[24]. The reasons for the apparent abandonment of the project included uncertainty concerning whether it would be funded by Arup or the government, as well as the failure of the developers to engage the local community in the planning process[25].

Curitiba, a city in Brazil that is also classified as an eco-city, provides an illustration of a more successful attempt at environmentally beneficial urban planning, with an especially notable feature being its innovative bus system[3,26]. As Arthur Lubow

describes in the New York Times Magazine, specially designed buses, operating on dedicated avenues and arriving at each stop once every 30 seconds during peak hours, transport 2.3 million people per day, more than two-thirds of the metropolitan region's total population[26]. In the 1970s, Curitiba converted its downtown shopping district into a pedestrian zone, blocked the entry of polluting industries, and increased the land area devoted to parks and open spaces dramatically, from 5 square feet for each person to 540 square feet[26]. Recently, nevertheless, the city's per capita car ownership has risen to the highest in Brazil, and its recycling rates have been decreasing, with some residents attributing these trends to low public participation in planning decisions and limited community willingness to adopt sustainable behaviors[26].

Not coincidentally, Curitiba's environmentally progressive master plan was formulated when a military dictatorship ruled the country[26], in a remarkable parallel with the authoritarian Chinese government's support for Dongtan, and Masdar's dependence on the hereditary emirate of Abu Dhabi. Comprehensively master-planned eco-city projects necessitate such large expenditures and strict regulations that only non-democratic governments might be capable of developing eco-cities on a large scale. For instance, as Danyel Reiche of the American University of Beirut notes in "Renewable Energy Policies in the Gulf Countries: A Case Study of the Carbon-Neutral 'Masdar City' in Abu Dhabi," political power in the United Arab Emirates remains mostly with the seven emirs[7]. At the federal level, foreign affairs and defense are coordinated by the emirs, and political parties are banned; oil and natural gas reserves are controlled by the individual emirates, accounting for the striking wealth of Abu Dhabi, where almost all the country's energy resources are located[7]. Governmental decisions are strongly influenced by personal access to the ruling family: Reiche reports that the idea for building Masdar was developed by Lebanese engineers in contact with the crown prince of Abu Dhabi, who conveyed it to the emir[7]. In another article, he writes that the Gulf monarchies, such as the United Arab Emirates, are impeded in efforts to advance sustainability by their reliance on oil revenues. Using income from oil exports, these regimes gain popular acquiescence to the lack of political rights by imposing low taxes, offering free medical care, and subsidizing energy supplies[14].

As the examples of Curitiba and Dongtan show, however, even with strong initial support from the government, attempts at ecologically conscious development can falter when they are not grounded in widespread public acceptance of sustainability measures and strong community involvement in the planning process. For example, Lurton Blassingame of the University of Washington at Oshkosh observes in a review of urban sustainability endeavors that residents must be willing to alter their ways of life, not only in terms of environmental impacts but also with respect to socioeconomic changes[27]. In Psychology of Sustainable Development, a collection of papers on environmental psychology, Elisabeth Kals and Jürgen Maes of the University of Trier in Germany identify several factors contributing to engagement in sustainable behaviors. Some of the most important elements include "ecological awareness" about environmental problems, the belief that one's actions exert a substantial influence on the extent of ecological damage, and "emotional affinity" toward nature[28]. Hence, the question of how people can be induced to accept sustainable practices becomes partially a question of how their awareness of the natural world and their emotional connection to it can be reinforced.

Nonetheless, some scholars in sustainability studies are unfortunately inclined to discuss environmental issues in ways that seem unlikely to promote greater public understanding of ecological concerns and support for sustainability initiatives. Wood, for instance, advocates the use of urban design in a "frankly manipulative" manner that involves "a process of 'seeding' consensual change"[18]. Later he writes, "Not surprisingly, consumer-centred, representative democracy has conspicuously failed to wean us from a way of life that threatens us with extinction."[18] Besides providing no scientific justification for his claim that the very survival of the human species is imperiled, Wood's statement may hold limited motivational force; Kals and Maes explain that "ecological fear" of detrimental effects from the environment on one's well-being has only a minor impact on behavior[28]. Another author, Peter H. Cock of Monash University, opens his paper in Psychology of Sustainable Development with a similarly startling sentence: "Hidden behind the blindness of our suicidal cultures is an insidious biophysical poisoning."[29] Cock also espouses a "Sacred Ecology" emphasizing "the spiritual journey of self-transcendence and inner reflection activated through encounters with the 'otherness' of other species and their habitats."[29] Calls for a transformation of religious morality and vituperative remarks about the societies in which people live will probably be unable to persuade large majorities of citizens in the Western world to modify their ways of life in favor of sustainability.

On the basis of both the theoretical and practical aspects of urban sustainability, one can develop a new framework for assessing efforts to create more sustainable communities. This proposal recognizes that climate change and ecological degradation occur on a global scale, and that efforts to protect and restore the environment require widespread public support across nations and social groups. Any solution must emphasize community participation in urban planning and, most importantly, combine environmental benefits with demonstrable improvements for residents' quality of life. All these criteria are satisfied by an "experimental" method that embraces many divergent, and even philosophically opposed, approaches to the task of reducing the worldwide environmental impact of cities. Indeed, as Solon L. Barraclough of the United Nations Research Institute for Social Development aptly states, "A unified approach to promoting sustainable development in diverse contexts is a utopian illusion."[30] With radical differences in economic, social, and political conditions between countries and even between neighborhoods in the same city, the "experimental" method acknowledges that each community must implement techniques and policies particularly designed for its own local circumstances.

Ultimately, the concerted and inventive application of these ideas in urban communities throughout the world should result in the development of what might be called "human-benefits" cities that would secure broad advantages for both the health of the natural environment and the standard of living for all inhabitants. These cities would embody the central principle that ecologically sustainable communities should simply be authentic and attractive cities in general, with lower pollution, improved health, greater social solidarity, and a strong appreciation for the beauty and complexity of nature. In contrast to the many sustainability proposals, such

Fig. 3 An artist's impression of the master plan for Masdar City.

as those of Wood or Myllylä and Kuvaja, that also seek improvements in aspects of urban life besides the environment, the "human-benefits" perspective does not fix any particular objectives other than ecological protection. Some cities might pursue social justice, some might strive for economic revitalization, and others might concentrate on cultural and educational opportunities: the "experimental" method allows each city to follow the priorities of its own residents. With this proposal, developers and municipal officials would eschew externally designed master plans in favor of broadly supported initiatives arising organically from recognized local concerns.

"Human-benefits" cities would focus on incentives for adopting sustainable behaviors rather than penalties for violating environmental regulations, seeking to maintain the interest and engagement of businesses and residents. This recommendation is inspired by the findings of Karen Umemoto and Krisnawati Suryanata, professors at the University of Hawaii at Manoa. They discuss the use of "informal social contracts" to develop trust and reciprocity between a mariculture company and local inhabitants and to resolve conflicts over environmental impacts without expensive litigation[31]. In the same manner, instead of an excessive reliance on legal constraints, cities can dissolve contention by implementing a problem-solving methodology that attempts to achieve advantages for all parties involved. Furthermore, the "human-benefits" approach would encourage the construction of new communities only in environments naturally suitable for human life, not in inhospitable deserts or ecologically fragile rainforests. Using locally sourced building materials and supporting research into sustainable technologies would yield not only more sustainable patterns of construction, but also benefits for the local economy. Most importantly, this proposal would offer more opportunities for innovation, greater public participation in civic decisions, and even more personal freedom. As Hall argues, "If there is an argument for higher densities and particular urban forms, it is far more that these give people more choice, not less: freedom to take good public transport rather than be car-dependent, freedom for their children to walk or bike to school rather than being ferried by car."[20] This approach solves the problem of ecological awareness and motivation, not by erecting an entirely new system of environmentally conscious values and ethics, but by drawing on the interest that humans have perennially placed in their own health, their communities, and improved living conditions for their families.

Applying this analysis to Masdar City specifically, one can identify clear disadvantages for the project, especially since the planning process has been controlled by the government for its own purposes, rather than responding to the concerns and desires of the local community. Indeed, the city is isolated from any other community by the surrounding deserts, and the high concrete base would prevent Masdar from being naturally integrated with any nearby neighborhoods that might exist in the future, raising the issue of social exclusion. As Hodson and Marvin suggest, moreover, while eco-cities are generally new developments, ecological security should be extended to the less privileged by focusing on retrofitting and behavioral change in existing cities[4]. Nevertheless, the "experimental" method does not demand that myriad ecological and social objectives all be advanced by a single initiative: as Blassingame notes in his article, sustainable development is a process, not a final product[27]. Even if it is lacking in certain respects, Masdar City represents profound progress in the central task of demonstrating the feasibility of a "human-benefits" city that purposefully connects environmental protection, economic opportunity, and an improved quality of life in a unified and potent vision. Specially designed, self-contained eco-cities can be effectively used to stimulate sustainable development in existing neighborhoods, but their architects should expect them to be incorporated into the larger community as environmentally conscious design is more broadly adopted in the future.

Acknowledgement

This paper was written for a course taught by Dr. Kimberly Moekle. Her support and advice are gratefully acknowledged.

References

1. Ouroussoff N. "In Arabian Desert, a Sustainable City Rises." *New York Times*. 25 Sept. 2010.
2. Bullis K. "A Zero-Emissions City in the Desert." *Technology Review* 2009; 56-63.
3. Joss S. "Eco-Cities—A Global Survey 2009." *WIT Transactions on Ecology and the Environment* 2010; 129: 239-50.
4. Hodson M. and Marvin S. "Urbanism in the Anthropocene: Ecological Urbanism or Premium Ecological Enclaves?" City 2010; 14: 298-313.
5. Myllylä S and Kuvaja K. "Societal Premises for Sustainable Development in Large Southern Cities." *Global Environmental Change Part A* 2005; 15: 224-37.
6. "What is Masdar City?" 2011. Masdar City. <http://www.masdarcity.ae/en/27/what-is-masdar-city>.
7. Reiche D. "Renewable Energy Policies in the Gulf Countries: A Case Study of the Carbon-Neutral 'Masdar City' in Abu Dhabi." *Energy Policy* 2010; 38: 378-82.
8. Fortson D. "Green City Rises from Desert—Abu Dhabi Is Building a $22 Billion Zero-Carbon Showcase City for 50,000 Residents. Danny Fortson Went to See Its Beginnings." *Sunday Times* [London] 1 Feb. 2009, 2nd ed.
9. Gavin J. "Taking the Lead on Clean Energy." *MEED: Middle East Economic Digest* 2009; 40-41.
10. James E. "Hidden Costs of Alternative Energy." *MEED: Middle East Economic Digest* 2008; 30-31.
11. "Official Opening of the Masdar Institute

Campus, First Solar Powered Building at Masdar City." 23 Nov. 2010. Foster & Partners. <http://www.fosterpluspartners.com/ News/428/ Default.aspx>.
12. "Knowledge Center." 2010. Masdar Inst. of Science and Technology. <http://www.masdar.ac.ae/campus/KnowledgeCenter.php>.
13. "Special Economic Zone." 2011. Masdar City. <http://www.masdarcity.ae/en/38/ special-economic-zone>.
14. Reiche D. "Energy Policies of Gulf Cooperation Council (GCC) Countries—Possibilities and Limitations of Ecological Modernization in Rentier States." *Energy Policy* 2010; 38: 2395-403.
15. Jabareen YR. "Sustainable Urban Forms: Their Typologies, Models, and Concepts." *Journal of Planning Education and Research* 2006; 26: 38-52.
16. *SURF: The Centre for Sustainable Urban and Regional Futures*. 2011. U of Salford.
17. Roseland M. "Sustainable Community Development: Integrating Environmental, Economic, and Social Objectives." *Progress in Planning* 2000; 54: 73-132.
18. Wood J. "Synergy City: Planning for a High Density, Super-Symbiotic Society." *Landscape & Urban Planning* 2007; 83: 77-83.
19. Zamagni S. "Sustainable Development, the Struggle against Poverty and New Structures of Governance in the Era of Globalization." In: Luigi Fusco Girard et al eds. The Human Sustainable City. Aldershot: Ashgate, 2003: 121-39.
20. Hall P. "The Sustainable City in an Age of Globalization." In: Luigi Fusco Girard, et al eds. *The Human Sustainable City*. Aldershot: Ashgate, 2003. 55-69.
21. Breheny M. "Urban Compaction: Feasible and Acceptable?" *Cities* 1997; 14: 209-17.
22. Breheny M. "The Compact City and Transport Energy Consumption." *Transactions of the Institute of British Geographers* 1995; 20: 81-101.
23. Brass K. "Breaking Ground on Eco-Cities near Shanghai." *New York Times*. 4 July 2007.
24. Sanford L. "A Tale of Two Eco-Cities." *Modern Power Systems* 2010; 20-22.
25. Brenhouse H. "Plans Shrivel for Chinese Eco-City." *International Herald Tribune*. 24 June 2010.
26. Lubow A. "The Road to Curitiba." *New York Times Magazine* 2007.
27. Blassingame, Lurton. "Sustainable Cities: Oxymoron, Utopia, or Inevitability?" *Social Science Journal* 1998; 35: 1-13.
28. Kals E and Maes J. "Sustainable Development and Emotions." In: Schmuck P, Schultz W eds. *Psychology of Sustainable Development*. Boston: Kluwer, 2002. 97-122.
29. Cock PH. "Partnerships for Sustainability: Psychology for Ecology." In: Schmuck P, Schultz W eds. *Psychology of Sustainable Development*. Boston: Kluwer, 2002: 175-95.
30. Barraclough SL. *Toward Integrated and Sustainable Development?* UNRISD Overarching Concerns 1. Geneva: United Nations Research Inst. for Social Development, 2001.
31. Umemoto K and Suryanata K. "Technology, Culture, and Environmental Uncertainty: Considering Social Contracts in Adaptive Management." *Journal of Planning Education and Research* 2006; 25: 264-74.

Figure References

1. "What Is Masdar City?" 8 June 2011 <http://www.masdarcity.ae/en/27/what-is-masdar-city>.
2. "Knowledge Center," Masdar Institute of Science and Technology, 8 June 2011 <http://www.masdar.ac.ae/campus/ KnowledgeCenter.php>.
3. "What's Next?" 8 June 2011 <http://www.masdarcity.ae/en/36/whats-next>.

Arthur Lau is a sophomore from Dallas majoring in Philosophy, Classics, and Mathematics. His principal interests are presently concentrated around the philosophy of action, the philosophy of government, and the history of philosophy, with an aspiration to conduct research on connections between property theory and concepts of agency. Arthur is also a member of Aisthesis, the newly founded Stanford undergraduate classics journal. He plans to obtain a dual J.D. and Ph.D. in philosophy.

The sordid economy: Gender and disease in nineteenth century labor struggles

Zoe Leavitt

Damning exposés of disease, degeneracy, and filth bubbling forth from the borders of Chinatown populate the pages of California books and newspapers from the 1850s-1900s. Anti-Chinese writers frequently portrayed the immigrants as a spreading cancer, a destructive force, or a poison. Moral arguments, political arguments, and economic arguments formed the three main pillars of anti-Chinese causes, while labor advocacy groups, such as the Workingman's Party of California and the Knights of Labor formed some of the strongest anti-Chinese critics. However, while hysteria about physical force and rape dominated discourse surrounding African Americans – the other major group unable to access "whiteness" well into the 1900s – disease became the driving force in defining Chinese threats and difference. How and why did this characterization develop? This paper will argue that white men built off the 1800s' gendered conceptions of race and nation to understand and construct Chinese men as effeminate, and then used this gendered portrayal to demonize Chinese labor as a vessel for perversion and disease. This formed the foundation of white men's attempts to protect their own economic interests.

The Feminization of Chinese Men

"To the Caucasian race, with its varied types, has been assigned the supremacy in elevation of the mind and beauty of form over all mankind."[1] With these words published in 1862, Dr. Arthur Stout displayed the Social Darwinist ideals that human races, like animals, evolved according to the dictates of survival of the fittest. Granting white males the highest rung on the evolutionary ladder, Dr. Stout and other Social Darwinists bolstered the confidence of an American manhood shaken by shifting economic environments, endowing it with the responsibility of leading the international system and the glory of taking the first steps toward racial perfection. Naturally appealing to a nation struggling with immigration and Western expansion, many Social Darwinist thinkers used Darwin's ideals to justify exploitive capitalism and colonization on the basis of the supposed weakness of non-whites.[2]

In America, theories of manliness came to imbue the Social Darwinist hierarchy with gendered characteristics – visions of races developing into higher races over time morphed into metaphors for the way boys grew into men over time. As the most highly developed of all the races, white men illustrated "the achievement of a perfect man, just as civilization of a perfect race."[3] Lower races, on the other hand, had not yet completed the boy-man development track, thereby supporting an inability to ever become true men in one lifetime. As one 19th century anthropologist stated, "The intellect – the mental strength – the moral beauty, all...harmonize perfectly with the growth of the beard, and when that has reached its full development, it is both the signal and the proof of mature manhood... The Caucasian," he emphasized, "is the only bearded race."[4] Such rhetoric illustrated the belief that "savage men were emotional and lacked a man's ability to restrain their passions...Savage men even dressed like women."[5] By focusing on the blurring of genders in other races and the effeminate habits of non-white men, white men could emphasize the continuation of their own manliness.

