The Penn State McNair Journal

Summer 2003, Volume 10

THE PENN STATE MCNAIR JOURNAL

Table of Contents

McNair Scholar Articles

WELCOME

Since 1991, the Penn State McNair Scholars Program has enriched the lives of students at both Penn State and Virginia State University, our partner in the McNair collaboration. The McNair Program holds a very special place in our lives as well as for the rest of the faculty and staff who work with our students. This publication celebrates their achievements and we offer it to our readers with pride and pleasure.

This is the tenth issue of the Penn State McNair Journal. We congratulate the Summer 2003 Penn State McNair Scholars and their Faculty Research Advisors! This journal presents the research conducted in the summer of 2003 by undergraduate students from Penn State and Virginia State Universities enrolled in the Penn State McNair Scholars Program. The articles within this journal represent many long hours of mutual satisfying work by the Scholars and their professors. The results of their research are published here and have also been presented at various research conferences around the country. We are especially proud to see how these students have grown as researchers and scholars. The hard work, dedication, and persistence required in producing new knowledge through research is most evident in these articles. We very much appreciate the guidance, expertise, caring and patience of our fine group of Penn State faculty research advisors.

We are also fortunate to have the support and encouragement of many faculty and staff who have worked with our students as social mentors or who have presented workshops and seminars on the many aspects of graduate and faculty life. You give the most precious of gifts to our students – your time in volunteering to support, encourage and nurture our Scholars' hopes and dreams.

Teresa Tassotti
Project Director

Curtis Price
Academic Coordinator

TRIO PROGRAMS ON THE NATIONAL LEVEL

Since their establishment in the mid-sixties as part of Lyndon Johnson's War on Poverty Program, TRIO Programs have attempted to provide educational opportunity and make dreams come true for those who have traditionally not been a part of the educational mainstream of American society. The TRIO programs are funded under Title IV of the Higher Education Act of 1965. While student financial aid programs help students overcome financial barriers to higher education, TRIO programs help students overcome class, social and cultural barriers to higher education. There are eight TRIO programs, which include the original three – Upward Bound, Talent Search and Student Support Services. The additional programs are Educational Opportunity Centers, Upward Bound Math & Science Centers, the Ronald E. McNair Post-Baccalaureate Achievement Program, a dissemination program and a training program for TRIO staff. McNair programs are located at 180 institutions across the United States and Puerto Rico. The McNair Program is designed to prepare participants for doctoral studies through involvement in research and other scholarly activities.

TRIO PROGRAMS AT PENN STATE

The TRIO Programs at Penn State comprise seven of the nine TRIO programs. There are two Educational Opportunity Centers in Philadelphia and Pittsburgh, Ronald E. McNair Program, Student Support Services Program, Talent Search, Upward Bound, Upward Bound Math & Science, and a TRIO Training Institute. These programs annually serve more than 4,000 students, from 6th graders through adults, with clear potential for academic success. The programs operate both at University Park and in communities across the state, often linking with middle schools, high schools, and community agencies. The programs focus on helping students overcome economic, social, and class barriers so that they can pursue education beyond high school.

MCNAIR SCHOLARS PROGRAM AT PENN STATE

Designed for low-income and first-generation college students, and students from groups underrepresented in graduate education, the McNair Scholars Program at Penn State encourages talented undergraduates from both Penn State and Virginal State Universities to pursue the doctoral degree. The program works closely with these participants through their undergraduate career, encourages their entrance into graduate programs, and tracks their progress to successful completion of advanced degrees.

The goal of the McNair Program is to increase graduate degree attainment of students from the above-mentioned underrepresented segments of society. McNair Scholars are presented with opportunities to study and do research in the University's state-of-the-art facilities to hone those skills required for success in doctoral education. Through both academic year and summer program components, McNair Scholars are required to complete a series of steps that lead to their application and enrollment in a graduate program of their choice.

Since 1991, the McNair Scholars Program at Penn State has helped 107 earn their baccalaureate degrees. Of these graduates, 81 or 76% have gone on to graduate school at institutions across the country, such as Harvard University, UC-Berkeley, Stanford, Cornell, UCLA, Boston University, Indiana University, University of North Carolina-Chapel Hill, University of Maryland-College Park, Penn State University, Emory University, University of Pennsylvania, American University, University of Iowa and Ohio State University.

ABOUT RONALD E. MCNAIR

Dr. Ronald Erwin McNair, the second African American to fly in space, was born on October 21, 1950, in Lake City, South Carolina. In 1971, he received a Bachelor of Science degree magna cum laude in Physics from North Carolina A&T State University. He continued his education at the Massachusetts Institute of Technology (MIT) where in 1976 he earned his Ph.D. in Physics.

While at MIT, McNair performed some of the earliest development of chemical and high-pressure CO lasers. He went on to study laser physics at E'cole D'ete Theorique de Physique in Les Houches, France. He was well published and nationally known for his work in the field of laser physics through the Hughes Laboratory.

In 1978, McNair realized his dream of becoming an astronaut when he was selected as for the first class of thirty-five applicants for the space shuttle program from a pool of several thousand applicants. Ronald McNair and six other astronauts died on January 28, 1986 when the space shuttle *Challenger* exploded after launching from the Kennedy Space Center in Florida.

McNair was the recipient of three honorary doctorates; a score of fellowships and commendations; an accomplished saxophonist; and held a sixth degree black degree belt in karate. He was married to the former Cheryl Moore and the father of two children, Reginald Ervin and Joy Cheray. After his death, Congress approved funding to honor the memory of McNair by establishing the Ronald E. McNair Post-Baccalaureate Achievement Program, which became the sixth program funded under the TRIO Programs umbrella.

Historians, who will write about McNair, the man, will discover that there was much more to him than his scholastics achievements. Friends who knew him, say he walked humbly and never boasted about his achievements. They say his commitments were to God, his family and to the youths he encouraged to succeed.
 (Ebony, May 1986)

EDITORIAL STAFF
Curtis Price
Teresa Tassotti

PROGRAM STAFF
Teresa Tassotti, Director
Curtis Price, Academic Coordinator
Judy Banker, Staff Assistant

RESEARCH METHODS COURSE INSTRUCTOR
Edgar Farmer

FACULTY RESEARCH ADVISORS AND EDITORS
Clemente K. Abrokwaa
J. Martin Bollinger
Squire J. Booker
Jeffrey Cohen
Henry Foley
Henry Giroux
Susan Searls-Giroux
Venkatraman Gopalan
Kyung-An Han
Patricia L. Johnson
Kenneth Lusht
M. Mercedes Maroto-Valer
Michael Micci
Murry Nelson
Stephen Piazza
David Shapiro
David B. Spencer
Elizabeth Walters
Keith B. Wilson

SOCIAL MENTORS
Clemente K. Abrokwaa
Vernon Carraway
Myrna Covington
James Davenport
Henry Foley
Cynthia Freeman-Fail
Ralph Gatrone
Joyce Hopson-King
Kenneth Lusht
Lee Newsom
Stephen Piazza
Curtis Price
Robert Pangborn
James Rambeau
David Spencer
Elizabeth Walters

Computer Simulation for the Optimal, Low-thrust LEO-TO-MOLNIYA Transfer Using NASA'S SEPSPOT Program

Felix Acon-Chen, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
David B. Spencer, PhD
Assistant Professor of Aerospace Engineering
College of Engineering
Pennsylvania State University

ABSTRACT

The nature of this work consists of running a computer simulation using NASA's SEPSPOT program to solve for the optimal low-thrust Earth-orbit trajectory for the LEO-to-Molniya transfer. For this scenario, a spacecraft is transferred from low Earth orbit to the final mission orbit by using various initial thrust accelerations ranging from 10^{-1} to 10^{-2} g. Furthermore, the numerical solutions obtained from the program for all the different cases are validated by comparing them with the analytic solutions derived from analytical blended control methods.