Extending gendered notion of race, nations as a whole became associated with their populations so strongly that America as a whole earned masculine traits in comparison to the feminine characters of other countries. America's perceived position of world caretaker cemented this vision, seen strongly in cartoons such as "The White Man's Burden" that portrayed a strong American male lifting a skirt- and jewelry-wearing non-white man toward civilization.[6] Indeed, political cartoons generally illustrated nations as people and America as Uncle Sam or another man, cementing the idea of nations as single bodies and of the American body as a masculine one.

While America became a masculine figure in such bodies, China became a feminine one. Political cartoonists drew Chinese men as weak, bent, twisted, and sleazy, in comparison to the straight, upright figure of the American male.[7] Even across the sea, several Chinese thinkers posited the "yellow" race and the "white" race in similar terms. Adapting Western ideas of Social Darwinism in their own way, writers such as Yi Nai, in his 1898 essay, "China Should Take Its Weakness for Strength," applied a male-female binary to the international trans-Pacific discourse. Feminine China, graceful and soft, provided a natural counterpart to the strength and aggressiveness of America in his vision.[8] While Nai's rhetoric used such differences to display interdependence rather than the insurmountable barrier to assimilation that many Americans saw, his work displays the strength that gendered races – specifically the feminine Chinese perception – held in the late 19th century.

The Labor of Chinese Men

In the fight against Chinese immigration, fears of Chinese male labor rose to the forefront, as anti-immigration activists painted Chinese laborers as a contagion that would undermine the integrity of American workers and the health of the nation itself. Rhetoric turned the contest between Chinese and American laborers into a battle between the two nations, pitting the races' supposed values, ideals, and living standards against each other. America's economic troubles of the late 1800s, therefore, became the source of much of the anti-Chinese sentiment of the period, as racialized and feminized portrayals of Chinese men rose as weapons in the fight for white employment.

When the first large wave of Chinese male immigrants arrived on US shores in the 1840s, they were welcomed as cheap labor, a supposedly more docile alternative to black slaves in the South and a new, affordable labor force for the sparsely settled West.[9] "Ingenious, industrious and submissive," Chinese men accepted low wages, menial jobs, and low living standards.[10] Employers perceived them as docile and diligent, and best of all, they sent much

of their money back to China, seemingly showing no desire to settle in America permanently.[11] Labor-scarce California especially looked toward Chinese immigrants as an important force in mining and building railroads; however, when Australian gold discoveries in 1853 caused California mining to collapse, the recent influx of Chinese miners became an immediate and visible scapegoat. As a 1924 textbook explained, "to these unemployed men, the presence of thousands of Chinese, thrifty, industrious, cheap, and above all, un-America, was obviously the cause of their plight."[12] The Panic of 1873, in which an agricultural drought, the burst of a speculation bubble, and the completion of the Transcontinental Railroad combined to topple Californian employment rates, only threw more kindling on the anti-Chinese-labor fire.[13]

Throughout this series of recessions, the stereotypes that had previously formed the strongest selling points for Chinese labor transformed into its fiercest condemnations. Chinese men's acceptance of low wages turned into a symptom of slimy servility, their thriftiness in living came to reveal their base nature – all traits rhetoricians used to paint them as unfit to work aside Americans. In 1870, well-known editorial writer John Swinton proclaimed, "We may look on the industrial contest between the Chinese and American races as a contest of the bunk against the bed, the roast rat against the roast beef – the baser side of life against the better."[14] As the "roast rat against roast beef" image became iconic in the anti-Chinese labor struggle, Chinese men were painted as the antithesis of the time's manly ideals. They ate and slept in packed communal areas, seemingly as animals, and had no visible families to support. Employers did all of America a disservice by hiring these Chinese laborers, such rhetoric implied, as the Chinese would eventually erode living conditions for all workers in America. The battle against Chinese laborers, therefore, became shaped as a battle to defend Americans' standards of living as men, not as beasts.

Just as anti-Chinese groups argued that Chinese laborers would weaken the standards of the entire labor market in America, they attempted to show Chinamen as using their labor as a treacherous way to gain Americans' trust before a grand betrayal. Rather than fighting upright like a true man, Chinese men would act in a cravenly, treacherous, and more feminine way to harm the American nation. Doctor Arthur B. Stout vividly decried what he saw as this underhanded method of attack: "Better would it be for our country that the hordes of Genghis Khan should overflow the land, and with armed hostility devastate our vallies with the sabre and the fire-brand, than that these more pernicious hosts, in the garb of friends, should insidiously poison the well-springs of life, and spreading far and wide, gradually undermine and corrode the vitals of our strength and prosperity."[15] Brought to America to contribute to the economy, Stout believed the Chinese would slowly eat away at the nation's prosperity from the inside. The act of an obedient laborer here becomes a mere ploy to gain access to America, showing employers the betrayal they commit to the entire future of their nation by hiring Chinamen.

Even the most anti-Chinese groups rarely portrayed the immigrants as instigators of direct aggression. The Wasp, a fiercely anti-Chinese San Franciscan publication, consistently caricatured the Chinese as deceitful parasites dependent on others' foolish trust. The Wasp's 1877 cartoon "The Modern 'Old Man and the Sea'" presents an excerpt from "Sinbad the Sailor" about a man who tricked Sinbad into carrying him by acting old and crippled, and creates a comparison to the Chinese by drawing a Chinaman atop the shoulders of Lady Eureka, a symbol of California.[16] In this image, not only must the Chinaman resort trickery to gain his life's support from others' strength, he has made himself dependent on a woman – a gender reversal strengthening the line between weak Chinese men and ideal white men who are strong enough to support others. Another satire, an 1879 play entitled "The Chinese Must Go," connected these themes more explicitly to the labor market. After equating money to blood, as money buys food that, once digested, becomes blood, the main character asks, "What would you think of a man who would allow a lot of parasites to suck every day a certain quantity of blood out of his body... And suppose these Chinese parasites should suck as much blood out of every State in the Union, destroying Uncle Sam's sinews and muscles, how many years do you think it would take to put him in his grave?"[17] Therefore, in the face of Chinese labor, the white race is not killed outright, but rather slowly bled out, weakened until it can stand no longer.

Through these threats to labor, the themes of feminization previously explored by this paper appear as a method by which to emphasize the danger and treachery inherent in Chinese labor. Anti-Chinese labor pamphlets continuously focused on the stereotypically feminine jobs that Chinese men took on – due to their position in the gendered racial hierarchy Chinese males taking on feminine work was both expected, and seen as suspect. This emphasis further serves to illustrate how "against an emergent heterosexual and dimorphic order, Oriental sexuality was constructed as ambiguous, inscrutable, and hermaphroditic,"[18] while enflaming deeper labor fears. In songs such as "Irish Widdy Woman" (1868) and stories such as "Poor Ah Toy," Chinese men replaced Irish women in domestic-quality jobs, while a Wasp cartoon drove home the consequences of such employment thievery: a line of desperate women snaking from an unemployment office door, a Chinese servant snickering at their despair.[18] These labor groups expressed that not only did Chinese men lower the wages and living standards of American men to that of animals, they shamelessly stooped to scoop up women's work.

Beyond exploring the women who lost their jobs to Chinese men, anti-Chinese labor groups continued to use the perversity and weakness they saw as inherent in Chinese men to render horrifying tales of the treachery posed by Chinese houseboys. In an 1879 narrative by the anti-Chinese Workingman's Party of California (WPC), the Chinese servant abandons his manhood in order to steal jobs from true Americans that have "sex, and shame, and sense of propriety," and then to worm his way into white families' trust. "He is so perfectly compliant and neutral that Madam and the children come to regard him as a thing, as a dog or a monkey," the pamphlet explains.[19] As middle-class white men labored to solidify their own image of masculinity, fear surrounded these seemingly inscrutable creatures, these seeming not-men who were willing to take on tasks from washing dishes to cleaning bathrooms. The mistress of the house will soon become too complacent around her seemingly passion-less aid, says the pamphlet, and then, insidiously, "the manners of the harem are introduced. The lady is taught all the luxuries of the East. She is now a pupil, and John her teacher... Thus our dames are acquiring strange, outlandish ways. Thus the luxury and debauchery of the Orient are brought into the houses of the rich."[20] This rheto-

ric concurrently shapes Chinese men into a silent yet lurking presence and echoes themes of manhood's reservation for the most highly evolved species. Too weak to fight the white race head-on as men, Chinese men must enter as Trojan horses into white homes, to take the cravenly route of subversion of women rather than upright confrontation.

Deepening California's recessionary mire by replacing American men at lower wage levels, sacrificing their dignity and sexuality to live in squalor and perform women's work, and using labor as a tool through which to undermine social values, the Chinaman, as portrayed by California labor groups, was a fearsome beast indeed. At the same time, by continuing to emphasize the weakness and cowardly tactics of the Chinese, such rhetoric showed that only when Americans voluntarily allowed Chinese entry, could they do the social order any harm. This idea encouraged white employers to present a united front against Chinese labor's seduction, discouraging such hiring by laying the responsibility on Americans to keep their nation pure and strong.

Chinese Labor and Disease

Beyond explaining explicit economic threats and the social perversion that Chinese men brought to America, anti-Chinese rhetoric decried the threat they posed to American bodies, and therefore to the nation itself. Just as bodies and bodily ideals grew in importance for manliness in the late 1800s, so did ideas of protecting the body of American superiority. When "white" becomes a synonym for "superior" and "race" becomes a synonym for "nation," an attack on any one of these aspects becomes an attack on all of them. Dr. Stout connected the circle between national sickness and bodily sickness in the late 1800s, writing that "The abuse of the human system, insidiously gnaws into each individual body, undermines the strength and beauty of God's noblest work, and thence penetrates, cancer-like, into the social, religious, and political system."[21] In this view, each body is a building block of the societal pyramid – on which, of course, white civilization rests at the top. The physical strength of white manhood protects the superiority of white society, while the feminine sickness of Chinese immigrants threatens it.

Associating syphilis and leprosy specifically with Chinese immigrants magnified the type of degeneration that Chinese lifestyles and habits would cause in America. Many people in the 1890s considered these diseases hereditary, a point that emphasized the unalterable difference of the Chinese race and an internal inability to assimilate. According to Shah, "By equating leprosy's disfiguration, disintegration, and dehumanization of the body with Chineseness, medical discourse could recalibrate notions of incommensurate racial difference."[22] Pamphlets and "tour guides" of Chinatown, provided the public of sensationalistic accounts of "forms of vice, which in other countries are barely named, are, in China so common that they excite no comment among the natives."[23] These "Chinese paradise[s] of filth" bred diseases and physical degeneracy that sat lurking inside every Chinese body.[24] Furthermore, diseases such as syphilis could sit dormant within bodies for years, evoking an image of Chinese-caused decay slowly eating America from the inside out – an image that echoes the treachery and betrayal seen in Chinamen. Portrayals of Chinese diseases as more virulent that those among whites, as seen in the segregation of lepers and smallpox patients by race[25] and in testimony to the Special Committee on Chinese Immigration in 1877-78[26] maintained unbreakable lines between Chinese people and their Chinatown backgrounds, and these backgrounds and disease.

Due to the equation of Chinese people as a single intrusion onto the American national body, the rhetoric of quarantine arose as the only way to keep the white race safe. Every American that hired a Chinese laborer opened a pore in the national body through which contagion could easily slide. The Anti-Chinese Council wrote that as Chinamen gained work among whites, "a perfect network of contagion and infection is created, a veritable octopus of disease, having its seat in Chinatown and its infectious arms thrust into every house of the city."[27] Labor groups railed at the idea that Americans might hire "diseased Chinamen" to "manufacture cigars in underground localities in an atmosphere that reeks with deadly vapors," or to wash American clothes amidst slime.[28] Slowly gaining Americans' trust, Chinese men entered the labor market only to penetrate white homes and thrust into white bodies incurable disease.

With their twisted gender performance, having achieved access to Americans' most intimate moments, Chinese houseboys carried their inherent connection to Chinatown diseases into the bodies of the children in their care. The WPC pamphlet emphasizes the physical, bodily threat of Chinese men as a tool to urge readers to never forget their houseboys' true nature: "But we say to the housekeepers themselves: Follow your Chinese servant to his abiding place in Chinatown... There he lies down with five or six others... Can he be a safe companion who has slept with the leper, and in the infectious breath of syphilis and small pox? Bah!"[29] Here, the way Chinamen use their daytime docility to induce their employers into forgetting their roots serves to underscore their deceitful ways. The connection to gender remains, disease being characterized by the deceitfulness of a false man rather than the visibility of a fair fight, while warning Americans against letting Chinese laborers anywhere near their homes.

Chinese habits outside of their labor, furthermore, seduced white men into anti-capitalist activities that threatened America's health in the eyes of anti-Chinese activists. Despite the unity the white race required to prevent Chinese disease, opium's lure was strong enough to induce white men and women to enter Chinatown of their own accord, and the act of its smoking implied all the sexual and moral transgressions that middle-class manliness worked so hard to repress. While smoking opium, men passed pipes from mouth to mouth, thereby accepting the threat of physical contagion in their partaking of the Chinese social vice. The Anti-Chinese Council wrote that leprosy "is infused into our healthy race by the using, the sucking of opium- pipes, which have been handled by those already afflicted."[30] Consistently reminding readers of the physical disease smokers accepted into their mouths along with the drug, the Council created a tight knot between social downfall and physical degeneration.

While the smoking of opium itself, in the dizziness, paleness, and weight loss it often caused, brought men farther and farther from the manly physical ideal of bulky strength, the feminine mental state it created seemed just as sickly. The emphasis on sucking sets the white men in dependency to the Chinese providers of the drug, creating a reversal of the racial power hierarchy. Furthermore, due to the whites as masculine, Asians as feminine gender binary, this power reversal implies an even more unsettling inversion of gender roles. Viewers emphasized the economic effects

of this decline, warning it brought "idleness, debauchery, poverty." Weakening the American workforce, luring them to spend their hard-earned money in unproductive, degenerating ways, the Chinese habit of opium developed into yet another alarm anti-Chinese activists could ring in visualizing Chinamen's debilitating effect on economic progress.

Conclusion

Enflaming the population with the rhetorical tactics of disease and economic decay, the WPC and other anti-Chinese labor groups spearheaded efforts through the 1870s that led to riots, lynchings of Chinese immigrants, and ultimately anti-Chinese legislation. By seemingly stealing jobs out from underneath white men's feet during the difficult economic times in which they needed them most, Chinese men threatened the authority to which white men felt Social Darwinism had granted them the right. Therefore, while increasingly pursuing an athletic physical ideal and high moral values themselves throughout this period, white men cast inherent sickness into the Chinese body at the same time it feminized its portrayal. Chinese men transmitted their threat in feminine ways, through bowing and scraping to enter white families' good graces; through the mental seduction of mothers in the home and men in the opium den; through the voluntary adoption of feminine work; through deceit, sneakiness, and sleaze. The one way Chinese men never threatened white men was through a solid, upright fight. By intertwining the factors of gender, race, and illness, American anti-Chinese rhetoric attempted to separate Chinese men permanently from the ability to contribute productively to the American workforce, and, therefore, from the healthy American body as a whole.

References

1. Stout AB. "Chinese Immigration and the Physiological Causes of the Decay of a Nation." San Francisco: Agnew and Deffebach Printers, 1862:7. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2).
2. Bederman *G.Manliness and Civilization: A Cultural History of Gender and Race in the United States,1880-1917*. Chicago: The University of Chicago Press, 1995:27.
3. DiPiero T. *White Men Aren't*. Durham: Duke University Press, 2002:186.
4. Bederman G.*Manliness and Civilization: A Cultural History of Gender and Race in the United States, 1880-1917*. Chicago: The University of Chicago Press, 1995:25.
5. "The White Man's Burden." Detroit: *Literary Digest* 1923; 18:180.
6. Teng E. "Eurasian Hybridity in Chinese Utopian Visions: From 'One World' to 'A Society Based on Beauty' and Beyond." *East Asia Cultures Critique* 2006; 14:146.
7. "Chinese Immigration to the US." Golden Venture. 2007.
8. Swinton J. "The Chinese-American Question." New York: American News Company, 1870:4. Stanford Special Collections (Chinese Immigration Pamphlets, Box 2) .
9. Jung M. *Coolies and Cane*. Baltimore: Johns Hopkins University Press, 2006.
10. Norton HK. *The Story of California from the Earliest Days to the Present*. Chicago: A.C. McClurg & Co., 1924.
11. Hicks J. "Unrest in California." *The American Nation: A History of the United States from 1865 to the Present*. Boston: Houghton Mifflin, 1937. http://www.sfmuseum.org/hist6/kearneyism.html
12. Swinton J. "The Chinese-American Question." New York: American News Company, 1870, Stanford Special Collections (Chinese Immigration Pamphlets, Box 2) p. 10.
13. Stout AB. "Chinese Immigration and the Physiological Causes of the Decay of a Nation." San Francisco: Agnew and Deffebach Printers, 1862.
14. "The Modern "Old Man of the Sea."" *The Wasp* 1878. In The Online Archive of California, No. 77:408-409.
15. Grimm H. "The Chinese Must Go: A Farce in Four Acts." San Francisco: A.L. Bancroft & Co, 1879:19. In The Online Archive of California, Chinese in California Collection.
16. Lee, RG. *Orientals*. Philadelphia: Temple University Press, 1999: 85.
17. "The Servant Question." *The Wasp* 1882; 9:440-441.
18. "The Labor Agitators, or, The Battle for Bread." San Francisco, The Workingman's Party of California; 1879:25. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2, File 10).
19. Stout AB. "Chinese Immigration and the Physiological Causes of the Decay of a Nation." San Francisco: Agnew and Deffebach Printers, 1862:7.
20. Shah N. *San Francisco's "Chinatown": Race and the Cultural Politics of Public Health, 1854-1952*. Chicago: U Chicago, 1995: 158.
21. "Speech delivered by Francis Tracy Tobin Esquire, of the Philadelphia Bar, at an Anti-Chinese meeting of citizens, held Saturday Evening April 15, 1882, At Broad and Sansom Sts, Philadelphia, Pennsylvania." Philadelphia: John D Avil and Co., Printers and Publishers, 1884:7. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2, folder 19).
22. Memorial on Chinatown, by an Investigating Committee of the Anti-Chinese Council." San Francisco: The Workingman's Party of California, 1880:16. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2). p. 16.
23. Shah N. *Contagious Divides: Epidemics and Race in San Francisco's Chinatown*. Berkeley: University of California Press, 2001:99.
24. "Chinatown Declared a Nuisance!" San Francisco: The Workingman's Party of California, 1880. Stanford Rare Books Collection (Chinese Immigration Pamphlets; 2-14).
25. "Memorial on Chinatown, by an Investigating Committee of the Anti-Chinese Council." San Francisco: The Workingman's Party of California:13. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2).
26. "Article on Chinese Cigar Making." *The Wasp* 1879. In The Online Archive of California, No. 134:467.
27. "The Labor Agitators, or, The Battle for Bread." San Francisco, The Workingman's Party of California, 1879:26. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2, File 10).
28. "Memorial on Chinatown, by an Investigating Committee of the Anti-Chinese Council." San Francisco, The Workingman's Party of California:12. Stanford Special Collections (Chinese Immigration Pamphlets; Box 2).
29. Shah N. *Contagious Divides: Epidemics and Race in San Francisco's Chinatown*. Berkeley: University of California Press, 2001:96.
30. Stout AB. "Chinese Immigration and the Physiological Causes of the Decay of a Nation." San Francisco: Agnew and Deffebach Printers, 1862:24.

Zoe Leavitt is a senior from Wayland, Massachusetts, studying international relations with a focus on East Asia. After scaling the Great Wall, train-hopping from Mao Zedong's tomb to Confucius's hometown, and karaoke-ing her heart out while abroad in Beijing, she was thrilled to have the chance to explore Chinese immigration issues in this paper.

The significance of mixed-race: Public perceptions of Barack Obama's race and its effect on his favorability

Samuel Sinyangwe

This research paper seeks to understand white, black, and mixed-race Americans' perceptions of President Barack Obama's racial identity and the influence that those perceptions have on favorability towards the President. While political pundits and academic researchers tend to emphasize Obama's status as the first black President, this research finds that the Americans generally conceptualize Obama as mixed-race. A majority of whites and mixed-race Americans, and a third of blacks, conceptualize Obama as mixed-race. These Americans identify Obama as mixed-race to distinguish him from black people, interests, and values. Moreover, Obama's perceived race has political significance; whites generally respond more favorably to a perceived mixed-race Obama while blacks respond more favorably to a perceived black Obama. Mixed-race Americans are more likely to perceive Obama to be mixed-race and also more likely to support Obama, presumably because of group commonalities they share. The group closest to Obama's identity, mixed-race Americans who self-identify as black, respond most favorably towards Obama. These findings indicate that mixed-race Americans are actively supporting mixed-race public officials who share distinct commonalities with them. More importantly, the way most Americans conceptualize race appears to be changing dramatically to accommodate this emerging mixed-race population.

Introduction

Americans often treat race as a given, preassigned quality that defines an individual's identity. The meanings assigned to individuals with certain phenotypes and skin-tone shape how these individuals are treated in society. Contrary to popular notions, racial theorists Omi and Winant (1994) demonstrated that these racial meanings are not static; they are constantly evolving.[1] Some argue that the racial hierarchy is undergoing one of the most prominent changes since the birth of the nation; a transition from a hierarchy based on race to a tri-racial "pigmentocracy" that differentiates whites, lighter-skinned individuals, and darker-skinned individuals.[2] As the fastest growing youth population in the United States, "mixed-race" Americans are poised to play a significant role in challenging or reinforcing America's existing racial hierarchy.[3]

As one of the most prominent mixed-race individuals in the world, President Barack Obama provides a unique opportunity to gauge Americans' reactions to mixed-race public officials. I hope to gauge how perceptions of President Barack Obama's race affect Americans' support for him to better understand whether, and under what conditions, perceived mixed-race public officials are advantaged or disadvantaged relative to perceived black public officials.

Determinants of White Support for Black Public Officials

There is extensive research examining the role of race on American public opinion and political behavior, particularly concerning white support for black political candidates. Race is an important predictor of white voting behavior; whites generally do not support black political candidates or officeholders.[4,5,6,7,8] Researchers have found two distinct dimensions of prejudice against blacks: implicit prejudice and explicit prejudice. [9,10,11] Explicitly prejudiced whites consciously view any candidate with nonwhite skin-tone or nonwhite phenotypes as inherently inferior to a white candidate. [1,12,13]

Though some whites still express explicit racist opinions, most whites today consciously conceal or reject these beliefs because they violate the prevailing post-Civil Rights "norm of equality."[1,10,14] Instead, most whites have been found to hold unconscious, implicit anti-black attitudes that often diverge from more socially acceptable explicit attitudes.[14,15,16,17,18,19] These attitudes have been found to disproportionately affect individuals who most resemble anti-black stereotypes; whites have been found to attribute more negative qualities to people with more prototypically "black" features and to penalize candidates with more prototypically "black" features.[11,14,20,21,22,23,24]

The Complexities of Obama's Race: How do Whites Respond to Racial Ambiguity?

While researchers have found that white voters penalized Obama because of his race in 2008, they have not investigated how Obama's racial ambiguity shapes this discrimination.[13,25,26,27,28] Born to a black father and white mother, Barack Obama could conceivably be viewed as black, mixed-race, or even white.

Since traditional phenotypical racial cues are more ambiguous for mixed-race individuals, there is evidence that other stereotypical cues become more meaningful in defining these individuals. Studies have shown that stereotypical emotional expressions, hairstyles, and even dress disproportionately affect the race categorization of the racially-ambiguous. [21,29,30,31,32]

Experiments have found that partisan identifiers also help whites define racially-ambiguous candidates. Whites have been shown to "lighten" mixed-race candidates, including Barack Obama, that share their partisan affiliation and "darken" mixed-race candidates that identify with the other political party.[33]

While many whites attempt to categorize black-white biracials as either black or white, some have a more complicated language and perception of race that includes a mixed-race category (or categories). This continuous race perception develops from exposure to this population's racial, genetic and behavioral diversity.[34] While Malahy et al. (2010) have found these individuals to be more likely to perceive Obama to be mixed-race, it is unclear how Obama's perceived race relates to his perceived political identity or how it affects favorability towards him.[34]

Determinants of Black Support for Obama

Compared to research on white attitudes towards black public officials, research on black attitudes towards black officials, particularly Barack Obama, is thin. Black public opinion is distinct from white public opinion in America, reflecting these groups' starkly different sociohistorical experiences.[35] While black skin may be viewed as a stigma to many whites, blacks uniquely embrace a black identity and generally favor black representatives.[8,36] Most blacks perceive other blacks as members of a common group identity with a common political interest in overcoming the persistent racial and economic inequality that affects the group (black utility heuristic).[8,35] The more a candidate or representative is viewed as sharing this common black identity or interest, the more likely blacks are to support them. While blacks overwhelmingly approve of President Obama and believe he shares their group interests, it is unclear whether perceptions of Obama's race as either mixed-race or black affect blacks' perceptions of his group loyalties. Since traditional racial norms would suggest that blacks should view Obama as black, members of the black community classifying Obama as mixed-race may be an indication of otherness, inauthenticity, or group difference.

Theory

The purpose of this research is to investigate how different racial groups perceive President Obama's race. I will then evaluate whether, and under what conditions, President Obama's perceived racial identity as either black or mixed-race affects favorability towards him.

Perceptions of Obama's Race

A. Individuals who perceive Obama to differ from the political interests and cultural values of blacks will be more likely to perceive Obama to be mixed-race.

Eduardo Bonilla-Silva (2004) theorizes that the American society is approximating a three-tiered "pigmentocracy" defined by color rather than race.[2] In this tri-racial hierarchy, most mixed-race Americans form a privileged "honorary white" racial class between whites and the "collective black."[2] According to this theory, black-white biracials differentiate themselves from blacks and the stigma associated with black identity.[2] In this light, I expect whites, blacks, and mixed-race individuals to distinguish Obama as mixed-race to the extent that they perceive Obama to differ from their conception of black group identity. Since black identity contains political and cultural dimensions, people who perceive Obama to share common racial identity with blacks should be more likely to assume that he shares black political interests and cultural values. Conversely, people who differentiate Obama from blacks by identifying him as mixed-race should be more likely to differentiate his interests and values from those common to the black community.

B. Mixed-race Americans will be more likely than monoracial respondents to identify Obama as mixed-race.

The tri-racial hierarchy theory presumes that black-white biracials will become part of an "honorary white" group identity that includes other mixed-race Americans.[4] As such, black-white biracials and other mixed-race Americans should conceptualize each other as members of the same group. Indeed, there are important distinctions between mixed-race Americans and other racial groups. Mixed-race Americans have been found to differ socioeconomically, politically, and culturally from individuals who identify monoracially.[2,3,34,37] In this light, mixed-race individuals should be more likely to distinguish Obama as mixed-race because they recognize distinct group commonalities they share with him.

C. Individuals in more diverse contexts will be more likely to perceive Obama to be mixed-race.

Perceptions of Obama's race may also be influenced by race contact.38 Individuals who live in more racially diverse contexts should be more likely to develop a language and understanding of the racial diversity around them. These individuals should be more likely to perceive potential distinctions between most blacks and Obama. Therefore, I expect individuals in more diverse contexts to be more likely to perceive Obama to be mixed-race.

Favorability towards Obama

A. Whites will generally favor a perceived mixed-race Obama.

The tri-racial hierarchy theory assumes that black-white biracials will differentiate from the negative stigma associated with black identity, joining a more privileged "honorary white" racial stratum with other mixed-race Americans.[2] White reactions to Obama's perceived race should therefore depend on the stigma they associate with black identity. In this light, whites with more anti-black prejudice should respond more favorably to a perceived mixed-race Obama because they are less likely to associate him with these negative stereotypes.

B. Whites who are less likely to perceive anti-black discrimination in society will be more likely to favor a perceived mixed-race Obama.

Whites that recognize the racial disadvantages that blacks face should be more likely to reject anti-black stereotypes because these stereotypes attribute blacks' economic disadvantage to cultural or biological deficiencies.[1] As such, this subset of whites should be less likely to stigmatize a black Obama relative to a mixed-race Obama.

C. Blacks will generally favor a perceived black Obama.

The tri-racial hierarchy presumes that blacks will remain supportive of public officials that conform to the interests and norms of their group.[2] Since blacks who perceive Obama to be mixed-race are expected to be less likely to believe he shares these interests, blacks should be less likely to support a mixed-race Obama.

D. Mixed-race respondents will favor a perceived mixed-race Obama.

Since mixed-race respondents are expected to distinguish Obama as mixed-race because they perceive him to share distinct group commonalities with them, mixed-race respondents should respond more favorably to a perceived mixed-race Obama.

Data and Measures

This research has two stages: a logistic regression model to predict cross-racial perceptions of Obama's race and then linear regression to predict the determinants of respondents' favorability towards Obama. For the purposes of this research, I used

survey data from the Pew Research Center's Racial Attitudes in America II Survey (N=2884, Whites N=1447; Blacks N=812; Hispanics N=376; Mixed N=484).[39] The survey was conducted between October 28 and November 30, 2009, during the end of President Barack Obama's first year in office. These data are useful because they explore sizeable samples of whites, blacks and mixed-race Americans' racial and political attitudes towards President Obama. Most importantly, these data uniquely measure respondents' perceptions of Obama's racial identity.[1] This Pew survey is the only survey publically available that combines all of these features, making this research possible.

Unlike most surveys that omit a mixed-race option, the Pew survey includes two measures of mixed-race self-identification. First, respondents were asked, "Do you consider yourself to be of mixed race?" Second, respondents were asked to choose between one or more traditional racial categories, making it possible to locate respondents who generally identify as monoracial but also identify as mixed-race when prompted. In the absence of a parental ancestry measure, however, mixed-race responses will remain relatively imperfect. Nevertheless, mixed-race respondents' decision to identify themselves as mixed-race should affect how this group conceptualizes a racially-ambiguous President's race and should make their evaluations of him informative both of how mixed-race Americans evaluate Obama and also how they view themselves.

Dependent Variables

Dependent variables are perceptions of Obama's race as either mixed-race or black and favorability towards Obama.

Independent Variables

To determine whether feelings of group difference influenced respondents' views of Obama, I constructed a variable measuring to what extent respondents felt that Obama "shared the values and interests of black people." I also included the interaction of Obama's shared interests and values with blacks and perceptions of Obama's race to examine whether Obama's perceived race shapes how his connections to black interests and values are weighed in respondents' calculations of favorability.

To gauge racial attitudes, I measured respondents' perceptions of anti-black discrimination in society. I then constructed an interaction of Obama's perceived race and perceived anti-black discrimination to determine how respondents' racial attitudes condition reactions to Obama's perceived race. Next, I measured whether respondents considered themselves to be mixed-race and whether respondents lived in diverse contexts.

Finally, I included controls for age, gender, education, income, Republican, Independent, South, political knowledge, and political ideology. Each variable was included because it reflected an important potential driver of respondents' attitudes towards Obama. All relevant question wording is included in Appendix A.

Results and Analysis

Table 1 Distribution of respondents' perceptions of Obama's race

Perceived Race	Whites (N=1,122)	Mixed-Race (N=381)	Blacks (N=623)
Mixed-Race	66.8%	64.3%	37.6%
Black	33.2%	35.7%	62.4%

Intergroup differences among respondents are clearly noticeable in Table 1; mixed-race and white Americans were much more likely to perceive Obama to be mixed-race compared to blacks. The proportions of groups' responses are almost perfectly polarized; about two-thirds of whites and mixed-race Americans conceptualized Obama as mixed-race, while 62% of black respondents conceptualized him as black.

Next, I estimated a logistic regression predicting perceptions of Obama's race. The most important finding was that respondents' perceptions of Obama's race were strongly related to their perceptions of Obama's racial interests and allegiances. Groups' perceptions that Obama shared common interests and values with black people were the strongest predictors of how those groups conceptualized Obama's racial identity; individuals and groups that were less likely to believe that Obama shared the values and interests of blacks were significantly more likely to identify Obama as mixed-race. Conversely, Obama's perceived race best predicted his perceived connection to black interests and values for all three groups (See Appendix C).