TABLE OF CONTENTS

Chapter 1

INTRODUCTION

1.1 Historical Perspective

During the past few decades, the aerospace industry has attempted to develop a solar-electric-propulsion planetary orbiter spacecraft, since such a vehicle would have a significantly increased propellant efficiency, greater maneuverability, larger payload capabilities, and a greater lifetime than a conventional chemical-propulsion space vehicle [1]. In the fall of 1998, NASA's New Millennium Program launched Deep Space 1, an ion propelled spacecraft, on an eleven month mission; however, the space vehicle exceeded NASA's expectations and kept running for a few more years, conducting testing on its ion engine [2].

Apart from successfully completing its primary mission, Deep Space 1 also flew by the comet Borrelly and transmitted the best close-up pictures and best scientific data ever gathered from a comet. Despite the fact Deep Space 1 was retired in December 2001, its spare ion engine has been running continuously at the Jet Propulsion Laboratory (JPL) in Pasadena, California since October 1998, which demonstrates that the ion engine is an excellent propulsion system for future space exploration missions. From a performance standpoint, the ion engine is capable of delivering ten times as much thrust per kilogram of fuel than conventional chemical engines [2].

Overall, Deep Space 1's spare ion engine ran for over 24,750 hours. To illustrate the effectiveness and efficiency of this engine imagine "if it had been an automobile engine instead of an ion engine, and it was driven for 24,750 hours at 80.5 kilometers per hour (50 mph), it would have traveled 1.93 million kilometers (1.2 million miles) without an oil change or tune up" [2].

1.2 The Problem

An important aspect of ion-propulsion technology is its high exhaust velocity, which allows an ion engine to run on a few hundred grams of propellant per day while allowing the ion-propelled spacecraft to travel faster and farther than any other space vehicle [2]. However, in order to achieve this high exhaust velocity, an ion engine must achieve very high specific impulse values, ranging in a few thousand seconds, along with relatively low-thrust levels in milli-Newtons [1].

As a result of the ion engine's limitations to low-thrust levels, the spacecraft's ion-propulsion systems will be required to operate for extended periods of time during orbit transfers in order to achieve the required mission orbit [1]. Moreover, the extended periods

required during orbit transfers, present the significant problem of computing an optimal trajectory while maintaining the spacecraft's solar panels pointed at the sun within tolerance levels; the solar-electric arrays provide power to the ion engine [3].

1.3 Significance of the Study

Many studies have been conducted on this issue in an attempt to find the best method for computing an optimal low-thrust Earth-orbit transfer. The nature of this work is to extend a study by Spencer and Herman [1] in which higher-order collocation methods and analytical blended control methods are applied to solve the optimal trajectory problem.

The contribution made to this study consists on running a computer simulation using NASA's SEPSPOT program on the LEO (low Earth orbit) to Molniya transfer, which was not included in the original project. In addition, the analytical data for the LEO-to-Molniya transfer computed by Spencer [4] is compared with the results obtained from the program.

The solution to this problem is imperative since it is preventing ion-propulsion technology, a very promising technology, from revolutionizing the face of space exploration by reducing orbit transfer through the ion engine's light weight, allowing spacecraft to travel at faster speeds with longer ranges [2], and designing spacecraft with larger payload capabilities and significantly greater lifetimes [1].

An example of how ion propulsion technology has the potential to enhance humanity's space capabilities is the combination of ion propulsion with solar arrays to create a space probe designed to escape from the solar system. Some individuals within the aerospace community predict that by combining these two technologies a spacecraft will be capable of traveling at an ultimate hyperbolic velocity of the order of 200 km/s. This means that such a space probe would be able to exceed by more than a factor of ten the ultimate velocity of Voyager 2, the space probe that did planetary flybys of Jupiter, Saturn, Uranus, and Neptune a few decades ago. Voyager relied on chemical propulsion [5].

In order to create such a spacecraft the solar arrays would have to be of "extremely low weight consisting of a photovoltaic thin film and conductors vapor-deposited on a thin mylar or kapton sheet." In addition, the sheet would have to be relatively small in width and very long in length. On the space probe, this sheet would be turned toward the Sun and stabilized by the ion engines, which would be arranged at the sheet edges [5].

Another possible configuration for the space probe would be for the sheet to be triangular, with the ion engines at the corners. In either case, the thrust vectors would have to be oriented such that a small fraction of the thrust can be used for stretching and stabilizing the sheet; moreover, the ion engines would also be used to unfold the folded sheet initially. However, all these predictions are still theoretical since the current solar arrays do not have the conversion efficiency required; it is predicted that in the near future solar arrays with the necessary conversion efficiency will be designed and manufactured [5].

1.4   Research Questions

This study will attempt to answer the following questions:

1. By doing a comparison of the results computed by SEPSPOT and the analytical data derived by Spencer during the original project, are both results closely related?
2. If the data is not closely related, what factors might have accounted for this?
3. What do the results look like when they are plotted?

Chapter 2

FUNDAMENTAL THEORY

2.1 Previous Work on Low-Thrust Optimization Methods

In past years, the trajectory optimization problem regarding the low-thrust propulsion systems has been investigated in order to find the best solution method. For example, multiple optimal and non-optimal transfer trajectories between specific initial and final orbits have been studied [6]. In addition, a method of averaging that provides a quick trajectory evaluation compared to methods based upon numerical integration of differential equations was developed [7].

Also, in another study, Lawden's "primer vector" theory was used to analyze impulsive and near-impulsive transfers in order to predict the conditions for low-thrust transfers. This study used algebraic approximations to compute the total time and gravity loss for relatively efficient transfers and to demonstrate that gravity losses for a transfer are reduced to a low level if enough burns are done [8].

2.2   Most Recent Work on Low-Thrust Optimization Methods

Herman and Conway [9] found optimal, low-thrust, Earth-moon orbit transfers by applying a method of collocation with nonlinear programming. The Earth orbit of the spacecraft and the final lunar orbit are both arbitrary while the moon is in its actual orbit. Furthermore, the total transfer time is minimized, but the trajectory is also propellant minimizing since the propulsion system operates continuously and prohibits a coast arc.

Also, Herman and Conway discovered that a very low initial thrust acceleration of 10^{-4} g yields flight times of approximately 32 days and requires many revolutions of both the Earth and the moon. In addition, if the problem is solved as two coupled two-body problems by ignoring the third body, then the optimal trajectory is changed slightly. The optimal trajectory is also insensitive to change in the engine specific impulse as long as the same initial thrust acceleration magnitude is used.

On the other hand, Prussing [10] examined minimum-fuel impulsive spacecraft trajectories in which long-duration coast arcs between thrust impulses are possible. If the coast time is long enough that it allows one or more complete revolutions of the central body then the solutions become complicated. This type of scenario presents Lambert's problem in which the determination of the orbit that connects two specified terminal points in a specified time interval brings about multiple solutions; a transfer time long enough to allow N revolutions of the central body has 2N + 1 trajectories that satisfy the boundary value problem.

Lambert's problem is a classical orbit boundary-value problem, which can be thought of as both an orbit determination problem and a spacecraft targeting problem. The solution to this problem in the two-body problem is the conic orbit that connects two specified terminal points in a specified time interval. In order to solve all the trajectories, Prussing developed an algorithm based on the classical Lagrange formulation for an elliptic orbit. Moreover, this procedure is applied to the problem of rendezvous with a target in the same circular orbit as the spacecraft, while the minimum-fuel optimality of the two-impulse trajectory is determined using primer vector theory [10].