Compared to monoracial respondents, mixed-race respondents were generally more likely to distinguish Obama's mixed-race identity just as they have ac-

	Model 1	Model 2	Model 3
(Intercept)	1.913***	2.187***	2.327***
	(0.233)	(0.306)	(0.475)
Obama.shares.black.interests	−0.470***	−0.357***	−0.328***
	(0.069)	(0.072)	(0.080)
respondent.mixedrace		0.176*	0.204**
		(0.068)	(0.078)
respondent.black		−0.783***	−0.984***
		(0.141)	(0.180)
perceive.antiblack.discrimination		−0.218*	−0.188*
		(0.085)	(0.094)
diverse.context		−0.076	0.076
		(0.213)	(0.233)
Republican			−0.095
			(0.210)
Independent			0.122
			(0.173)
education			−0.074
			(0.051)
age			−0.002
			(0.004)
South			0.001
			(0.151)
N	1181	1135	949
AIC	1542.309	1430.625	1213.063
BIC	1582.902	1571.587	1504.387
$\log L$	−763.155	−687.312	−546.531

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

*Some variables were included in Model 3 but were not significant and are not displayed here: see Appendix B for full table

Table 2 Factors that predict cross-racial perceptions of Obama as mixed-race

	Blacks	Whites	Mixed-Race
(Intercept)	3.151***	3.530***	3.708***
	(0.183)	(0.197)	(0.286)
mixedrace.obama	0.341	0.533*	0.129
	(0.209)	(0.236)	(0.360)
Obama.shares.black.interests	0.153***		
	(0.045)		
Obama.shares.black x mixedrace.obama	−0.123*		
	(0.058)		
perceive.antiblack.discrimination		0.224**	0.103
		(0.068)	(0.094)
perceive.antiblack x mixedrace.obama		−0.180*	−0.097
		(0.081)	(0.115)
respondent.mixedrace	0.069**	0.097*	
	(0.025)	(0.038)	
Republican	−0.404***	−1.103***	−1.029***
	(0.109)	(0.079)	(0.146)
Independent	−0.299***	−0.544***	−0.406***
	(0.060)	(0.072)	(0.105)
conservatism		−0.299***	−0.056
		(0.031)	(0.047)
education	−0.035*	0.018	0.006
	(0.017)	(0.020)	(0.032)
female	0.186***	0.149**	0.092
	(0.050)	(0.056)	(0.090)
N	344	969	297
R^2	0.207	0.390	0.288
adj. R^2	0.181	0.382	0.258
Resid. sd	0.444	0.853	0.762

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

*Some variables were included but were not significant and are not displayed here; see Appendix B for full table

Table 3 Factors that predict favorability towards Obama

knowledged their own. There were important differences between the perceptions of respondents who identified as black (N=164) and those who did not (N=240). Compared to mixed-race respondents who did not identify as black, mixed-race respondents who identified primarily as black were more likely to identify Obama as black just as they have chosen to identify themselves.

At the same time, mixed-race respondents who identified as black were about 18% more likely than blacks to consider Obama mixed-race. By being more likely to acknowledge mixed-race identity, these mixed-race respondents are challenging the more traditional racial categorization norms of their minority half.

In addition to mixed-race self-identification, groups' perceptions of anti-black discrimination in society shaped perceptions of Obama's race. Black, white, and mixed-race respondents who perceived higher levels of anti-black discrimination in society were more likely to perceive Obama to be black. These race-conscious individuals perceive the force that continues to define many black-white biracials as black: being subjected to anti-black discrimination in society.[37] Even Barack Obama himself attributes his decision to identify as black to his experience of being treated as a black man in America. Without understanding this important component of Obama's "black" experience, it may be more intuitive to identify Obama as mixed-race.

Contrary to my hypothesis, respondents who lived in more diverse contexts were not more likely to distinguish Obama as mixed-race. Though these respondents likely have more contact with blacks and mixed-race individuals, contact theorists point out that contact with an out-group can either encourage opinion change or reinforce existing preconceptions depending on a number of contextual factors.[40] White respondents in poorer contexts, for example, may encounter heavy competition for resources from both blacks and mixed-race individuals, discouraging these respondents from distinguishing between groups.

Interestingly, partisanship shaped perceptions of Obama's race among whites.[2] Though white Democrats were significantly more likely than white Republicans to believe that Obama shared the values and interests of blacks, they were about 4% less likely than white Republicans (32%) to perceive Obama to be black. Obama's perceived racial interests predicted white Republicans' perceptions of Obama's race but did not predict white Democrats' perceptions (See Appendix C). In other words, white Democrats tended to depict Obama as mixed-race, whether or not they believed Obama to share black interests and values.

Similarly, while white Republicans were significantly less likely than mixed-race Americans to believe that Obama shared the values and interests of blacks, they were still about equally as likely (36%) to perceive Obama to be black (see Appendix C). These findings indicate that there are factors in addition to the perception of Obama's racial interests that influence white Republicans to perceive Obama to be black and influence white Democrats to perceive Obama to be mixed-race.

There are three possible explanations for these interparty differences in Obama's perceived racial identity. White Democrats could call Obama mixed-race because they want to promote a "post-racial" ideal that rejects traditional racial categorization while white Republicans hold more traditional notions of racial categorization (i.e., the one-drop rule).[3,41] It is also possible that party identification itself influences whites to view greater racial commonality between co-partisans (the in-group) and greater racial difference between members of the other party (the out-group).

Alternatively, the way that Obama's race and racial interests are communicated to whites could differ by party in an era of polarized politics. There is evidence that conservative media outlets and political figures, on the one hand, exaggerate Obama's "black" attributes and, on the other hand, frame Obama as a socialist or radical outside the mainstream American political consensus, black or white[1]. This could conceivably lead white Republicans to conceptualize Obama as both black and too extreme to share most blacks' interests and values.

I estimated a linear regression predicting favorability towards Obama. As hypothesized, white respondents generally favored a perceived mixed-race Obama. Whites' racial attitudes conditioned responses to Obama's race: whites that perceived the most anti-black discrimination in society favored a perceived black Obama. By recognizing the racial disad-

vantages that blacks face, this small subset of whites likely rejects anti-black stereotypes and supports black public officials as a means to address racial inequality.

Among blacks, Obama's perceived race appears to act on favorability indirectly by affecting the salience of Obama's racial credentials. Specifically, blacks are less likely to translate Obama's perceived commonalities with blacks into support for the President when they perceive him to be mixed-race. In other words, Obama's perceived connection to black interests may become more consequential to blacks' political evaluations of Obama when they also perceive him to be black.

Mixed-race respondents, by contrast, generally responded favorably towards Obama. Interestingly, mixed-race favorability was not contingent on Obama's perceived race. There are two plausible explanations for this finding. Mixed-race individuals could value aspects of Obama's personality, politics, or ancestry that are particularly relevant to the mixed-race experience without attributing these factors to Obama's mixed-race identity. Alternatively, mixed-race respondents might be favorable towards Obama whether they identify him as black or mixed-race because they do not conceptualize mixed-race identity exclusively. Rather, this group might believe that Obama can identify as both mixed-race and black without compromising either identity, a view that creates critical space for mixed-race individuals to embrace both sides of their heritage as they see fit.

Favorability towards Obama differed by racial combination or subgroup of mixed-race respondents. Mixed-race respondents who self-identified as more than one race responded more favorably than those who self-identified as white, while mixed-race respondents who self-identified as black responded more favorably to Obama than other mixed-race respondents. These findings support past research on mixed-race attitudes that has found mixed-race individuals who identify as black to exhibit more liberal preferences than other mixed-race individuals.[37,42] It also suggests that the decision to identify as more than one race is associated with political opinions that are in-between white and black.

Given extraordinarily high support for Obama among blacks, it was particularly significant that mixed-race respondents self-identified as black responded more favorably towards Obama than blacks. Mixed-race respondents who self-identify as black may be more likely to support Obama because they perceive Obama, who is also a mixed-race individual self-identified as black, to share commonalities with them. Though mixed-race Americans broadly support Obama, this finding demonstrates that mixed-race individuals also feel a particular group connection to those who share their racial background.

Conclusion

This research sought to explain cross-racial perceptions of Barack Obama's race and the influence that those perceptions have on favorability towards the President. Contrary to those who still conceptualize race in terms of black and white, this research establishes that a majority of whites and mixed-race Americans, and a third of blacks, conceptualize President Barack Obama as mixed-race. These Americans distinguish Obama as mixed-race to express his perceived difference from black identity, interests, and values.

This perceived racial difference clearly mattered to respondents. Consistent with the tri-racial hierarchy theory, this research finds that the dominant white population privileges Obama's mixed-race identity while the minority black population stigmatizes it. This research also finds that mixed-race individuals are embracing mixed-race identity and a mixed-race President. While these developments indicate that the American racial hierarchy has become more complex, they do not necessarily imply the preservation of white dominance. While mixed-race individuals generally supported Obama, those who were closest to Obama's identity (those who self-identified as black) responded most favorably towards him. This indicates that there remain important divisions within the mixed-race population based on racial background and identification that challenge the formation of a cohesive mixed-race or "honorary white" political identity and interest. Furthermore, mixed-race Americans are not merely conforming to white norms or rejecting black norms by supporting a mixed-race President. Instead, mixed-race Americans are challenging America's dominant racial paradigm with a new, multiracial conception of race that allows individuals to choose to embrace the particular political or cultural aspects of one racial identity without compromising other identities.

There are several areas where this research can be further developed. More precise measures of race perception (i.e., Implicit Categorization Task), implicit prejudice (i.e., Implicit Association Test), and explicit prejudice (i.e., Modern Racism Scale) would improve this research. Subsequent research should also explore the effect of Obama's perceived race on perceptions of Obama's political priorities and on vote choice. Given the finding that partisanship shapes white perceptions of Obama's race, further research and experimentation should explore the relationship between party identification, perceptions of Obama's race, and political support from all racial groups.

Since Obama's connection to black interests and his race were so closely related across racial groups, future experiments should be done to disentangle the relationship between them. It could be that Obama's perceived race drives how Americans perceive his racial interests or, alternatively, Obama's perceived racial interests could determine his race. In addition to experiments, one way to explore this relationship would be to track Obama's perceived racial identity over time as perceptions of Obama's support or opposition to black priorities change.

References

1. Omi M, Winant H. Racial Formation in the United States: From the 1960s to the 1990s. New York: Routledge, 1994.
2. Bonilla-Silva E. From bi-racial to tri-racial: towards a new system of racial stratification in the USA". Ethnic and Racial Studies 2004; 27(6):931-950.
3. Masuoka N. "The 'Multiracial' Option: Social Group Identity and Changing Patterns of Racial Categorization". American Politics Research, 2011.
4. Canon DT. Race, Redistricting, and Representation: The Unintended Consequences of Black-Majority Districts. Chicago: University of Chicago Press, 1999.
5. Hajnal, Z. Changing White Attitudes toward Black Political Leadership. Cambridge [England] ; New York: Cambridge University Press, 2007.
6. Highton, B. "White Voters and African American Candidates for Congress." Political Behavior, 2004; 26 (1): 1-25.
7. Lublin, D. The Paradox of Representation: Racial Gerrymander and Minority Interests in Congress. Princeton, N.J.: Princeton University Press 1997.
8. Tate, K. Black Faces in the Mirror. Princeton: Princeton University Press, 1993.
9. Davis D, Wilson D. "The Measurement of Racial Resentment: Conceptualization and Content." 2008.
10. Kinder D, Sears D. "Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life." Journal of Personality and Social Psychology, 1981; Vol. 40, No. 3, 414-431.
11. Livingston RW. The role of perceived nega-

tivity in the moderation of African Americans' implicit and explicit racial attitudes. *Journal of Experimental Social Psychology*, 2002; 38 (4), 405-413.
12. Carmines E, Stimson J. "Issue Evolution." Princeton, N.J., Princeton University Press, 1989.
13. Schaffner, B. "Racial Salience and the Obama Vote." Amherst, Ma: University of Massachusetts, 2011.
14. Jabon M, Plaut E, Messing S. Bias in the Flesh: Attack Ads and the Effects of Visual Cues in the 2008 Presidential Campaign. Paper presented at the International Communication Association annual conference in Singapore, 2010.
15. Bargh JA, Chen M, Burrows L. Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. Journal of Personality and Social Psychology 1996; 71:230–244.
16. Dovidio JF, Kawakami K, Gaertner SL. Implicit and explicit prejudice and interracial interaction. Journal of Personality and Social Psychology, 2002; 82 (1), 62-68.
17. Greenwald AG, Bana ji MR. Implicit social cognition: Attitudes, self-esteem, and stereotypes. Psychological Review, 1995; 102 (1), 4–27.
18. Klatzky R, Martin G, Kane R. Influence of social-category activation on processing of visual information. Social Cognition, 1982; 1, 95–109.
19. Spencer SJ, Fein S. Wolfe CT et al. Automatic activation of stereotypes: The role of self-image threat. Personality and Social Psychology Bulletin, 1998; 24, 1139–1152.
20. Bailenson J, Hahn K., Iyengar S et al. Do Explicit Racial Cues Influence Candidate Preference? The Case of Skin Complexion in the 2008 Campaign. 2010.
21. Blair I, Jenkins C. The role of Afrocentric features in person perception: Judging by features and categories. Journal of Personality & Social Psychology 2002; 83(1):5-24.
22. Gray SA, Maddox KB. Cognitive representations of Black Americans: Re-exploring the role of skin tone. Personality and Social Psychology Bulletin, 2002; 28 (2), 250-259.
23. Terkildsen N. "When White Voters Evaluate Black Candidates: The Processing Implications of Candidate Skin Color, Prejudice, and Self-Monitoring." American Journal of Political Science, 1993; 37: 1032-1053.
24. Weaver V. "Race, Skin Color, and Candidate Preference." Paper presented at the 2005 Annual Meeting of the Midwest Political Science Association, Chicago, IL.
25. Knowles E, Lowery B, Schaumberg R. "Racial Prejudice predicts opposition to Obama and his health care reform plan." Journal of Experimental Social Psychology, 2009; Volume 46, Issue 2, Pages 420-423.
26. Mas A, Moretti E. "Racial Bias in the 2008 Presidential Election." American Economic Review, 2009; 99 (2): 323-329.
27. Pasek J, Tahk A, Lelkes Y et al. "Determinants of Turnout and Candidate Choice in the 2008 U.S. Presidential Election: Illuminating the Impact of Racial Prejudice and Other Considerations." Public Opinion Quarterly, 2009 73(5): 943-94.
28. Sears D, Tesler M. Obama's Race: The 2008 Election and the Dream of a Post-Racial America. Chicago: University of Chicago Press, 2010.
29. Bodenhausen G, Hugenberg K.. Ambiguity in social categorization: The role of prejudice and facial affect in racial categorization. Psychological Science 2004; 15, 342-345.
30. Bodenhausen G, Macrae CN. Stereotype activation and inhibition. In Wyer, J. R. (Ed.) Advances in social cognition. Mahwah, NJ: Erlbaum, 1998:1–52.
31. Freeman JB, Penner AM, Saperstein, A et al. Looking the Part: Social Status Cues Shape Race Perception. PLoS ONE, 2011; 6(9): e25107.
32. MacLin OH, Malpass RS. Racial Categorization of Faces: The Ambiguous Race Face Effect. Psychology, Public Policy and Law, 2001; 7(1), 98-118.
33. Caruso E, Mead N, Balcetis E. Political partisanship influences perception of biracial candidates' skin tone. PNAS, 2009; 106;20168-20173.
34. Malahy LW, Sedlins M, Plaks J, et al, Black and White, or Shades of Gray? Racial Labeling of Barack Obama Predicts Implicit Race Perception. Analyses of Social Issues and Public Policy, 2010; 10: 207–222.
35. Dawson M. Behind the Mule: Race and Class in African-American Politics. Princeton: Princeton University Press, 1994.
36. Darity W, Dietrich J, Hamilton D. Passing on Blackness: Latinos, Race, and Earnings in the USA." Applied Economics Letters, 2002; 9:847-853.
37. Davenport L. Politics Between Black and White: The Identity and Political Attitudes of Biracial Americans. 2011
38. Malahy LW, Sedlins M, Plaks J et al. Folk Beliefs About Human Genetic Variation Predict Discrete Versus Continuous Racial Categorization and Evaluative Bias. Social Psychological and Personality Science, January 2012: 3: 31-39.
39. Pew Research Center (2009). "Racial Attitudes in America II" project (N = 2,850).
40. McClain et al."Racial Distancing in a Southern City: Latino Immigrant Views of Black Americans." Journal of Politics, 2006; 68, 3: 571-584.
40. Davis J. "Who is Black? One Nation's Definition." University Park: The Pennsylvania State University Press, 1991.
41. Masuoka, N. "Political Attitudes and Ideologies of Multiracial Americans: The Implications of Mixed-Race in the United States". Political Research Quarterly, 2008; 61(2): 253-267.

Appendices
Appendix A: Question Wording

Dependent Variable: mixedrace.obama
Do you mostly think of Obama as?
1. A black person [OR mostly as]
2. A person of mixed race

Dependent Variable: favorability towards Obama
"Now I'd like your views on some people. As I read some names, please tell me if you have a favorable or unfavorable opinion of each person. First, Barack Obama would you say your overall opinion of Barack Obama is very favorable, mostly favorable, mostly unfavorable, or very unfavorable?"

Independent Variable: Obama.shares. black.interests
"How much would you say Barack Obama shares the values and interests of black people in this country?"

Independent Variable: perceive.antiblack. discrimination
Please tell me how much discrimination there is against each of these groups in our society today. How about Blacks? Would you say there is a lot of discrimination, some, only a little, or none at all?

Independent Variable: racial.progressivism
"Which of these two statements comes closer to your own views, even if neither is exactly right.
1. Our country has made the changes needed to give blacks equal rights with whites 2. Our country needs to continue making changes to give blacks equal rights with whites"

Independent Variable: diverse.context
"Thinking about the neighborhood where you live, are all, most, some, or only a few people (if respondent white, then black; if respondent black or more than one race, then white)?"