Kechichian [11] also studied the minimum-time low-thrust rendezvous and transfer using the epoch mean longitude formulation. The study shows the state and adjoint differential equations as explicit functions of time that include natural orbital elements that stay constant if no perturbations are applied. In addition, the optimal Hamiltonian is time varying while the function that defines the transversality condition at the end time in minimum-time problems is illustrated as constant during the optimal transfer.

Coverstone-Carroll and Williams [12] developed a direct optimization method based on differential inclusion concepts and used the formulation to compute low thrust trajectories. This procedure removes explicit control dependence from the problem statement, which reduces the dimension of the parameter space and requires fewer nonlinear constraints in the resulting nonlinear programming problem. Moreover, the study presents simulations for a two-dimensional gravity-free trajectory, which involves a maximum velocity transfer to a rectilinear path, an Earth-Mars constant specific impulse transfer, an Earth-Jupiter constant specific impulse transfer, and an Earth-Venus-Mars variable specific impulse gravity assist.

In another study, Betts [13] used the direct transcription method, one of the most effective numerical techniques, to solve the trajectory optimization and optimal control problems. This method combines a sparse nonlinear programming algorithm with a discretization of the trajectory dynamics. Furthermore, the vehicle dynamics are defined by using a modified set of equinoctial coordinates while the trajectory modeling is described using these dynamics. Also, in order to demonstrate some special features of this method such as alternate coordinate systems during the transfer and mesh refinement to produce a high fidelity trajectory, the solution for the transfer from Earth to Mars including a swingby of the planet Venus is presented using the direct transcription method.

In addition, Kechichian [14] explored the optimal low-Earth-orbit-Geostationary-Earth-orbit intermediate acceleration orbit transfer by analyzing the problem of minimum-time orbit transfer using intermediate acceleration through precision integration and averaging. In his study, continuous constant accelerations of the order of $10^{-2}g$ are considered for applications using nuclear propulsion upper stages; in addition, the acceleration vector is optimized in direction with its magnitude held constant throughout the flight. The scenarios examined have trajectories that circle the Earth for only a few orbits before reaching geostationary Earth orbit, and these trajectories have demonstrated to be sensitive to departure and arrival points, requiring the use of the full six-state dynamics for satisfactory and meaningful results. Also, the ΔV losses with respect to very low-acceleration transfers are shown to be small.

2.3   Background on the Current Study

Spencer and Herman [1] focused on optimal low-thrust Earth-orbit transfers using higher-order collocation methods in which several Earth-orbit transfers, the LEO-to-GEO, LEO-to-MEO, and LEO-to-HEO transfers, were computed and then compared to the solutions found through analytical blended control methods. For each of these scenarios, a spacecraft is transferred from LEO to the final mission orbit by using various initial thrust accelerations (TA) ranging from 10^0 to 10^{-2} g. Refer to Figure 1 for the classification of these orbits.
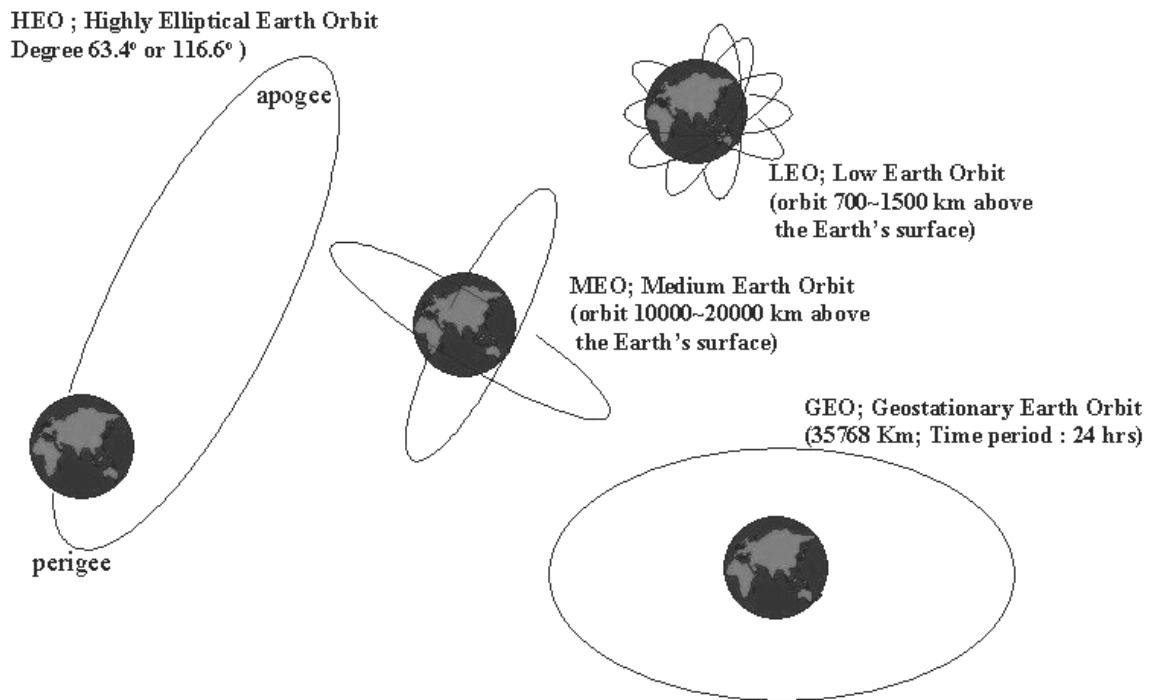
Figure 1: Classification of the GEO, LEO, MEO, and HEO Earth Orbits [15].

Their study involved determining the control time histories of a set of states, a system of first-order ordinary differential equations, from specified initial conditions to the desired final conditions while minimizing a function of the final values of states and/or time. These time histories are determined through a performance function, a scalar function consisting of the values of the states at the final time and the initial and final times, which is minimized while meeting the initial and final conditions of the system of differential equations [1].

2.3.1 Numerical Solution Method

For this method the problem is transformed into a mathematical programming problem (MP) by discretizing the time history solutions and then applying an approximate integration method. By discretizing the time history solutions into L subintervals, not necessarily of equal length, the endpoints of these subintervals are denoted as $\{t_0, t_1, …, t_{i-1},$ $t_i, t_{i+1}, …, t_{L-1}, t_L\}$. Within a given subinterval $[t_{i-1}, t_i]$ the time history of each solution is approximated by a numerical integration of the system of dynamics [1], which is given by Eq. (2.1) as

$$\dot{\bar{x}} = \bar{f}(\bar{x},\bar{u},t) \tag{2.1}$$

where the initial conditions for the states are

$$\dot{\bar{x}}\,(t_I) = \bar{x}_I \tag{2.2}$$

and the desired final conditions are represented by

$$\bar{\Psi}[\bar{x}(t_f)] = 0 \tag{2.3}$$

The original optimal control problem is formulated as a mathematical programming problem (MP) where the controls, $\bar{u}$, are determined to minimize the performance function, $\tilde{J}$, given in Eq. (2.2)

$$\tilde{J} = \phi\,[\,\bar{x}\,(t_F),t_I,t_F] \tag{2.4}$$

Furthermore, in this mathematical programming problem, higher order collocation 7th degree system constraints are applied to solving the system differential equations given in Eq. (2.1), which results in a non-linear programming problem (NLP) where the constraint Jacobian, exhibits a high degree of data sparseness. The software package SNOPT is used to solve the NLP-formulated problem due to its usefulness with sparse matrix problems. This entire method is called DHOC7 [1].