Independent Variable: political.knowledge
Index ranging from 0-2, combining three questions:
1. And thinking about the news…do you or not? Next, do you regularly read a daily newspaper or not?
2. And thinking about the news…do you or not? Next, do you regularly watch the local television news about your area or not?
3. And thinking about the news…do you or not? Next, do you regularly watch the national news on major network or cable channels or not?

Appendix B: Full tables of Statistical Analyses

Table 4: Logistic regression predicting perceptions of Obama as mixed-race

Table 5: Linear regression predicting respondents' favorability towards Obama

Appendix C: Additional Statistical Analyses

Table 6: Linear regression predicting the perception that Obama shares the values and interests of blacks

Table 7: Mixed-race respondents who self-identify as black are more likely to perceive Obama to be black

Table 8: The effect of partisanship on white perceptions that Obama shares black interests and values

Table 9: Logistic regression predicting perceptions that Obama is mixed-race among whites

Table 10: Connection between partisanship and race perception among whites

	Model 1	Model 2	Model 3
(Intercept)	1.913***	2.187***	2.370**
	(0.233)	(0.306)	(0.484)
Obama.shares.black.interests	−0.470***	−0.357***	−0.348**
	(0.069)	(0.072)	(0.081)
respondent.mixedrace		0.176*	0.205**
		(0.068)	(0.078)
respondent.black		−0.783***	−0.996**
		(0.141)	(0.181)
respondent.other.minority		0.278	−0.029
		(0.196)	(0.226)
perceive.antiblack.discrimination		−0.218*	−0.192*
		(0.085)	(0.094)
diverse.context		−0.076	0.067
		(0.213)	(0.234)
Republican			−0.114
			(0.211)
Independent			0.112
			(0.173)
conservatism			0.113
			(0.074)
education			−0.073
			(0.051)
income			−0.082†
			(0.047)
age			−0.002
			(0.004)
female			0.104
			(0.143)
South			0.016
			(0.151)
political.knowledge			0.020
			(0.078)
N	1181	1135	946
AIC	1542.309	1430.625	1208.721
BIC	1582.902	1571.587	1519.265
$\log L$	−763.155	−687.312	−540.361

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 4

	Blacks1	Blacks2	Whites1	Whites2	Mixed-Race1	Mixed-Race2
(Intercept)	3.796***	3.151***	2.553***	3.530***	3.221***	3.708***
	(0.026)	(0.183)	(0.057)	(0.197)	(0.090)	(0.286)
mixedrace.obama	−0.099*	0.341	0.136*	0.533*	−0.163†	0.129
	(0.042)	(0.209)	(0.070)	(0.236)	(0.097)	(0.360)
respondent.mixedrace		0.069**		0.097*		
		(0.025)		(0.038)		
perceive.antiblack.discrimination				0.224**		0.103
				(0.068)		(0.094)
perceive.antiblack x mixedrace.obama				−0.180*		−0.097
				(0.081)		(0.115)
obama.shares.black.interests		0.153***				
		(0.045)				
obama.shares.black x mixedrace.obama		−0.123*				
		(0.058)				
Republican		−0.404***		−1.103***		−1.029***
		(0.109)		(0.079)		(0.146)
Independent		−0.299***		−0.544***		−0.406***
		(0.060)		(0.072)		(0.105)
conservatism				−0.299***		−0.056
				(0.031)		(0.047)
education		−0.035*		0.018		0.006
		(0.017)		(0.020)		(0.032)
income		0.028†		−0.033†		−0.025
		(0.016)		(0.018)		(0.029)
age		0.001		−0.001		−0.004
		(0.001)		(0.002)		(0.003)
female		0.186***		0.149**		0.092
		(0.050)		(0.056)		(0.090)
South		0.087†		−0.108†		0.017
		(0.050)		(0.060)		(0.099)
self-identified.black					0.599***	0.365***
					(0.096)	(0.101)
N	614	344	1084	969	365	297
R^2	0.009	0.207	0.003	0.390	0.119	0.288
adj. R^2	0.007	0.181	0.003	0.382	0.114	0.258
Resid. sd	0.506	0.444	1.088	0.853	0.866	0.762

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 5

	Model 1	Model 2	Model 3
(Intercept)	3.400***	3.246***	3.078***
	(0.044)	(0.053)	(0.172)
mixedrace.obama	−0.406***	−0.308***	−0.259***
	(0.057)	(0.058)	(0.063)
respondent.mixedrace		−0.053†	−0.037
		(0.028)	(0.032)
respondent.black		0.356***	0.228**
		(0.060)	(0.076)
perceive.antiblack.discrimination			0.030
			(0.039)
diverse.context			0.038
			(0.099)
Republican			−0.382***
			(0.089)
Independent			−0.205**
			(0.073)
conservatism			0.012
			(0.032)
education			−0.010
			(0.021)
income			0.032
			(0.020)
age			0.002
			(0.002)
female			0.038
			(0.061)
South			0.049
			(0.064)
N	1181	1161	951
R^2	0.041	0.070	0.090
adj. R^2	0.041	0.067	0.078
Resid. sd	0.959	0.943	0.919

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 6

	Mixed-Race
(Intercept)	3.046***
	(0.664)
obama.shares.black.interests	−0.578**
	(0.190)
self-identified.black	−0.782*
	(0.347)
more.than.one.race	−0.650
	(0.540)
N	196
AIC	230.867
BIC	283.317
$\log L$	−99.434

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 7

	Model 1
(Intercept)	3.227**
	(0.072)
Republican	−0.395**
	(0.099)
Independent	−0.275**
	(0.095)
N	675
R^2	0.024
adj. R^2	0.021
Resid. sd	1.012

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 8

	White Democrats	White Republicans
(Intercept)	3.189*	3.823**
	(1.430)	(1.409)
Obama.shares.black.interests	−0.290	−0.461*
	(0.264)	(0.183)
diverse.context	−0.605	−0.767
	(0.799)	(0.720)
perceive.antiblack.discrimination	0.049	−0.026
	(0.307)	(0.255)
conservatism	0.139	0.100
	(0.225)	(0.253)
education	−0.312†	0.103
	(0.189)	(0.132)
income	0.015	−0.320*
	(0.129)	(0.139)
age	0.009	−0.021†
	(0.012)	(0.011)
female	−0.535	0.516
	(0.446)	(0.368)
South	−0.799†	−0.458
	(0.435)	(0.371)
N	127	156
AIC	166.637	206.084
BIC	280.405	328.079
$\log L$	−43.319	−63.042

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 9

	White Respondents
(Intercept)	0.814***
	(0.080)
Republican	−0.337*
	(0.134)
N	1099
AIC	1394.965
BIC	1434.982
$\log L$	−689.483

Standard errors in parentheses

† significant at $p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$

Table 10

Samuel Sinyangwe is a senior political science major with a concentration in American Government. Sam is fascinated by the way race shapes our lives and our politics. A multiracial person himself, he finds it particularly interesting to examine how other multiracial people negotiate racial identity in the midst of social pressure to 'pick a side.' Sam is currently studying abroad in Cape Town, South Africa, where he works with disadvantaged young men to help them create opportunities to overcome the barriers they face. After graduation, he hopes to continue complicating what he sees to be an outdated, binary racial discourse to more closely reflect the racial complexity of 21st century America.

United States drug policy: The scientific, economic, and social issues surrounding marijuana

Matt Winterbourne

The United States currently enforces one of the harshest drug policies in the world. This policy has been directly linked to increases in nonviolent incarceration rates and racial profiling by law enforcement. This report examines the current "War on Drugs" in terms of three issues: the scientific findings on marijuana regarding its medical viability, lack of addictive properties, minimal behavioral and health consequences, and zero annual death-rate; the unsustainable economic costs of enforcement, convictions and incarceration compared to treatment and rehabilitation programs; and the social implications of racially targeted enforcement efforts directly contributing to minority over-representation in stop-searches, arrests, convictions, and incarcerations for nonviolent, drug-related offenses. This meta-analysis of the War on Drugs culminates with a request to revisit U.S. policies towards marijuana, to refocus drug enforcement on treatment for drug-related crime, addiction, and abuse, and to re-educate law enforcement officials to reduce the drastic racial disparities currently present in arrest rates through an understanding of drug usage across ethnicities.

Effects of the Current Drug Policy

The United States government currently implements one of the world's harshest programs related to the use of illicit substances—the "War on Drugs." Arrests, convictions, and incarceration rates have skyrocketed in the U.S. since the implementation of this policy in the 1980s. Despite only comprising 5% of the global population[1], the U.S. currently holds one-quarter of the world's incarcerated population[2], largely due to arrests for nonviolent, drug-related offenses. Additionally, there is striking evidence of discriminative enforcement against African Americans and Hispanic Americans, who currently account for over 85% of arrestees in some states[3]. Non-white drug offenders also account for nearly 75% of all drug offenders in U.S. prisons[4]. The increased arrest rates, targeting of minority citizens, and immense allocation of federal spending for the War on Drugs necessitates an examination of its legitimacy.

A comprehensive review of scientific research on marijuana, economic models of the costs and benefits of drug prohibition, and writings on the social effects of the War on Drugs reveal that this policy is based on false premises and misrepresented scientific findings, is an unnecessary drain on the United States economy, and is resulting in unconstitutional and excessive legal sanctions against U.S. Hispanic and African American minorities. While this report will not offer a formal outline of necessary drug policy changes, this analysis of the War on Drugs will address its current implications for American society.

The first portion of this report offers an examination of scientific findings on marijuana. These findings undermine the premises of the War on Drugs policies regarding marijuana and debunk common misconceptions regarding marijuana's "gateway" effects on drug use and lethality. The analysis pulls from scientific findings and laboratory tests performed by a variety of national and private institutions; some of which come from committees appointed by Presidents Richard Nixon and Ronald Reagan in the 1970s and 1980s.

Next, an analysis of the economic costs of enforcement, convictions, and incarcerations will demonstrate the War on Drugs' excessive drain on the economy. Support for this claim will be drawn from data presented by the first director of the National Institute on Drug Abuse on costs of the War, as well as a study conducted by Michael Grossman and colleagues on the elasticity of drug consumption based on price since the 1970s. The impact of the reduced instance of court-ordered rehabilitation brought on by the War on Drugs' zero-tolerance, minimum sentencing, and three-strike policies is explored and compared with the implications of the current Dutch policy regarding marijuana.

Finally, a review of drug-related arrests, convictions, and incarceration records demonstrates stark evidence of racial discrimination against Hispanics and African Americans. The analysis of the demographics of drug users, arrestees, and convicts will include quantitative data from government self-report surveys on use, as well as private studies conducted on drug-testing results in the military and the workplace. Included in this analysis are self-report data collected by the National Institute on Drug Abuse from various middle and high schools throughout the United States. Arrest rates will be drawn from multiple studies on the law enforcement, including both findings from public records and personal testimonies from judges and police officers.

The investigation culminates in a discussion of U.S. drug policy and identification of necessary changes in the current policy regarding federal marijuana sanctions to more accurately reflect the scientific data regarding the drug, to increase alternative punishments for nonviolent drug crimes and consequently reduce both economic costs and the instance of drug-use relapse, and to better educate law enforcement officials to combat racial profiling. This concluding section will also include a consideration of the weaknesses and limitations of this study, and outline avenues for future research.

Marijuana Use and Effects

This paper focuses specifically on the United States' War on Drugs policies regarding marijuana, since marijuana is unquestionably the most widely used illicit substance in the world today[4,5–11]. Marijuana use is prevalent in American society, more so than any other illicit drug. Figure 1 shows U.S. Department of Health and Human Services data on the use of alcohol, cigarettes, marijuana, and non-marijuana illegal drugs by individuals within the year and month preceding the survey. As illustrated by the graph, nearly one in five U.S. adults has admittedly used marijuana within the past year. Furthermore, the use of marijuana in the past year exceeds the use within the year of all other illicit drugs combined. The use of marijuana in the past month (one of every ten adults) is also far greater than for

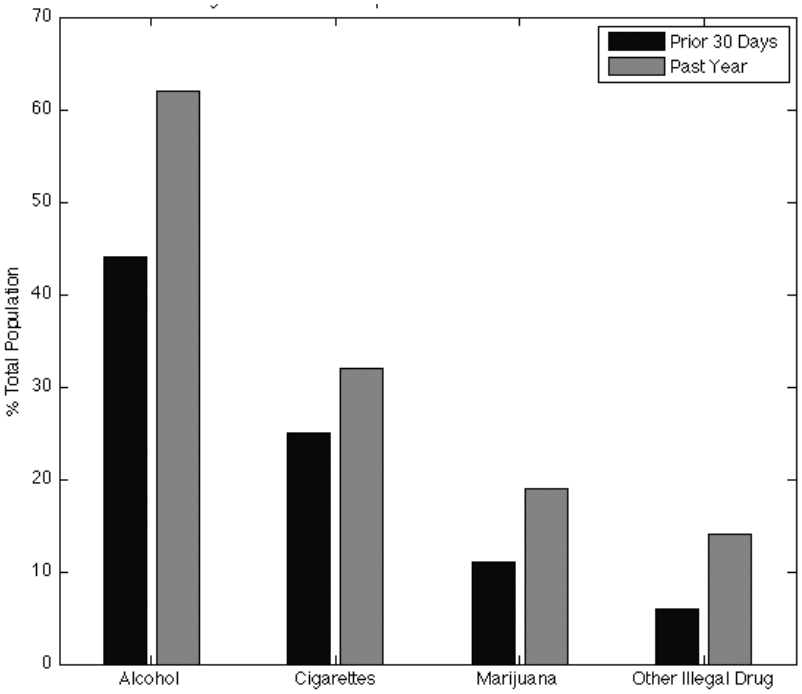


Figure 1. 2010 U.S. Population Substance Use[27]

all other illicit drugs. Usage within the past month is commonly regarded as "regular" use, and thus presents a particularly important variable for drug policy literature on marijuana.

Research suggests that first exposure to marijuana use for U.S. citizens is likely to occur during high school. A 30-year study conducted on nationally-representative samples of high school seniors, summarized in Figure 2, has led to estimates that three-quarters of current adults over age 18 in the U.S. have tried marijuana at least once during their lives[9]. Because illegal drugs are selectively marketed to youths[4], the fact that usage numbers in high schools are twice the population average is unsurprising. Even with elevated drug exposure, however, high school seniors currently display minimal past-year use of cocaine and heroin (approximately 5% and 1% respectively). Marijuana and other illicit drug usage among high school seniors has remained reasonably constant in the U.S. since the 1990s (see Figure 2), despite the introduction of legalized medicinal marijuana distribution and decriminalization in several states.

Because of marijuana's prevalence of use, as well as its seemingly equal use across ethnicities[9], marijuana is an optimal focus for a study on the effects of the current War on Drugs. Marijuana is also the most widely tested and historically analyzed illicit drug to date, and thus represents the greatest avenue for accurate analysis of the scientific, economic, and legal implications of the War on Drugs. Furthermore, marijuana is perceived by students at the 7th through 12th grade levels as being the most easily accessible drug, further heightening the importance of investigating its potential harms to U.S. citizens[10].

Drugs: The "Science" Behind the Schedule

The United States Drug Enforcement Administration (DEA), in an effort to formulate clear, identifiable sanctions against the use, possession, and sale of certain chemical compounds, produced a federal "Drug Schedule" of rankings from 1 to 5. Where a substance fell in this schedule was supposed to be a reflection of its threats to the well-being of the American citizenry. Schedule 1 chemicals allegedly serve no medicinal function and pose high risks of addiction and abuse. Schedule 2 drugs serve some medicinal purpose, but are also highly addictive. Drugs in Schedules 3-5 all serve medicinal purposes, but have decreasing levels of risk related to addiction and abuse. Alcohol, nicotine (tobacco), and caffeine were not placed in this schedule[11].

Scientific research on the dangers of each substance should inform its scheduling placement. However, these decisions were not made by scientific researchers or medical experts. Rather, these choices were left to the discretion of the United States Justice Department, Attorney General John Mitchell, and the Bureau of Narcotics and Dangerous Drugs during the development of the Controlled Substance Act of 1970[5].