2.3.2 Analytical Solution Method

This method minimizes the propellant usage for a given transfer by assuming that the propellant usage rate is constant during a burn, which means, the burn times for a given maneuver are minimized by maximizing the time rate of change of the particular orbital parameter that governs the burn. During the first burn, thrusting is performed in the orbit plane to increase the apogee radius to the desired circular GEO (geostationary) radius value; in addition, the rate of change of the semimajor axis, da/dt, is maximized by determining the in-plane (α) motion of the thrust direction. Also, during the first burn, the out-of-plane component of the thrust vector does not exist. Afterwards, a coast is initiated and lasts until there is a second burn [1].

For the second burn, the first change from the initial value to the desired final value is the inclination of the spacecraft's orbit, followed by the orbit being circularized to correspond to the GEO. Next, the change in inclination, di/dt, is maximized in order to minimize the burn time. This results in the out-of-plane thrust angle (β) being near ± 90 degrees while the inclination change maneuver is centered about the apogee. Figure 2 illustrates the thrust vector and angle definitions [1].
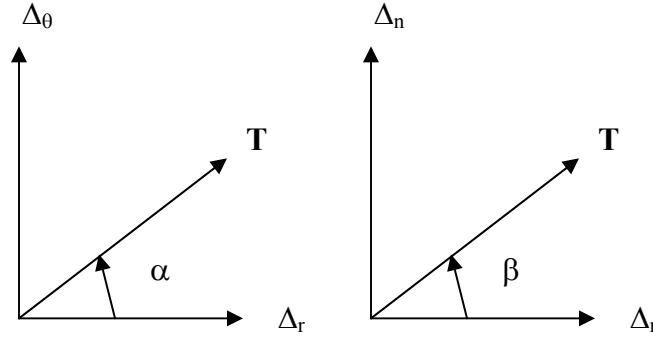
Figure 2: Thrust Vector and Angle Definitions [1]

For various cases of the LEO-to-GEO transfers with a range of TA values, a specific impulse of 1000 sec was used. When a comparison between the analytical solution (Spencer's method) and the numerical solution (DHOC7 method) was done for these cases, the results showed that Spencer's method provides a near-optimal performance by assuming his thrusting strategy. Table 1 denotes the change in total effective velocity between the two solutions [1].

Table 1 Comparison of LEO-to-GEO transfers methods [1]

Case	First Burn	Second Burn	Third Burn	Total	Change in Total Effective ΔV
Effective ΔV, m/s					
Spencer's results	3785	1237	782	5804	------------
3 Burns ─ Separate Controls	3751	1244	674	5668	2.40%
2 Burns ─ Combined Second Burn	3749	1499	------------	5248	10.60%
2 Burns ─ Optimized Thrust Directions	3944	1133	-----------	5077	14.30%

Here, the case named "3 Burns-Separate Controls" where the first and third burns are forced to raise the orbit, and the second burn rotates the orbit plane and zeros out the inclination, shows there is a percent error of 2.40% between the numerical and analytical solutions [1].

2.3.3 Equations Used

In order to avoid the singularities that occur in the modified classical orbit elements (a, e, i, Ω, ω, M) when e = 0 and i = 0 deg, modified equinoctial orbit elements must be used to describe the orbit transfers [1]. Therefore, the modified equinoctial orbit elements (p, f, g, h, k, L) must be defined in terms of the modified classical orbital elements as:

$$p = a(1 - e^2) \tag{2.5}$$

$$f = e\cos(\omega + \Omega) \tag{2.6}$$

$$g = e\sin(\omega + \Omega) \tag{2.7}$$

$$h = \tan(i/2)\cos \Omega \tag{2.8}$$

$$k = \tan(i/2)\sin \Omega \tag{2.9}$$

$$L = \Omega + \omega + \upsilon \tag{2.10}$$

In addition, the equations of motion of a thrusting spacecraft in an inverse square gravity field in terms of the modified equinoctial orbit elements are:

$$\dot{p} = (2p/\omega)(p/\mu)^{1/2}\Delta_\theta \tag{2.11}$$

$$\dot{f} = (p/\mu)^{1/2}\{\Delta_r \sin L + [(\omega+1)\cos L + f](\Delta_\theta/\omega) \\ - (h\sin L - k\cos L)(g\Delta_h/\omega)\} \tag{2.12}$$

$$\dot{g} = (p/\mu)^{1/2}\{-\Delta_r \cos L + [(\omega + 1)\sin L + g](\Delta_\theta/\omega) \\ + (h\sin L - k\cos L)(f\Delta_h/\omega)\} \tag{2.13}$$

$$\dot{h} = (p/\mu)^{1/2}(s^2\Delta_h/2\omega)\cos L \tag{2.14}$$

$$\dot{k} = (p/\mu)^{1/2}(s^2\Delta_h/2\omega)\sin L \tag{2.15}$$

$$\dot{L} = (\mu p)^{1/2}(\omega/p)^2 \\ + 1/\omega(p/\mu)^{1/2}(h\sin L - k\cos L)\Delta_h \tag{2.16}$$

$$\dot{m} = -T/c \tag{2.17}$$

$$\dot{\eta} = -(T/m_0)(1/c) \tag{2.18}$$

where

$$\omega = 1 + f\cos L + g\sin L, \quad s^2 = 1 + h^2 + k^2, \text{ and } \eta = m/m_0 \text{ and } \cdot = \frac{d}{dt}.$$

Also, the change in effective velocity is defined as

$$\Delta V_{eff} = -\left(\frac{T}{m_0}\right)\left(\frac{\ln[\eta(t_i)] - \ln[\eta(t_{i-1})]}{\eta(t_i) - \eta(t_{i-1})}\right)\Delta t_i \tag{2.19}$$

While the thrust vector $\overline{T}$ is computed by using two angles α and β, which represent the in-plane and out-plane components of the thrust direction,

$$\bar{T} = \mathrm{T} \begin{Bmatrix} \sin(\alpha)\cos(\beta) \\ \cos(\alpha)\cos(\beta) \\ \sin(\beta) \end{Bmatrix} = \begin{Bmatrix} \square_r \\ \square_\theta \\ \square_h \end{Bmatrix} \qquad (2.20)$$

2.3.4 Results

In this study, the scenarios where a spacecraft is transferred from the LEO-to-GEO, LEO-to-MEO, and LEO-to-HEO orbits while the final mass is maximized are analyzed. The conditions for this analysis are illustrated in Table 2 [1].

Table 2 LEO and GEO/MEO/HEO conditions for transfer trajectories [1]

Orbital Element	LEO	GEO	MEO	HEO
Semimajor axis, km	7003	42287	26560	26578
Eccentricity	0	0	0	0.73646
Inclination, deg	28.5	0	54.7	63.435
Right ascension of the ascending node, deg	0	0	0	0
Argument of perigee, deg	0	0	0	0
Mean anomaly, deg	Free	Free	Free	Free

These orbits were specifically chosen since the MEO orbit is representative of a global-positioning-system-type orbit while the HEO orbit portrays the Molniya orbit; in addition, different configurations for the spacecraft are assumed by varying thrust-accelerations (TA) ranging from 10^0 to 10^{-2} N/kg. The different thrust-accelerations considered are 1, 10, 10^{-1}, and 10^{-2} N/kg, which result in a total of 12 orbit transfer cases. Furthermore, in all of these cases, a burn-coast-burn thrusting structure is a priori determined for the transfer trajectories that duplicate the burn structure [1].