Research on marijuana's chemical properties, as well as its medicinal and recreational uses has, throughout history, challenged the categorization of this substance by the federal government as a Schedule 1 drug. In fact, between 1840 and 1900 "more than 100 articles about the therapeutic value of cannabis were published in Europe and North America"[5] alone. With instances dating as far back as the

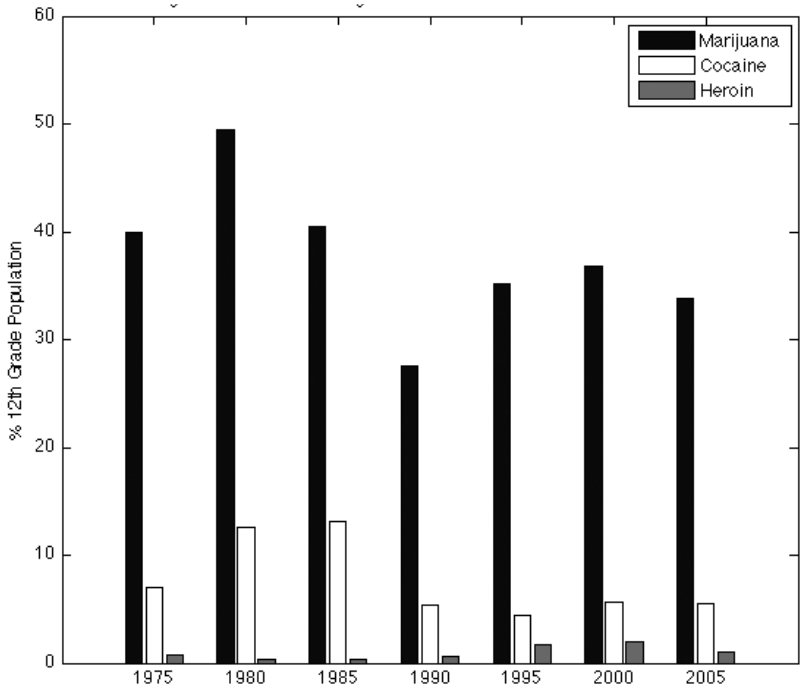


Figure 2. 12th Grade Drug Use in U.S. from 1975 to 2005[9]

first century A.D., marijuana has been used medicinally in China, India, the Middle-East, Central Asia, Greece, Rome, Africa, Europe, and America. Cannabis is also currently being used in various states across the U.S. to treat ailments such as nausea from chemotherapy, wasting-syndrome in AIDS victims, optical pressure from glaucoma, depression, anxiety, and insomnia, among many other afflictions[5,8,11]. Historic and current medical practices and reports support the argument that marijuana has legitimate medical purposes. These practices and reports directly rebut the assertion implied by the DEA's Drug Scheduling that marijuana lacks medical value.

The second criterion for categorization as a Schedule 1 drug is that risks of addiction and abuse are too high to merit legal distribution or prescription by doctors. Addiction refers to drug use that adversely affects the user's social standing, ability to perform civil duties, and capability to reduce or eliminate use despite a desire to do so[5]. An addicted individual is compelled to continue their high-usage of a substance despite negative consequences he or she faces from using the drug.

Addiction often occurs through tolerance or dependence, in which the chemical composition of bodily functions becomes so familiarized to the presence of a drug that stopping use results in adverse physical and psychological reactions. These reactions, called withdrawals, range in severity from mood swings to organ malfunctions. A current scientific study investigated the severity of addiction and withdrawal across six psychoactive drugs (caffeine, nicotine, alcohol, heroin, cocaine, and marijuana). This study concluded that marijuana and caffeine were the least addictive of these substances, asserting that marijuana is "slightly less addictive than caffeine."[8]

Other arguments for the placement of marijuana as a Schedule 1 drug highlight the negative consequences associated with marijuana use, such brain damage, biological defects, deviant behavior, and crime. Harry Anslinger, the first Commissioner of the Federal Bureau of Narcotics, issued a statement that marijuana "addicts" accounted for "fifty percent of the violent crimes committed…by Mexicans, Turks, Filipinos, Greeks, Spaniards, Latin-Americans and Negroes."[8] Noticeably absent from Anslinger's list of "violent criminals" are whites. Claims that marijuana causes violent behavior have been refuted in numerous studies conducted by independent researchers, national scientific organizations, and presidential commissions throughout the 20th and 21st centuries[5,8]. In fact, marijuana has been proven, in numerous controlled scientific experiments, to reduce aggressive behaviors, even among those addicted to harder drugs8. In addition, there is a wealth of reports on valid empirical studies refuting the claims of brain damage, biological defects, and criminal activity[5,8,11,12,13].

A further example of a fictitious detrimental effect of marijuana use is the "gateway theory," which suggests that the use of marijuana leads to experimentation with more dangerous illegal drugs. This argument is based upon government reports, such as the 1994 Center on Addiction and Substance Abuse report which stated that marijuana users were 85 times more likely than non-marijuana users to try cocaine[14]. These assertions were based on statistics of marijuana use by cocaine users and non-users, which indicated that 17% of marijuana users also tried cocaine, compared to the 0.2% of cocaine users who had never used marijuana[8]. This comparison does not accurately depict drug use motivations. Not only is marijuana the most widely used illegal drug, but it is also the least dangerous and most readily accessible, according to self-report data from U.S. youths10. The fact that 0.2% of cocaine users have never used marijuana simply reflects a logical usage progression, not a predisposition resulting from marijuana use. In fact, less than 1% of people who have used marijuana currently use cocaine[8].

Another important consideration for the enforcement policies regarding marijuana is the number of deaths caused by its use or abuse. This is one of the most unfounded pieces of evidence used to support current law enforcement policies, as deaths directly associated with the abuse and overdose of marijuana have not been documented in any historical record[5,6,8]. Estimates of death directly related to the abuse or overdose of tobacco, alcohol, and marijuana are displayed in Figure 3. The estimated annual death rates associated with the use of tobacco and alcohol in the United States are 430,000 and 80,000 respectively. While these substances are excluded from the Federal Drug Schedule, marijuana—among the highest scheduled substances—has zero instances of death annually due to abuse or overdose. In fact, laboratory research findings indicate that an individual would have to consume over 10,000 marijuana cigarettes within a couple of hours in order to reach lethal levels of THC[5,8]—a feat that can reasonably be deemed impossible.

It appears that marijuana, contrary to assertions by government officials and the DEA, does not fit either requirement for Schedule 1 categorization and does not cause detrimental health problems, induce criminal behavior, or cause death. Despite these glaring contradictions between the

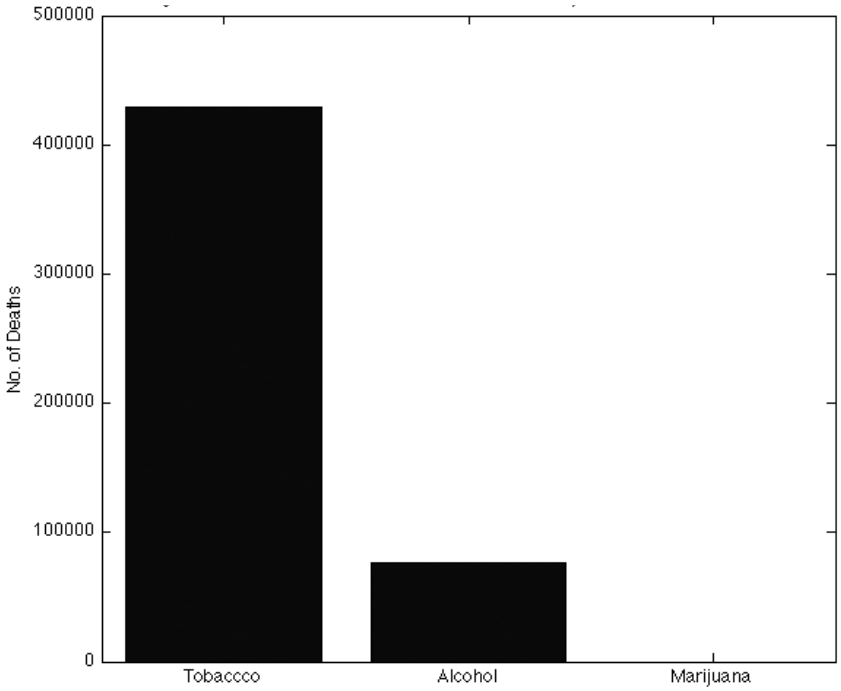


Figure 3. 2010 Annual U.S. Preventable Death Rate by Substance[27]

federal policy guidelines and scientific data, a panel of three law judges in the U.S. District Court of Appeals in Washington, D.C. decided in 1994, when presented with evidence for the rescheduling of marijuana, that "in their opinion [marijuana] has no medicinal value—none."[5] Since 1994, 16 states have legalized medicinal marijuana, yet marijuana continues to be inappropriately categorized as a Schedule 1 drug.

The Market: Implications of Marijuana's Removal from the Black Market

One argument for the continued prohibition of marijuana is that usage will increase if the substance can be legally purchased. Some officials claim that legalization of marijuana would increase use due to a reduction in price brought forth by the removal of risks involved in production and distribution[4]. These assertions have been challenged by researchers like DiNardo and Lemieux, who assert that decriminalization of marijuana throughout the U.S. has shown no effect on the prevalence of marijuana use for states that have enacted these policies[15]. Evidence against the price-sensitivity of marijuana use can be found in the Netherlands, where marijuana possession, sale, and use were recently legalized. Despite two decades of government-sanctioned marijuana sales, use among youths and young adults continues to be lower than in the United States[8].

Unlike the U.S., the Netherlands chose to decriminalize marijuana, and make it available to the general public through regulated dispensaries. Dutch society had not always accepted marijuana use, nor has it elected to legalize all drugs. This is a society that considered the evidence, implemented policy change, and has not experienced drastic negative consequences. In fact, the use of hard drugs among Dutch citizens is significantly lower than among American citizens, according to self-report data from each nation[16,17]. Furthermore, according to self-report surveys conducted in 1994, Dutch youths aged 12 to 19 were nearly six times less likely to use cocaine than American youths of the same age range[8] (see Figure 4). The vast disparity displayed in Figure 4 has resulted from the removal of marijuana from the same market as harder drugs such as cocaine and heroin, thus eliminating chance exposures to these harder drugs when an individual seeks to obtain marijuana[4].

Current costs for funding the War on Drugs are estimated between $26 and $58 billion annually, including enforcement efforts and incarcerations[4,6]. The funding requirements for continued enforcement use resources with little potential for revenue returns to the government or its citizens. Add to these costs the money required for drug education programs—estimated at around $1.3 billion in 2001[18]—and treatment facilities, and it becomes clear that America is dumping money into a system which has failed to display any significant deterrence or reduction in illegal drug use by either the country's youth or adult populations[9,10,19].

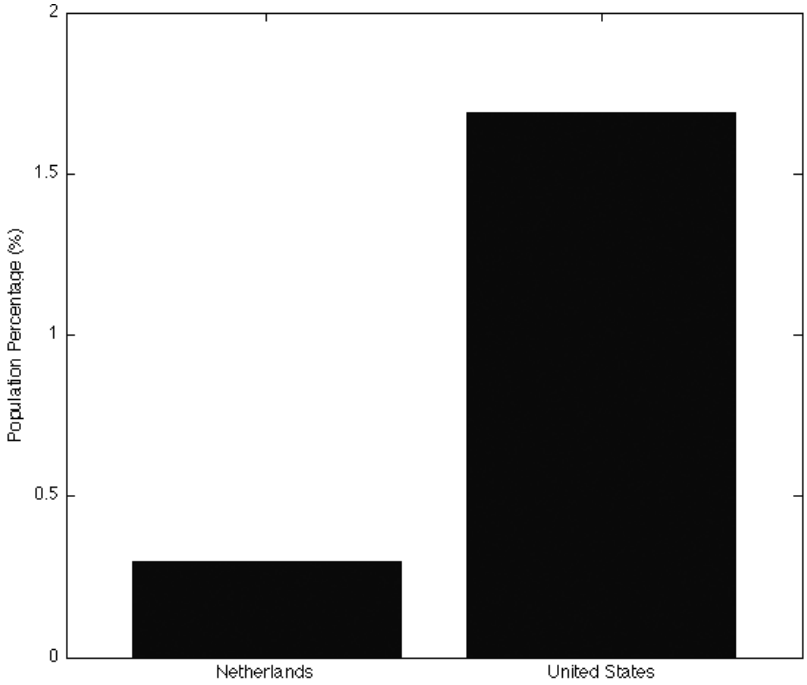


Fig. 4 1994 Cocaine Use for Youths Aged 12-19 in U.S. and Netherlands[16,17]

By contrast, government-sanctioned taxation and distribution of marijuana in the Netherlands has effectively provided the country with tax revenue to support drug education and treatment facilities, while also eliminating the costly processes of police training, enforcement, arrests, legal trials, and incarcerations[8]. Drug education programs in the Netherlands teach tolerance-based, responsible use of drugs, unlike the zero-tolerance, "Just-Say-No" approaches coined in the 1980s and still used today in the U.S. These foreign strategies to education, accompanied by the societal acceptance of responsible drug use, have undoubtedly contributed to the lower rates of use and abuse found in self-report data collected from Dutch youths[16].

Another issue presented by the War on Drugs is the increase of incarceration rates in relation to court-ordered drug treatment programs. Data from the Substance Abuse and Mental Health Services Administration, as well as work conducted by Delaney and colleagues, indicate that court-issues treatment facilities save tax payers 3-1 for in-patient rehabilitation, and upwards of 13-1 in out-patient rehabilitation compared to the expenses required for incarceration[18,20]. Further benefits of rehabilitation programs suggested by these studies include a reduction in the instance of positive urine-analysis drug screens and a nearly 20% reduced rate of re-arrest within six months after release from rehabilitation programs (compared to those released from prison facilities).

Demographics: Arrests, Convictions and Incarcerations

In the U.S., African Americans and Hispanics are perceived to constitute the majority of the drug-using population. Police officers across the country have openly admitted to "rational racial profiling" in stop-searches of minority civilians during routine traffic stops, claiming that disproportionate searches between races are the "unfortunate byproduct of sound police policy."[21] A likely result of the selective targeting of minority groups is the fact that African American men are five times more likely to enter prison than their white male counterparts[22]. The practice of rational profiling fuels these continued disparities. Citizens from a racial minority group are more likely to be arrested than their white peers, more likely to be convicted of drug offenses during criminal trial, and more likely to receive severe punishments from

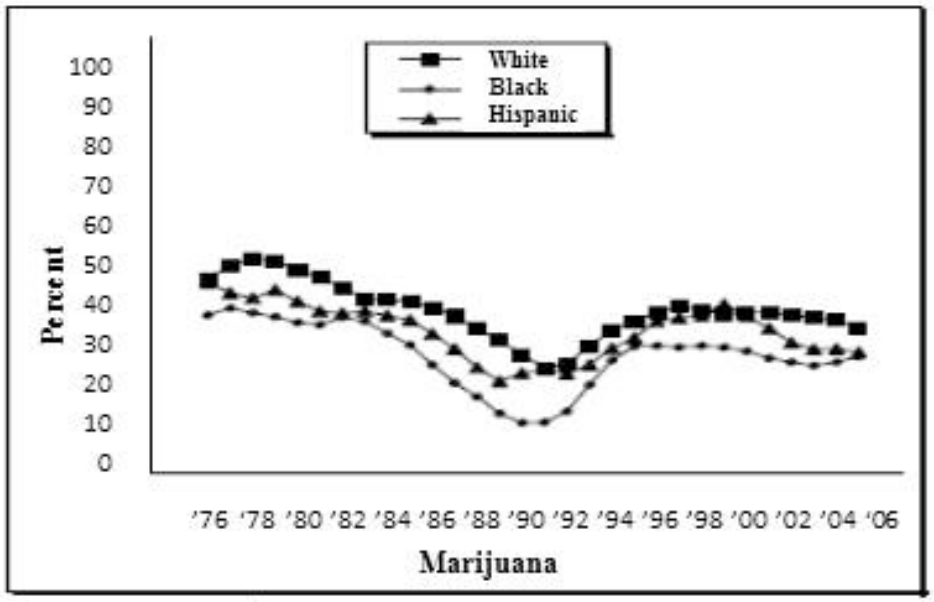


Fig. 5 7-12th Grade Population Marijuana use in the U.S. from 1975 to 2006[9]

court officials.

Statistics gathered from the U.S. Department of Health and Human Services in 2010 indicate that just over 18% of Americans admittedly used marijuana in the past year. Nearly 11% admitted to use in the 30 days preceding the survey[10]. But what is the ethnic composition of these users? In national surveys conducted on the employees of several corporations who decided to institute urine-analysis drug screens, lifetime marijuana use was significantly lower in almost every age category for both Hispanics and African Americans than in the white populace[23]. Wright and Rogers estimate that African Americans account for approximately 12% of the United States' "regular drug users," while also comprising approximately 13% of the entire U.S. population[25]. Survey data gathered from American students in 7th through 12th grade display lower usage of both licit and illicit drugs in the African American students, as well as lower rates of marijuana use in the Hispanic populations compared to white students for almost every year since 1975[9]. Figure 5 graphically displays marijuana use trends across these racial categories.

The available data suggest that claims of higher marijuana usage among minority populations are unsubstantiated and misleading. Instances of prejudiced police action, such as those discovered by Andrew Golub in his analysis of New York City's police records—where African Americans comprised 51% of traffic stops despite only composing 26% of the population—further display the bias plaguing police efforts[3]. Similarly, Maryland police records of stop searches show that 70% of those stopped are African American, while this ethnicity only accounts for 17% of drivers in Maryland[21]. These disparities provide clear evidence of highly discriminative law enforcement practices disadvantaging minorities.