Thus, for each of the optimal orbit transfer cases a solution is found using each of the different thrust levels; moreover, the results for the maneuvers are analyzed and all cases are compared on the basis of their effective velocity change, ΔV_{eff}. The results for the LEO-to-GEO, LEO-to-MEO, and LEO-to-HEO transfers are illustrated in Tables 3-5 [1].

Table 3 LEO-to-GEO transfer results [1]

Initial Thrust Acceleration, N/kg	Effective ΔV, m/s			Total Transfer Time, hours
	First Burn	Second Burn	Total	
10^1	2366	1761	4127	5.40
10^0	2592	1716	4308	6.06
10^{-1}	4079	1088	5167	18.32
10^{-2}	5698	------------	5698	149.59

Table 4 LEO-to-MEO transfer results [1]

Initial Thrust Acceleration, N/kg	Effective ΔV, m/s			Total Transfer Time, hours
	First Burn	Second Burn	Total	
10^1	2008	1856	3863	3.06
10^0	2137	1834	3970	3.60
10^{-1}	3717	1014	4731	14.56
10^{-2}	5122	------------	5122	135.23

Table 5 LEO-to-HEO transfer results [1]

Initial Thrust Acceleration, N/kg	Effective ΔV, m/s			Total Transfer Time, hours
	First Burn	Second Burn	Total	
10^1	2434	836	3271	6.03
10^0	2666	890	3555	6.55
10^{-1}	4146	1125	5271	18.59
10^{-2}	6109	------------	6109	159.75

In all of these cases, the highest thrust acceleration, 10 N/kg (approximately 1g), results in a transfer where the burn duration is small compared to the coast arc which indicates that the performance of the transfer is near the performance for a transfer that uses a high-thrust impulsive approximation; a burn duration that is small compared to the coast arc is a characteristic of a transfer based on a high-thrust impulsive approximation [1].

A noticeable aspect of the results for the three transfer types is that as the thrust-acceleration decreases, the ΔV_{eff} and transfer time increases. For instance, a case that

results from a thrust-acceleration (TA) of 1 N/kg has a slightly higher ΔV_{eff} and transfer time than a case resulted from a TA of 10 N/kg. Another observation from the results is that for the scenario that uses a TA of 10^{-1} N/kg, the ΔV_{eff} required is 20-25% more for the LEO-to-GEO and LEO-to-MEO cases [1].

For all three-transfer types, the three-dimensional trajectories are illustrated in Figures 3-5 as near optimal, low-thrust transfers taking on the shape of spirals with increasing radius. Here, the thick lines represent a burn arc, and the thin line denotes coast arcs [1].
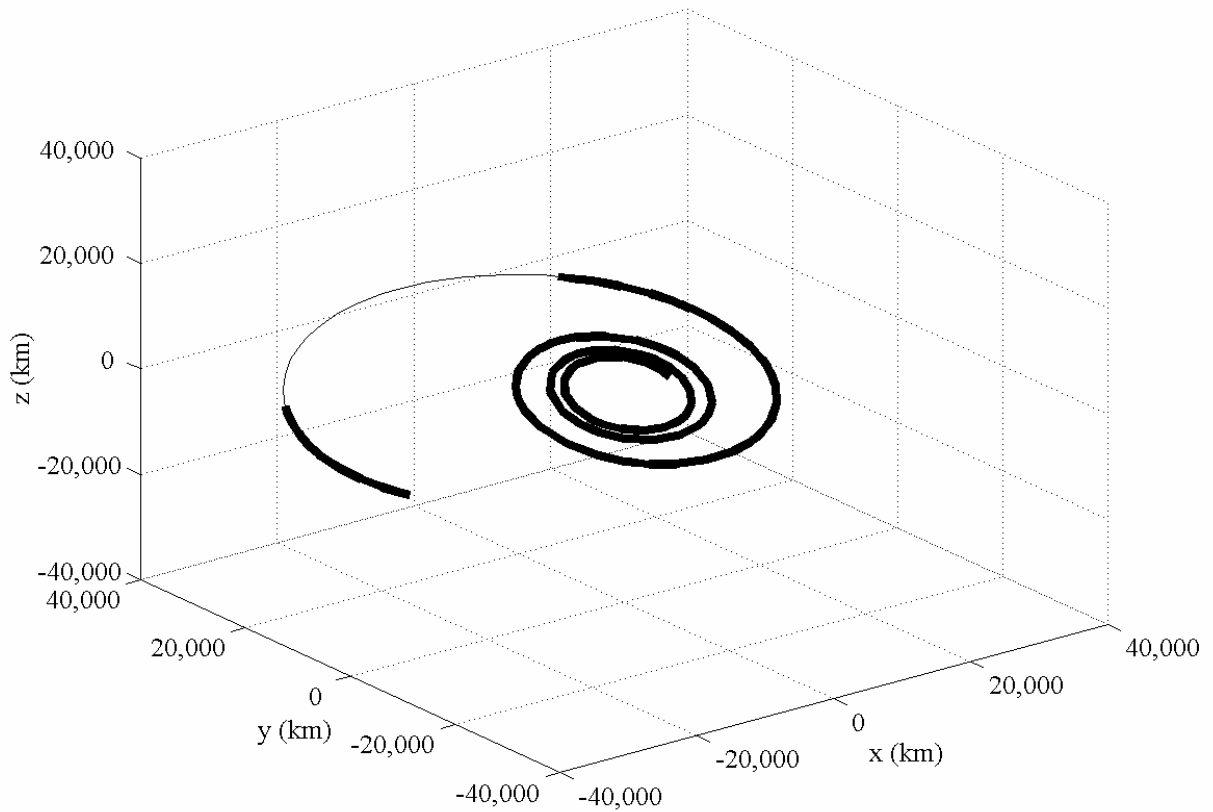

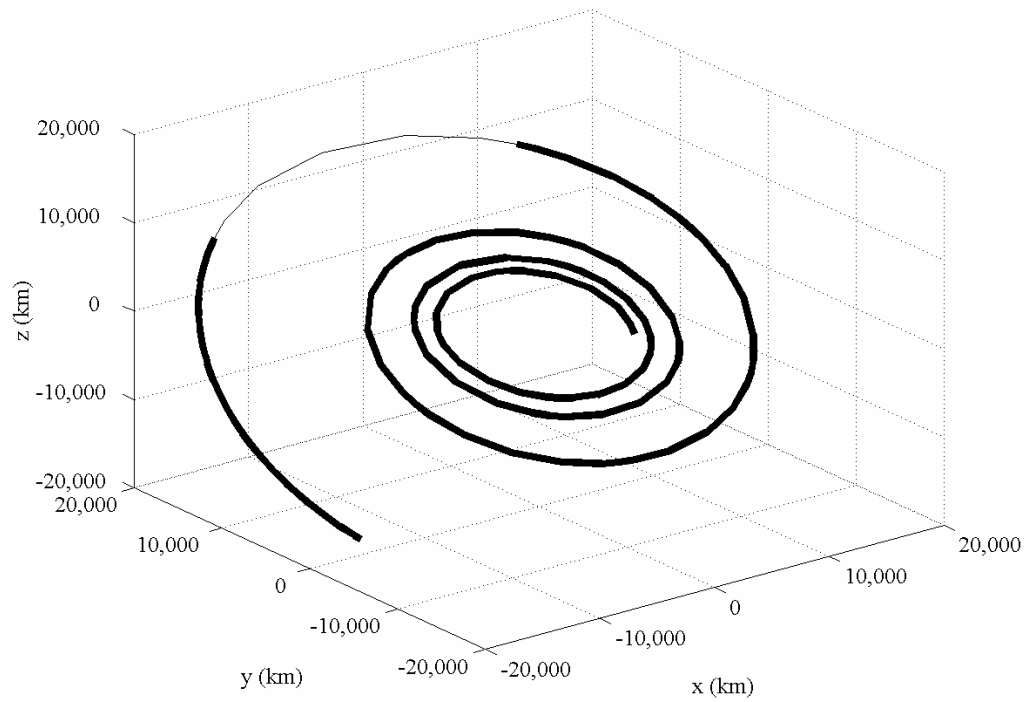**Figure 3: LEO-to-GEO transfer with initial thrust acceleration of 10^{-1} N/kg [1]**

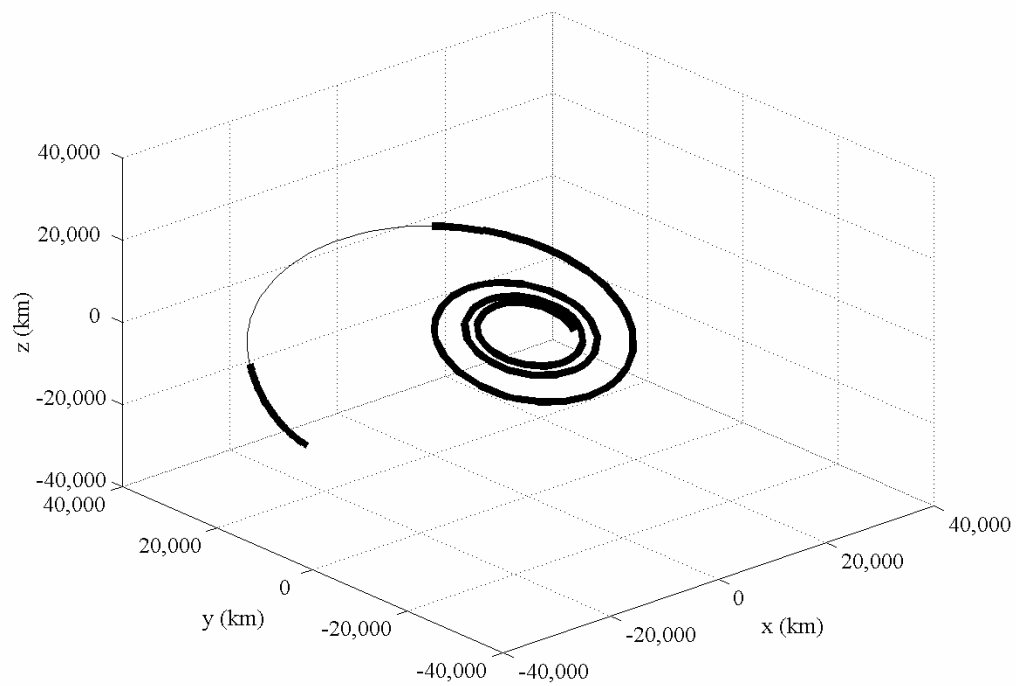**Figure 4: LEO-to-MEO transfer with initial thrust acceleration of 10^{-1} N/kg [1]**

Figure 5: LEO-to-HEO transfer with initial thrust acceleration of 10^{-1} N/kg [1]

The ΔV_{eff} for the LEO-to-GEO and LEO-to-MEO transfer type are approximately 35% greater than the corresponding transfer types with a TA of 10 N/kg, and the LEO-to-HEO transfer type is approximately 85% more than the corresponding transfer with a TA of 10 N/kg [1].

2.4 SEPSPOT

SEPSPOT is a modified version of the SECKSPOT (Solar Electric Control Knob Setting Program by Optimal Trajectories) computer program. The program is written in Fortran IV with double precision. A costate formulation is used which results in a two point boundary value problem which is solved using a Newton iteration on the initial unknown parameters and the unknown transfer time. Also, a Runge-Kutta method is used to integrate the state and costate equations and averaging is done using a Gaussian quadrature [16, 17].

SEPSPOT is designed to calculate time optimal or nearly time optimal geocentric transfers for a solar electric spacecraft with or without attitude constraints. The program has the option to use initial high thrust or low thrust. For the initial high thrust stage one or two impulses of fixed total ΔV can be included, and the initial orbit is assumed to be circular. For the low thrust stage, a nonsingular set of orbital elements and an averaging method are used. In addition, the low thrust phase is applicable to general geocentric elliptical orbits [16,17].

The program also includes options for oblateness, solar motion, shadowing with or without delay in thruster startup, and an analytic radiation and power degradation model. The main modifications done to the original SECKSPOT program include the altitude constraint solution, a new radiation and power loss model, a revised shadow model, and extended output. Also, one key aspect of using the altitude constraints option is that it causes power to become a function of thrust direction and sun direction, and the time optimal thrust direction becomes a complex function of primer vector direction [16,17].

SEPSPOT's input is entered in the form of a data file (input file), which contains initial values of unspecified states and costates and a guess for the transfer time. The initial and desired orbit are specified in terms of semimajor axis (km), eccentricity, angle of inclination (degrees), longitude of ascending node (degrees), and argument of perigee (degrees). The initial thrust acceleration must be included in terms of initial mass (kg), initial power (kw), specific impulse (sec), and total constant efficiency (ε) [16, 17].

2.4.1 SEPSPOT's Equations

SEPSPOT operates on a set of equinoctial elements. These equations [17] are shown in Eqs. 2.21-2.30:

a) **Equinoctial Orbital Elements (in terms of classical elements):**

$$a = a \qquad\qquad (2.21)$$

$$h = e \sin(\omega + \Omega) \qquad\qquad (2.22)$$

$$k = e \cos(\omega + \Omega) \qquad\qquad (2.23)$$

$$p = \tan(i/2) \sin \Omega \qquad\qquad (2.24)$$

$$q = \tan(i/2) \cos\Omega \qquad\qquad (2.25)$$

b) **Inverse Relationships:**

$$a = a \qquad\qquad (2.26)$$

$$e = \sqrt{h^2 + k^2} \qquad\qquad (2.27)$$

$$i = 2\tan^{-1}\sqrt{p^2 + q^2} \qquad\qquad (2.28)$$

$$\Omega = \tan^{-1}(p/q) \qquad\qquad (2.29)$$

$$\omega = \tan^{-1}(h/k) - \tan^{-1}(p/q) \qquad\qquad (2.30)$$

The costate equations [17], used for the optimization, are shown in Eqs. 2.31-2.40:

a) **In terms of Equinoctial Orbital Elements (λ and ψ = adjoints):**

$$\lambda_a = \psi_a \qquad\qquad (2.31)$$

$$\lambda_h = \psi_e \sin(\omega + \Omega) + \psi_\omega \cos(\omega + \Omega)/e \qquad\qquad (2.32)$$

$$\lambda_k = \psi_e \cos(\omega + \Omega) - \psi_\omega \sin(\omega + \Omega)/e \qquad\qquad (2.33)$$

$$\lambda_p = \psi_i\, 2\sin\Omega\, \cos2(i/2) + \psi_\Omega\, \cos\Omega/\tan(i/2)$$
$$- \psi_\omega\, \cos\Omega/\tan(i/2) \tag{2.34}$$

$$\lambda_q = \psi_i\, 2\cos\Omega\, \cos2(i/2) - \psi_\Omega\, \sin\Omega/\tan(i/2)$$
$$+ \psi_\omega\, \sin\Omega/\tan(i/2) \tag{2.35}$$

b) In terms of Classical Orbital Elements:

$$\lambda_a = 0 \tag{2.36}$$

$$\lambda_h = \psi_e\, \frac{h}{\sqrt{h^2 + k^2}} + \psi_\omega\, \frac{k}{h^2 + k^2} \tag{2.37}$$

$$\lambda_k = \psi_e\, \frac{k}{\sqrt{h^2 + k^2}} - \psi_\omega\, \frac{h}{h^2 + k^2} \tag{2.38}$$

$$\lambda_p = \psi_i\, \frac{2p}{\sqrt{p^2 + q^2}}(1 + p^2 + q^2)$$
$$+ \psi_\Omega\, \frac{q}{p^2 + q^2} - \psi_\omega\, \frac{q}{p^2 + q^2} \tag{2.39}$$

$$\lambda_q = \psi_i\, \frac{2q}{\sqrt{p^2 + q^2}}(1 + p^2 + q^2)$$
$$- \psi_\Omega\, \frac{p}{p^2 + q^2} + \psi_\omega\, \frac{p}{p^2 + q^2} \tag{2.40}$$