Beyond the heightened occurrence of stops, searches, and arrests, African Americans and Hispanics also face further disadvantage in court proceedings. These minority groups experience elevated rates of pre-trial incarceration, hindering their ability to formulate a proper defense, impacting employment and family lives, and further enforcing the false image of their culpability[23]. African Americans are subjected to significantly higher rates of conviction for drug offenses, higher rates of incarceration, and elevated severity in punishment sanctions than their white drug-offender counterparts[24–26]. Hispanics face similarly heightened likelihood of incarceration and severity in punishment, especially when their offense occurs in an area with a large racial minority population[26]. When not facing incarceration, these groups are subjected to harsher monetary sanctions despite their generally lower socio-economic statuses, often resulting in the acquisition of their property upon failure to provide payments to the courts[25].

While it may be premature to accuse racism as the cause for these disparities, the evidence makes it difficult, if not dishonest, to rule it out completely. From selective pursuit of minorities by police to harsher criminal sanctions in the court room, it is clear that African Americans and Hispanics carry the greatest burdens of the societal and economic harm brought forth by the War on Drugs. The effects of such blatant targeting likely "exacerbate tension between racial minorities and law enforcement agencies," further fueling the instance of deviant behaviors and social backlash from these disenfranchised populations[22]. In essence, racial profiling may itself be providing the War on Drugs self-reinforcing behavior from minority and police entities, resulting in the statistics used in claims of enforcement effectiveness and the rationality of continued profiling strategies.

My research has three major implications. First, marijuana has no place among Schedule 1 drugs according to the DEA's own guidelines. Second, the enforcement costs of the War on Drugs coupled with the concurrent increase in incarcerations in relation to court-ordered rehabilitations are draining exorbitant amounts of U.S. tax revenue from the federal budget. Finally, police stops, arrests, convictions, and incarcerations of African Americans and Hispanics are disproportionately large in comparison to the portions of the drug-using population these ethnicities constitute.

These findings call for a reconsideration of drug enforcement policies regarding marijuana. Because of the scientific findings for marijuana regarding medical uses for nausea, wasting-syndrome, optical-pressure reduction, etc., as well as a lack of evidence to support claims of addictive properties, marijuana is a substance that should not be addressed with criminal sanctions, especially the severe sanctions for Schedule 1 drugs. Furthermore, in contrast to the annual death rates associated with abuse or overdose of alcohol and tobacco (Figure 3), there are no identifiable cases of death directly related to the abuse or overdose of marijuana. Yet alcohol and tobacco may be sold legally to citizens above a certain age, while marijuana may not[27]. Although an argument for complete legalization of marijuana is beyond the scope of this paper, this research opens an important conversation to be addressed with further investigation into the consequences of regulated, taxed sale of marijuana on crime, death rates, and deviant behavior.

In contrast to the zero-tolerance and minimal-sentencing laws of the War on Drugs, the monetary, criminal, and so-

cietal benefits of rehabilitation-focused approaches to drug treatment and punishment make these treatment-based court sanctions necessary in the United States' fight against drug abuse and addiction. The reduction in costs related to housing, processing, and feeding jail and prison inmates that results from placing these individuals in care facilities would provide large savings in federal tax revenues. Additionally, the significant reduction in re-arrest rates and drug usage relapses from successful treatment cannot be ignored. Evidence of these benefits can be drawn from the current Dutch model of drug treatment, which is funded almost entirely by the tax-revenue generated from legally sanctioned marijuana sales.

Lastly, the drastic disparity between arrests, convictions, and incarcerations of Hispanics and African Americans and the statistical use of drugs by these populations provides ample evidence of discriminatory practices in drug-related enforcement and judicial sanctioning. These disparities must be addressed with drug-policy reform, and educational material must be provided to law-enforcement officials and judiciary members highlighting these incongruences if this problem is to be resolved. The tensions created between minority populations and legal authorities because of unfounded "rational profiling" cannot be ignored.

References

1. US & World Population Clock. Available at: http://www.census.gov/population/www/popclockus.html. Accessed May 13, 2011.
2. Harrison PM, Beck AJ. Prison and jail inmates at midyear 2005. *Bureau Justice Stats Bulletin*. 2006:1–12.
3. Grossman M, Chaloupka FJ, Shim K. Illegal drug use and public policy. *Health Affairs*. 2002;21(2):134.
4. Golub A, Johnson BD, Dunlap E. The Race/Ethnicity Disparity in Misdemeanor Marijuana Arrests in New York City*. *Criminology & public policy*. 2007;6(1):131-164.
5. Mathre ML. *Cannabis in medical practice: a legal, historical, and pharmacological overview of the therapeutic use of marijuana*. McFarland & Co Inc Pub; 1997.
6. DuPont RL, Voth EA. Drug legalization, harm reduction, and drug policy. *Ann Intern Med*. 1995;123(6):461.
7. Blank DL, Command NM., Fenton JW. Early Employment Testing for Marijuana: Demographic and Employee Retention Patterns1. *Drugs In The Workplace*.139.
8. Zimmer LE, Morgan JP. *Marijuana myths, marijuana facts: a review of the scientific evidence*. Bookworld Services; 1997.
9. Johnston LD, Bachman JG, Schulenberg JE. Monitoring the Future: National Survey Results on Drug Use, 1975-2006. Volume I: Secondary School Students. NIH Publication No. 07-6205. *(NIDA)*. 2007:729.
10. U.S. Dept. Health and Human Serv. Subst Abuse Ment Health Serv. Admin. Office of Applied Studies. National Survey on Drug Use and Health, 2009. 2010. Available at: http://dx.doi.org/10.3886/ICPSR29621.
11. Richards DA. *Sex, Drugs, death and the law: an essay on human rights and overcriminalization*. Rowman & Littlefield Inc; 1982.
12. Model KE. The effect of marijuana decriminalization on hospital emergency room drug episodes: 1975-1978. *Journal Amer. Stat. Assoc*. 1993;88(423):737-747.
13..Bonnie RJ, Whitebread CH. The forbidden fruit and the tree of knowledge: an inquiry into the legal history of American marijuana prohibition. *Virginia Law Review*. 1970;56(6):971-1203.
14. Center on Addiction and Substance Abuse, Cigarettes, Alcohol, and Marijuana: *Gateways to Illicit Drugs*. New York; 1994:9.
15. DiNardo J, Lemieux T. *Alcohol, marijuana, and American youth: The unintended effects of government regulation*. Natnl. Bureau Econ. Research Cambridge, Mass., USA; 1992.
16. Sandwijk JP, Cohen PDA, Musterd S et al. Licit and illicit drug use in Amsterdam II. *Amsterdam: University of Amsterdam*. 1995.
17. U.S. Dept. Health and Human Serv. Substance Abuse Mental Health Serv. Admin. Office of Applied Studies. National Household Survey on Drug Abuse, 1994. 2008. Available at: http://dx.doi.org/10.3886/ICPSR06949.
18. Coffey RM, (US) C for MHS, (US) C for SAT. *National estimates of expenditures for mental health and substance abuse treatment, 1997*. U.S. Dept. Health and Human Serv, Substance Abuse Mental Health Serv. Admin; 2000.
19. Meier RF, Johnson WT. Deterrence as social control: The legal and extralegal production of conformity. *American Sociological Review*. 1977;42(2):292-304.
20. Delaney T, Gable A, Kendell N, Legislatures NC of S, (US) C for SAT. Treatment of alcoholism and drug addiction: What legislators need to know. In: National Conference of State Legislatures; 2000.
21. Harris DA. Racial profiling revisited: Just common sense in the fight against terror. *Crim. Just*. 2002;17:36.
22. Banks RR. Beyond Profiling: Race, Policing, and the Drug War. *Stanford Law Review*. 2003;56(3):571-604.
23. Gust SW, Walsh JM. Research on the Prevalence, Impact, and Treatment of Drug Abuse in the Workplace. *DRUGS IN THE WORKPLACE*.3.
24. Wright EO, Rogers J. *American Society: How It Really Works*. New York, NY: W.W. Norton & Company, Inc. 2011.
25. Harris A, Evans H, Beckett K. Courtesy Stigma and Monetary Sanctions. *American Sociological Review*. 2011;76(2):234.
26. Uhlman TM. Racial Justice. *Lexington, Mass.: Lexington*. 1979.
27. Center Disease Control and Prevention. Available at: http://www.cdc.gov/. Accessed May 19, 2011.

Matt Winterbourne, graduated in 2012, was born and raised in Cornelius, Oregon. He loves camping, fishing, lifting weights, and spending time with friends. Matt plans to become a wrestling coach and personal trainer in the future, along with a career in business or marketing.

Land and labor in the Negev: Palestinian Bedouin citizens of Israel

Rena Zuabi

This paper seeks to broadly examine the relationship between the State of Israel's transfer and resettlement policies targeted at Palestinian Bedouin of the Negev region and the subsequent position of this population in the labor market of Beer Sheva – the region's largest urban center. In 1948, the newly established state set out a clear policy agenda to consolidate state land ownership in peripheral regions of the country, to promote Jewish settlement in these regions, and ultimately further control densely populated Palestinian spaces. The Negev region was a marker of this policy agenda. This study centers on El Huashla vs. The State of Israel as the landmark 1984 Supreme Court case that legalized state claims over the majority of Bedouin land in the region and ultimately set a powerful precedent for future land disputes between this indigenous population and the state. The purpose of the research aims to ultimately illustrate the various means by which the state continues to utilize the law as a primary tool to reconfigure regional geographic and political landscapes and empower state policy goals. This area of study remains significant due to the way in which the struggle over land between the state and Palestinian Bedouin of the Negev has lead to the involuntary entrance of this population into the region's urban labor market and its continued inability to adequately access and integrate into it.

1. Introduction

The State of Israel has implemented expansive transfer and resettlement projects since the 1950s, significantly impacting the socioeconomic development of the Palestinian Bedouin population in the southern Negev region. A historical marker in this continuing struggle for land is the 1984 Supreme Court case El Huashla vs. The State of Israel, which solidified state land claims over the vast majority of land in the region at the expense of Bedouin claimants. The impact of these developments will be illustrated through the subsequent "proletarianization" of the Bedouin population, specifically in the region's urban capital Beer Sheva[1]. Following the formation of Israel in 1948, the newly established state sought to strengthen and consolidate Jewish presence throughout its territory particularly in targeted, peripheral and densely populated Palestinian areas such as the Galilee and the Negev. Land law has served as one of the state's most formidable tools to achieve this policy goal. Israeli land use planning ultimately served to strengthen Jewish presence throughout the newly established state. These policies remain controversial because they place Palestinians and Palestinian Bedouin communities in systemic cycles of underdevelopment and underrepresentation, which prompts several questions: how did the Supreme Court's empowerment of state land claims in the El Huashla case facilitate larger state policy goals in the Negev? how does the position of Bedouin citizens in this region's urban labor market reflect the impact of this ruling, and rulings like El Huashla?

The Supreme Court's ruling in the 1984 El Huashla case was largely based on a reinterpretation of existing land laws, notably the 1921 Mewat Land Ordinance. This ruling functioned to empower state land claims and facilitate the larger goal of territorial control and increased Jewish settlement in this densely populated Arab region. El Huashla was one case amongst many that facilitated the legalization of widespread state land expropriation during this time period. Furthermore, the subsequent landlessness of the affected Bedouin tribes coincides with the loss of traditional livelihood, leaving a large percentage of Bedouin to enter Beer Sheva's urban market as medial laborers.

1.1 Palestinian Bedouin & the Land Question

Bedouin Palestinians in the Negev region have become a recent focus in modern scholarship. Traditionally, this population has been studied for its culture, folklore, and post-nomadic experience with modernization. However, recent scholars in the field of political geography, legal geography, and historical geography have begun to study this population through its experience with state land planning policies.

The Bedouin faced a rapid reconfiguration of their traditional livelihood following the establishment of the State of Israel. Until 1948, the Palestinian Bedouin in the Negev relied on cattle, herding, rain-fed agriculture, and local commerce as their source of livelihood[2]. Although they fell under the jurisdiction of both the Ottoman government and the British Mandate, the Bedouin population traditionally followed its own system of land ownership and was generally not subject to land registration[2]. After the 1948 war and the establishment of Israel, however, only 19 of the original 95 tribes remain in the Negev region to be recognized by the state[3]. Currently, there are around 120,000-140,000 Palestinian Bedouin living in the Negev region. They constitute 15 - 20 percent of the region's population and own around 1.3 percent of the land, the majority of which is considered under illegal possession in the eyes of the state. Traditionally, the Palestinian Bedouin citizens of Israel have been subject to extreme poverty, high crime rates, and increased tensions with state authority, particularly due to the ongoing struggle over land in the region.

Lustik points out that among the many grievances that form the basis of Palestinian discontent with the status of their citizenship rights, expropriation of land is one of the most pressing alongside economic discrimination and problems facing educated Arab youth. He elaborates on this point of land extensively as a root cause of intra-state ethnic conflict between Arab and Jews, emphasizing "without a doubt the extensive and continuous expropriation of Arab lands has aroused the passionate antagonism of Israeli Arabs to a greater degree than any other single issue"[4].Quite a few years after Lustik published his book, Shafir and Peled also point out that "the struggle for land has been at the heart of the frontier conflict in Palestine/Israel, as elsewhere, and continues to be the most hotly disputed issue in the relations between Israel and the Palestinian, both citizens and non-citizens, to this day."[5]

1.2 Land Planning in the Negev: 'Making the Desert Bloom'

The Negev constitutes an internal frontier

in the state of Israel. Largely left untouched by state authorities well into the 1960s, this region has become the new subject of government development with land planning policies being the integral government instrument. Longstanding slogans of "making the desert bloom" packaged this endeavor in frontier romanticism[6]. As early as 1948, the Negev region was seen as the future of the State of Israel – an unchartered territory for Jewish settlement and revival [6,7].

Upon the establishment of a military government in Israel in 1948, 12 of the 19 Negev Bedouin tribes were involuntarily transferred to a designated territory in the northeastern part of the Negev called the Siyag[8,9]. Not only was this region already populated by seven Bedouin tribes, but is also infamous for its low agricultural fertility[2]. Until 1966, Palestinian Bedouin were also denied freedom of movement like their Palestinian counterparts in the northern Galilee. This forced enclosure of Bedouin society induced the community's rapid sedentarization and ultimately forced many tribes to give up their tradition of herding[6,7,9]. Unable to rely on their customary means of livelihood coupled with the Siyag's close proximity to the only urban center in the Negev, Beer Sheva, these Palestinian Bedouin quickly became another segment of the Arab proletariat in Israel as they sought employment from the Jewish sector in the city[9].

Bedouin Palestinians of the Negev were largely left in the Siyag by the planning authorities until the mid 1960s. In these twenty years no socioeconomic development took place for the newly transferred Bedouin communities[5]. Bedouin tribes were forbidden from establishing permanent settlements in the restricted area, and instead created makeshift tents and shacks to accommodate basic living needs[7]. The state considered the second resettlement project in the 1960s a compromise with the Bedouin because it provided Bedouin with limited state services and the opportunity to live in "modern" villages with electricity and running water, but the intention of land planning policies in the region remains clear[7]. As Ben-Gurion stated:

"Negev land is reserved for Jewish citizens, whenever and where ever they want… we must expel Arabs and take their places…and if we have to use force, then we have force at our disposal not in order to dispossess the Arabs of the Negev, and transfer them, but in order to guarantee our own right to settle in those places."[4]

1.3 The State's Legal Claims in the Negev

Through a number of legal tools, nearly the entire Negev region was declared state land by the Israeli government. The state's claim was based on an interpretation of the 1921 Mewat Land Ordinance. Mawat land (spelled mewat in British documents) was a legal category created in the 1858 Ottoman Land Code, and is defined as land left for the public use of nearby villages and could be recategorized as miri (land with private title) if it came into continuous cultivation[10]. This land was traditionally used as a source of water, timber, or other resources for public needs. Through the 1921 Mewat Land Ordinance, such public land was no longer available for free use by local villages. In fact, under the Ordinance, any person known to use these areas without the consent of the Director of Lands would face punishment as a trespasser[11].

Of the 2,560,000 dunams of mawat land expropriated in the region, Bedouin Palestinians previously owned and cultivated 1.2-2.2 million dunams until their forced resettlement[4,7]. The state could legally assert this claim based on the enforcement of the 1921 Mewat Land Ordinance registration stipulation, which mandated that any mawat land be registered with the Mandate administration by April 1921. Due to the fact that the Negev never underwent settlement of title under the British Mandate, this registration system never reached the Bedouin. Moreover, both the Ottoman administration and British Mandate sought to respect the general autonomy of the Bedouin land ownership system[3]. It was through the 1921 Mawat Land Ordinance that the state justified its claims to land in the Al-Huashla case, which will be examined in detail later on.