Chapter 3

ANALYSIS

3.1 Preliminary Work

A zip file containing the SEPSPOT program files and the program's manual and analysis in the form of a pdf file were obtained from NASA/Glenn Research Center. The zip file was uncompressed and Jeff Nucciarone, the Pennsylvania State University senior research programmer for the High Performance Computing Group, ITS/ASET, compiled the program with the Linux operating system. After the program was compiled a cluster was opened, LIONXL, in Pennsylvania State University's High Performance Computing Center and all the program files were copied into a folder named "sepspot."

To run and execute SEPSPOT a program called "SSH Secure Shell" was used to connect via the Internet to the LIONXL cluster, which uses the Linux system. The SSH Secure Shell program was installed and ran in both Intel Celeron (533 MHz with 319 RAM) and Intel Pentium III (930 MHz with 256 RAM) computers with Windows ME and Windows XP operating systems.

3.2 Variables: Independent and Dependent

Tables 6 and 7 show the initial conditions required to run the computer simulation. The initial conditions used to declare the initial and final orbit consist of the semimajor axis (km), eccentricity, inclination (degrees), right ascension of the ascending node (degrees), and argument of perigee (degrees) for both the LEO and Molniya orbits. In addition, the initial mass (kg), initial power (kw), thruster specific impulse (sec), and the total transfer time (hours) are considered in order to compute the initial thrust acceleration using the following relation [16,17,18]:

$$P = \frac{g_0 T I_s}{2\varepsilon} \tag{3.1}$$

Also, the final conditions are the total effective change in velocity, the total transfer time (hours), the semimajor axis time history, the eccentricity time history, the inclination time history, the apogee and perigee radius time history, and the energy time history.

Table 6 Initial Conditions for the Initial and Final Orbit

Element	Initial Value (LEO)	Final Value (Molniya)
a	7000 km	26578 km
e	0	0.73646
i	28.5°	63.435°
Ω	0°	0°
ω	0°	0°

Table 7 Initial Conditions for Initial Thrust Acceleration

Element	TA = 10^{-1} N/kg	TA = 10^{-2} N/kg
Initial Mass (kg)	1	1
Initial Power (kw)	0.4905	0.04905
ISP (sec)	1000	1000
Estimated Time of Arrival (hours)	13.5	139.2

For this study the initial conditions are regarded as the independent variables since all other data is derived from these conditions. The final conditions are considered to be the dependent variables since they are derived from the initial conditions or the independent variables. Figure 6 depicts the orientation of the right ascension of the ascending node (Ω), the argument of perigee (ω), and the inclination (i).
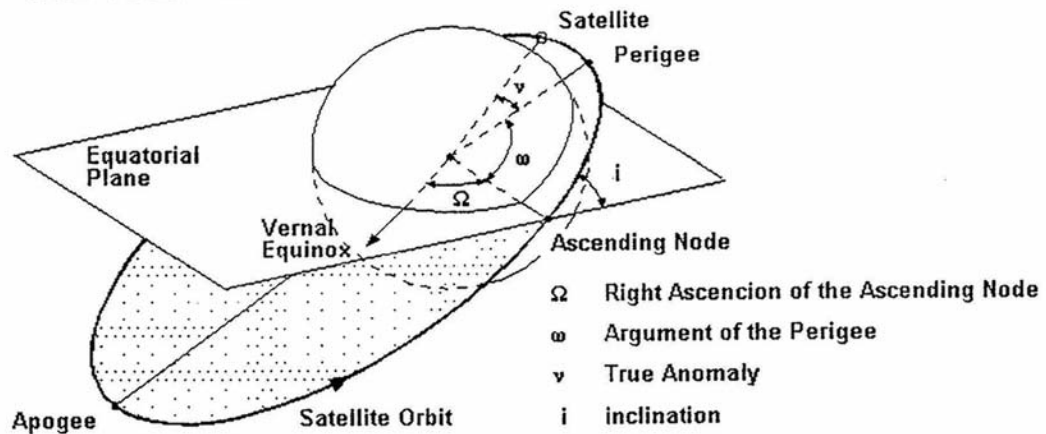
Figure 6: Orbit elements

3.3   Instrumentation

A quantitative numerical research method is used in this study. The main instrument used during this project is the computer program, SEPSPOT, since it collects the data by computing the final conditions from the initial conditions. This instrument can be considered reliable since it will always be consistent with its solutions; the program is based on mathematical equations.

Furthermore, another instrument that is used in the study is the analytical data that was produced during the original project since it serves as secondary data. The analytical solutions obtained from Spencer's research study serves as a tool to validate the data produced by SEPSPOT.

3.4   Data Collection and Analysis

For this project, the final conditions are derived from the solutions produced by SEPSPOT. In addition, the final data is analyzed by using Microsoft Excel to plot and compare the results with Spencer's analytical solutions. Overall, there are a total of five plots for each initial thrust acceleration (10^{-1} N/kg and 10^{-2} N/kg). For these graphs, the semimajor axis (km), eccentricity, inclination (degrees), apogee and perigee radius (km), and energy are all plotted versus time (hours).

Chapter 4

RESULTS

4.1 LEO-Molniya Transfer: $T/m_0 = 10^{-1}$ N/kg

For the case of an initial thrust acceleration of 10^{-1} N/kg, SEPSPOT accomplishes the transfer in one burn while achieving an overall effective change in velocity (ΔV) of 5814.69 m/s in 12.18 hours. On the other hand, Spencer's analytical solution completes the transfer using two burns in 13.75 hours at a ΔV of 6896.00 m/s. Refer to Table 8 to view a comparison of the two solutions, which shows that SEPSPOT's trajectory is slightly more efficient by completing the transfer in less time; in addition, there is a percent error of 12.89% between the numerical solution provided by SEPSPOT and Spencer's analytical solution.

Table 8 Initial Thrust Acceleration of 10^{-1} N/kg

	Overall Effective Change in Velocity (ΔV)	Overall Time	Percent Error (%)
Spencer's Results	6896.00 m/s	13.75 hrs	**12.89%**
SEPSPOT's Results	5814.69 m/s	12.18 hrs	

Five figures are now presented for this case. Figures 7-11 show a comparison of the time history of the semimajor axis, eccentricity, inclination, apogee and perigee radius, and energy between SEPSPOT's numerical data and Spencer's analytical data. A key aspect of the comparison that should be noticed is that SEPSPOT's trajectory manages to complete all the desired conditions in approximately the amount of time it takes Spencer's trajectory to complete the first burn.