Although state attempts towards resettlement to townships emerged by the 1960s, these plans have largely failed through the active resistance of the Bedouin population. Upon the end of military rule in 1966, these displaced Arabs had brought forward 3,200 legal claims regarding expropriated land based on traditional Ottoman or British records. With their citizenship status, Palestinian Bedouin have been able to exercise campaigns for land rights and equality more broadly than under other regimes. Nevertheless, presently not one Bedouin claimant has received full land ownership rights in a land dispute case[7]. 95 percent of Arabs claims to land have not been settled in the court system, and these claims cover around eight hundred thousand dunams; Jewish citizens have already settled half of this land.

2. Al-Huashla vs. The State of Israel

Al-Huashla was the first significant Supreme Court land case to derive from the Negev and sets an important precedence for future Supreme Court cases involving Palestinian Bedouin land disputes. In 1984, the Supreme Court upheld a ruling brought to them in 1974, involving 13 Bedouin tribes who asked for court's recognition of their title to a delineated area of land. In El-Huashla vs. The State of Israel, these tribes asserted their title based on prescription - the historic use of this land for generations through herding and seasonal cultivation. The state's claim to land was based on the argument that the tribe did not in fact use the land in question – it was uncultivated and should be categorized as mawat, thus belonging to the state. The state further pointed to a registration stipulation in the 1921 Mawat Land Ordinance. Due to the fact that this law was never formally enforced in this peripheral region of Mandate Palestine, the Bedouin tribes had not fulfilled this registration requirement.

Without official legal title over the land, the only legal recourse for the appellants was to convince the court that the land in question did not qualify as mawat, but miri[3]. However, the state succeeded in rebutting this claim as well. Its claim was based in a previous Supreme Court ruling from the Galilee, which established the One-and-a-half Mile rule. This legal interpretation of Article 78 of the Ottoman Land Code deemed any land that lay One-and-a-half Miles outside of an inhabited village or settlement to be categorized as mawat. In the Al-Huashla case the nearest village to this land was 20 miles away and was a Jewish "development town" established after 1948. Moreover, the court dismissed the Bedouin's claim that an Arab village (Kurnov) had historically existed there before 1948 because it maintained only one permanent structure and a number of Bedouin tents. The judge further stated that the lack of agriculture in the area amounted the land to "a barren desert." [3]

The court's rulings empowered state land claims through a number of important legal mechanisms. First, the use of a very narrow definition of a village or

town dismissed the notion that a Bedouin encampment could be considered a legitimate form of settlement. The imposition of these definitions, although seemingly objective, ultimately devalues the Bedouin way of living – reducing it to a few tents and wild vegetation instead of a historical cultural lifestyle and community. The court's treatment of the Bedouin encampment paints it as socially invisible, another part of the wild vegetation and desert[3]. The Bedouin pastoral lifestyle also constituted an "abstract" form of possession in the state's legal framework, dismissing seasonal agriculture and herding as a valid use of land, and thus made it easier for the Court to justify expropriation.

Through the court's defense of state land claims, the Bedouin are left to confront the redefinition of time and space: the court began their legal history of land ownership at the moment of their involuntary displacement and when the state does permits them to leave their enclosure, they are treated as newcomers on the land that they had once inhabited for generations. The refusal of many Bedouin to lease state land and resettle for a second time in the state's "development towns" categorizes them as illegal squatters under Israeli law[3]. The court essentially retells Palestinian Bedouin history through the framework of the state's legal system in order to justify and defend the legitimacy of state land claims. Although seemingly neutral in language, the use of evidentiary preferences, legal categorizations, burden of proof on Bedouin appellants, as well as procedural rules ultimately function to empower the state in court at the expense of Palestinian claimants.

3. Impact: Bedouin in the Urban Labor Market, Beer Sheva

The Israeli land tenure system centrally transformed the sociological characteristics of Palestinian society in Israel. The decline of village life and the virtual stagnation of the Palestinian economy turned this minority population into a rural proletariat virtually overnight[1].

Generally, Palestinians in Israel make up 20 percent of the population yet constitute 11 percent of the workforce in Israel. Most of the employed Arabs occupy labor-intensive parts of the Israeli economy (unskilled industry, construction, and agriculture). The Arab population is almost completely absent from the "prestigious" parts of the Israeli economy that offer higher salaries (hi-tech, banking, electricity, water, etc). Moreover, the representation of Arabs in other "prestigious" branches of the economy such as academia, technical professions, and managers is significantly lower than their Jewish counterparts. These barriers to Arab employment reflect both discrimination from Jewish employers and also a shortage of jobs in Arab centers of the country.

The following general facts regarding the Beer Sheva District economy construct a framework through which Bedouin employment in the urban labor market will be examined. These facts are from the Main Findings of the Negev Center for Regional Development's most recent 2007 Report.

- In the year 2003 there were 159,200 people employed in the Beer-Sheva district.
- In the Beer-Sheva district there is a high percentage of employment in agriculture, food services and accommodation, transportation storage communication and education, in relation to the national percentage. The high percentage of employment in the food services and accommodation can be explained by the two large tourism centers in the area: The Dead Sea and the city, Eilat .
- There is a relatively low percent in the fields of wholesale commerce, banking, insurance, financing and business services.
- In the settlements that are the strongest economically and socially, Lehavim, Meitar and Omer (all Jewish settlements), and also in the weakest ones (the Bedouin settlements) there are few people who are in professional training.
- Half of the people in professional training in the Negev were residents of Sederot, a Jewish locality.
- Unemployment allowance receivers in the urban settlements within the area of the Negev Development Authority constitute 10.5% of the total unemployment allowance receivers in Israel in 2004, as opposed to 9.4% in 2003; 50% of them are men.

Ultimately, the economy of the Negev region is largely built off of tourism and industry. Although there is a lack of a professional market for both the Jewish and Bedouin Arab sector in the region, this study will highlight the way in which Palestinian Bedouin face disproportionate challenges in the region's labor market.

Historically, the Bedouin built their economy from land-intensive agrarian and herding activity. With this sort of economic activity no longer feasible due to state resettlement plans, urban types of employment were made necessary for this population. Although state reports recognize Bedouin state townships as "urban or suburban centers", they have no independent economic base; very few jobs can be provided locally[12]. Although even adjacent Jewish suburbs largely commute to city centers such as Beer Sheva, most Bedouin have no means to get to the country's main cities because there is no public transportation into Bedouin towns. Therefore, these impoverished communities must resort to their own means of transportation or systems of collective taxis[13]. Abu-Saad points out that in 2000, the Bedouin state township of Rahat had 5 cars per 100 residents, while the Jewish settlement of Omer had 41.1 cars per 100 residents[4]. For the Bedouin workers who do have the means to commute to city centers such as Beer Sheva, prospects for employment are dismally low.

Harvey Lithwick asserts that by 1995, five of the seven official Bedouin state townships in the Negev (Rahat, Arara, Tel Sheva, Kseifa, Segev Shalom) hold the lowest socioeconomic ranking in Israel. This can be compared to their neighboring Jewish settlements, such as Omer, Lehavim, and Metar, which hold some of the highest rankings in the country[13] (Table 1). illustrates the economic disparities between Arab and Jewish localities in the region[13]. Generally, average urban Bedouin's income is lower because of a lower rate of pay and also due to lower rates of employment throughout the course of one year. Moreover, since Bedouin women work an average of seven months in a year as opposed to the average Israeli woman who works nine months out of the year, gross family annual income is also seriously affected: the family salary income of Bedouins in Negev towns is ultimately less than half of the average Jewish family in Beer Sheva; family salary per person (per capita earned income) can be as low as 20 percent of the Israeli average[12]. Although there are Jewish towns whose income gaps are closer to that of Bedouin towns such as the Jewish town, Ofaqim, Lithwick asserts "the bottom line is that it is significantly harder for a breadwinner to support a family if one is a Bedouin." [16]

A recent survey of industries in

the Negev confirms that most of the larger, more successful firms, such as Motorola, Dell-Vishay, the Dead Sea conglomerate, and so forth, employ few, if any, Bedouin. According Lithwick's publication from Ben Gurion University in 2000, Bedouins held 2.5% of all the industrial jobs in the Negev in 1998. This publication further noted that a recent follow up inquiry confirmed no change in this percentage amongst the biggest firms[13]. In 2000, the Center for Research on Bedouin Society and Development wrote that in the entire Negev region, which holds a work sector of 15,000, fewer than 400 Bedouin are employed in manufacturing jobs. A few are employed in larger modern industry, and none are employed in high-tech industry[4]. The extremely low levels of employment in the main industrial companies in Beer Sheva illustrate this reality (Table 2).

The lack of employment in larger industrial companies have left Palestinian Bedouin men to search for menial jobs in construction, gardening, car repair, and service trades. These jobs offer little pay or security and provide lower opportunity for upward movement in the economy. With such little employment in the private sector, one might expect to find some sort of improvement in public sector employment. Due to the fact that this population constitutes around 15-20 percent of the region's population, one might expect there to be high employment in industries that serve the Bedouin sector. However, a 1998 survey shows that in Negev government offices, Muslims (almost all Bedouin) occupied 32 positions out of 25,000[13]. Out of the 32 Bedouin employees working for the government, 13 were in Beer Sheva where there is a much larger capacity for employment positions. This also means that there were only 17 Bedouin employees working for the service of Bedouin towns.

The question of land has held a potent impact on the Bedouin population in the Negev. State urbanization policies have not only left Bedouin Palestinian landless, but continuously subject to institutional discrimination and policies of neglect. The position of Bedouin Palestinian in the region's labor market, particularly in urban center such as Beer Sheva, illustrate the impact of state land policies on this minority population.

4. Conclusion

The struggle over land in Israel has centrally shaped the inability of Palestinian citizenry to develop and grow relative to Jewish citizenry in the state, rendering development both separate and unequal between the two populations. In a peripheral region such as the desert Negev, economic disparities are particularly acute. As state expropriation of land continues and Bedouin Palestinians increasingly resist state efforts to resettle and further transfer this population, the negative impact on this landless indigenous population continually poses a serious threat to intra-state ethnonational conflict on the national level.

Historically, the state has utilized a number of legal mechanisms to further its land policies in regions such as the Negev. The court has traditionally functioned to empower state land claims at the expense of Palestinian Bedouin claimants. For instance, using rules of evidence, prioritizing the burden of proof on the native landholder, and using legal categories to undermine Palestinian land rights have been means by which judges have been able to rule in favor of the state. Although the court provides a forum through which Bedouin tribes can indeed bring forth their evidence and dispute state claims, no Arab claimant has won a land dispute case in Israel's Supreme Court since 1948. The sheer scope of state land in the Negev region can be attributed to the early policies of transfer and resettlement in the 1950s and 1960s, which left Palestinian Bedouin in complete isolation for years. Unable to assert their title over land during the time of expropriation, Bedouin would not even be able to face the state in court until decades after their land was taken from them. By this time, large areas of this land were already settled and developed into Jewish localities. Bedouin who remain on their expropriated land have been subsequently labeled "illegal" – their communities' subject to state transfer policies or demolition at any time.

Although the impact of land tenure policies in the Negev can be analyzed in a number of different ways, the status of Bedouin employment in the region clearly illustrates the obstacles to development that this population continues to face. Staggering unemployment levels and general exclusion from private industries have left Palestinian Bedouin to work in the service sector and construction, which offer a lack of stability or opportunity for upward economic mobility.

A number of policy proposals have been put forth regarding the possible expansion of the urban labor market for Palestinian Bedouin. However, the call for an improvement of the legal status of this population in regards to land and labor has yet to be answered in the Israeli government.

References

1. Zureik E. Transformation of Class Structure Amongst the Arabs in Israel: From Peasantry to Proletariat. Journal of Palestine Studies 1976; 6.1:39-66.
2. Yiftachel O. Ethnocracy: Land and Identity Politics in Israel/Palestine. Philadelphia: University of Pennsylvania Press, 2006:193-198.
3. Shamir R. Suspended in Space: Bedouin under the Law of Israel. Law and Society Review 1996; 30.2:231-258.
4. Lustik I. Arabs in the Jewish States: Israel's Control of a National Minority. Austin: University of Texas Press, 1980:12-17.
5. Shafir G, Peled Y. Being Israeli: The Dynamics of Multiple Citizenship. Cambridge: University of Cambridge, 2002:112.
6. Abu Saad I. Spatial Transformation and Indigenous Resistance: The Urbanization of the Palestinian Bedouin in Southern Israel. American Behavioral Scientist 2008; 51.12:1713-1754.
7. Yiftachel O. Epilogue: Studying Naqab/Negev Bedouin – Towards a Colonial Paradigm? HAGAR Studies in Culture, Politics, and Identity 2008; 8.2:83-108.
8. Falah G. Israeli State Policy Towards Sedentarization in the Negev. Journal of Palestine Studies 1989; 18.2:71-91.
9. Meir A. Nomads and the state: the spatial dynamics of centrifugal and centripetal forces among Israeli Negev Bedouin. Political Geography Quarterly 1988; 7.3:251-270.
10. Article 103 of the 1858 Ottoman Land Code. Translated by F. Ongley. London: William Clowers and Sons, Limited, 1892.
11. 1921 Mewat Land Ordinance. Laws of the State of Israel; Volume 1.
12. Negev Statistical Yearbook. Negev Center for Regional Development. Beer Sheva: Ben Gurion University of the Negev, 2007.
13. Lithwick H. An Urban Development Strategy for the Negev's Bedouin Community. Ben Gurion University - the Center for Bedouin Studies and Development and the Konrad Adenaur Foundation 2000; 142:11-27.

	Average Monthly Family Salary for the Year	
	NIS	Index Beer Sheva = 100
Bedouin Towns		
Rahat	3989	45
Arara	3583	41
Tel Sheva	3860	44
Kseifa	3878	44
Segev Shalom	3787	43
Hura	3955	45
Jewish Towns		
Beer Sheva	8792	100
Ofaqim	5593	64
Dimona	7791	89
Arad	9237	105
Metar	16707	190
Lehavim	19048	217
Omer	21739	247

Table 1 Family Salaries for Bedouin and Jewish Towns, 1996[13]

Major Firms and Cities	*Industry Sector*	*Main Product(s)*	*Workers No.*	*Bedouins No.*
Beer Sheva				
Feuchtwanger Asher Ltd.	Electrical and Electronic	Controls	100	
Kirur Benny	Electrical and Electronic	Refrigerators	98	
Israel Aircraft Ltd – Ramta Division	Electrical and Electronic	Engines	480	
Harsa	Electrical and Electronic	Sanitary equipment	180	2
Fiber Technik	Rubber and Plastic	Chemical fiberglass equip	60	
A.A. Plast	Rubber and Plastic	Containers	60	5
Bel Art Israel Ltd.	Rubber and Plastic	Containers for chemicals	70	9
Triumph International	Textiles and Clothing	Undergarmets	100	
Nimrod	Textiles and Clothing	Shoes and sandals	80	4
Mishkan Hatchelet	Textiles and Clothing	Prayer shawls	40	2
Arihant Textiles	Textiles and Clothing	T-shirts	100	
Solog Knitting	Textiles and Clothing	Children's clothing	140	

Table 2 Total and Bedouin Employment in Negev Industry

Major Firms and Cities	*Industry Sector*	*Main Product(s)*	*Workers No.*	*Bedouins No.*
Dunhill S. General Industries	Chemical Products	Building material, marble	140	4
Hasin – Esh Ceramic Products	Chemical Products	Ceramics	51	6
Thermal Isolation Center	Chemical Products	Thermal insulation material	43	3
Makhteshim Chemical Words Ltd.	Chemical Products	Crop protection material	901	
Klir Marketing and Manufacturing Ltd.	Chemical Products	Chemicals	85	4
Zer Ma'adanim	Food and Beverage	Catering services	300	
Papa Michel Catering	Food and Beverage	Catering services	120	2
Harishonim Bakeries	Food and Beverage	Baked goods	50	17
Isa Khoury Metal Ind. Ltd.	Metals	Metal products	120	10
Zinor Darom (82) Industries Ltd.	Metals	Metal piping	80	10
Nisir Metal Ind.	Metals	Metal products	50	20
Ozgam Ltd.	Metals	Faucets	70	
M.T.D. Industries	Metals	Metal products	100	3
Razpal	Metals	Aluminum material	89	3
Mitrani	Metals	Bathtubs	47	
Gold & Honey	Accessories	Gold jewelery	114	
Oknin Brothers	Accessories	Stands	200	
Total			4891	160

Table 2 (continued) Total and Bedouin Employment in Negev Industry

Rena Zuabi is a Palestinian-Italian-American freelance writer and activist currently working between Nazareth and Ramallah, Israel/Palestine. She received a BA in political science from UC San Diego where she completed over two years of research and fieldwork on political-legal geography; namely the naturalization of ethno-religious spatial divides and inequalities in development across urban and rural landscapes inside the Green Line. Rena hopes to expand this research, through current and future work, towards questions of equitable food and water access and security in the broader Levant region.