Figure 7 shows how SEPSPOT manages to achieve a semimajor axis of 26578 km (Molniya Orbit) from a starting semimajor axis of 7000 km (LEO Orbit) in 12.18 hours by using one burn. However, Spencer's trajectory shows that a burn is performed for approximately 12 hours, followed by a coast arc of 4.5 hours, and then a second burn is made which takes about 1.5 hours to complete the trajectory. One should also notice that SEPSPOT's trajectory seems almost parabolic while Spencer's trajectory has more of a oscillatory shape which could account for SEPSPOT's reduced time to complete the trajectory.
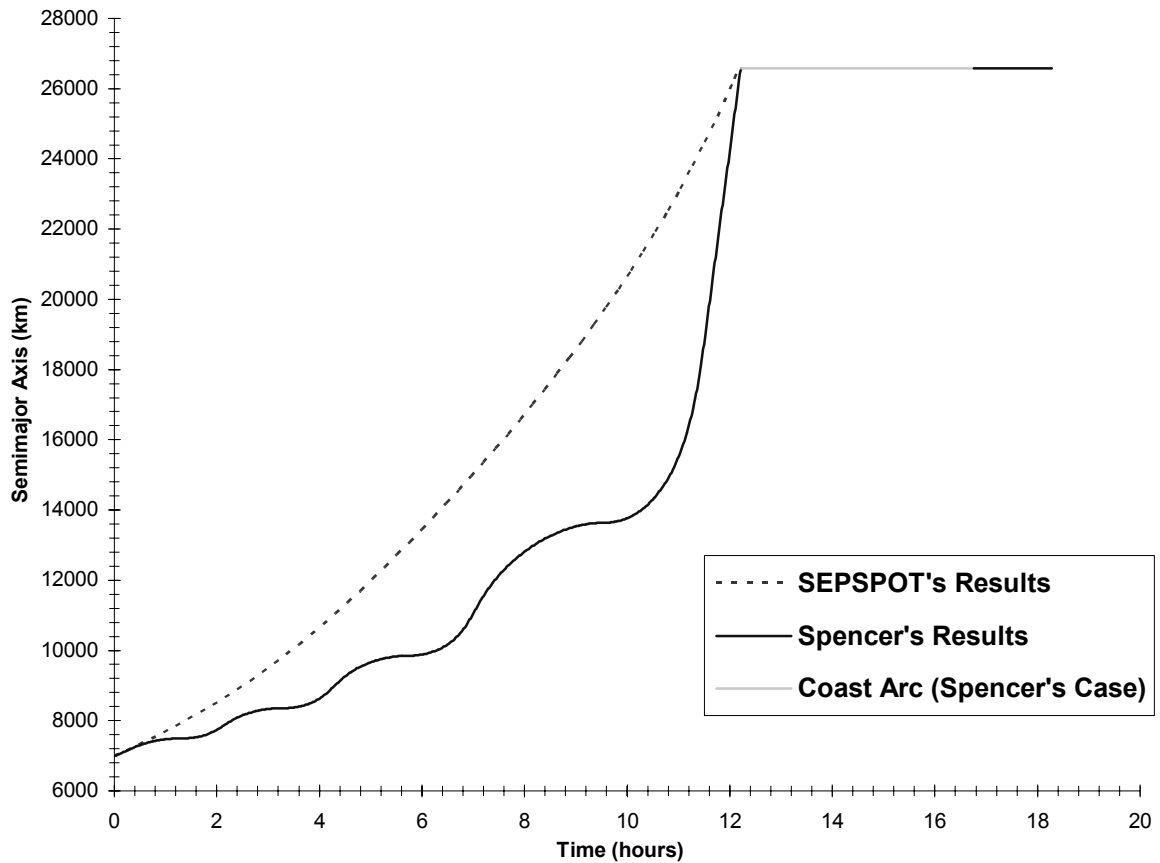
Figure 7: Semimajor Axis Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-1}$

Figure 8 shows how Spencer's eccentricity curve slightly oscillates, but increases at a steady rate while SEPSPOT's curve increases slowly at the beginning and then the eccentricity starts increasing at a faster rate, resulting in a smooth parabolic curve. The oscillations in Spencer's trajectory could account for additional time required to achieve a final eccentricity of 0.73646.
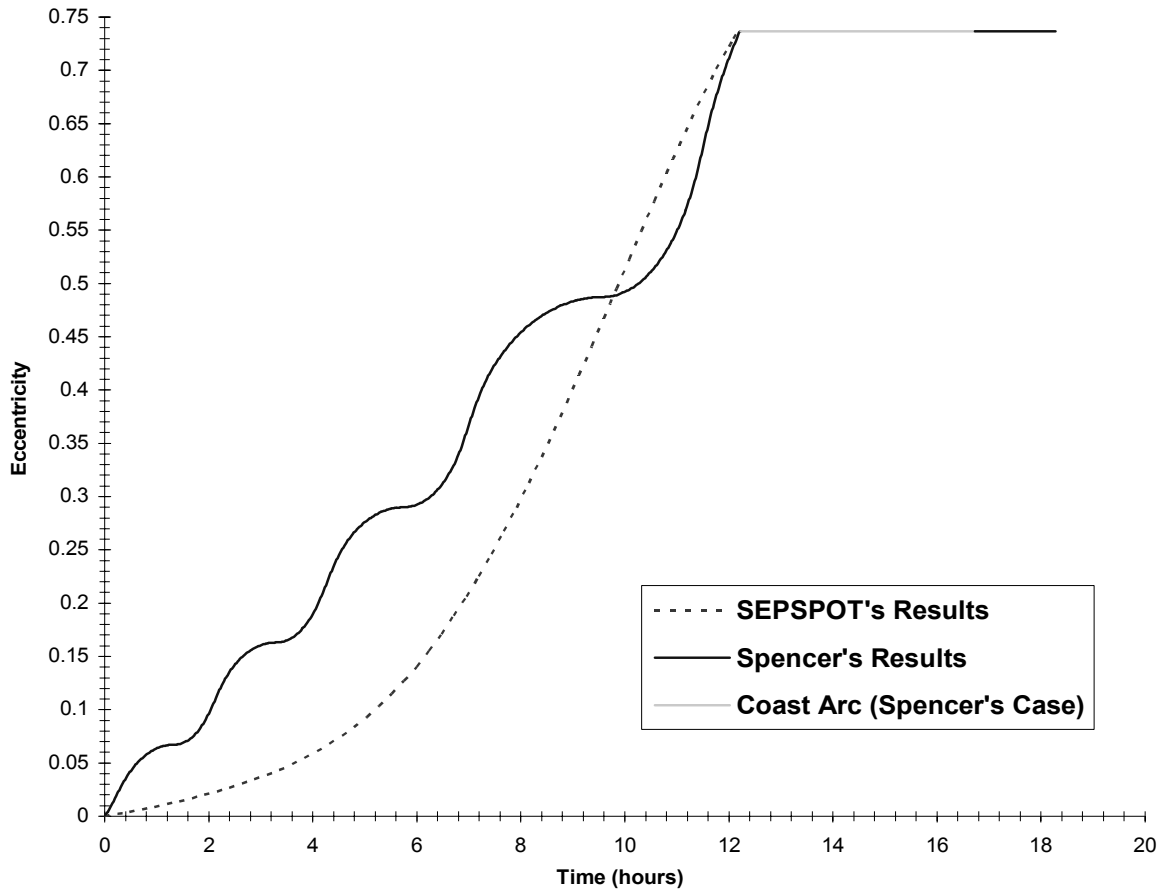
Figure 8: Eccentricity Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-1}$

Figure 9 illustrates how for Spencer's results the angle of inclination stays constant at 28.5° during the first burn and then it rapidly increases to 63.435° in the second burn. SEPSPOT's results show how the angle of inclination is increased over time and a parabolic curve is formed.
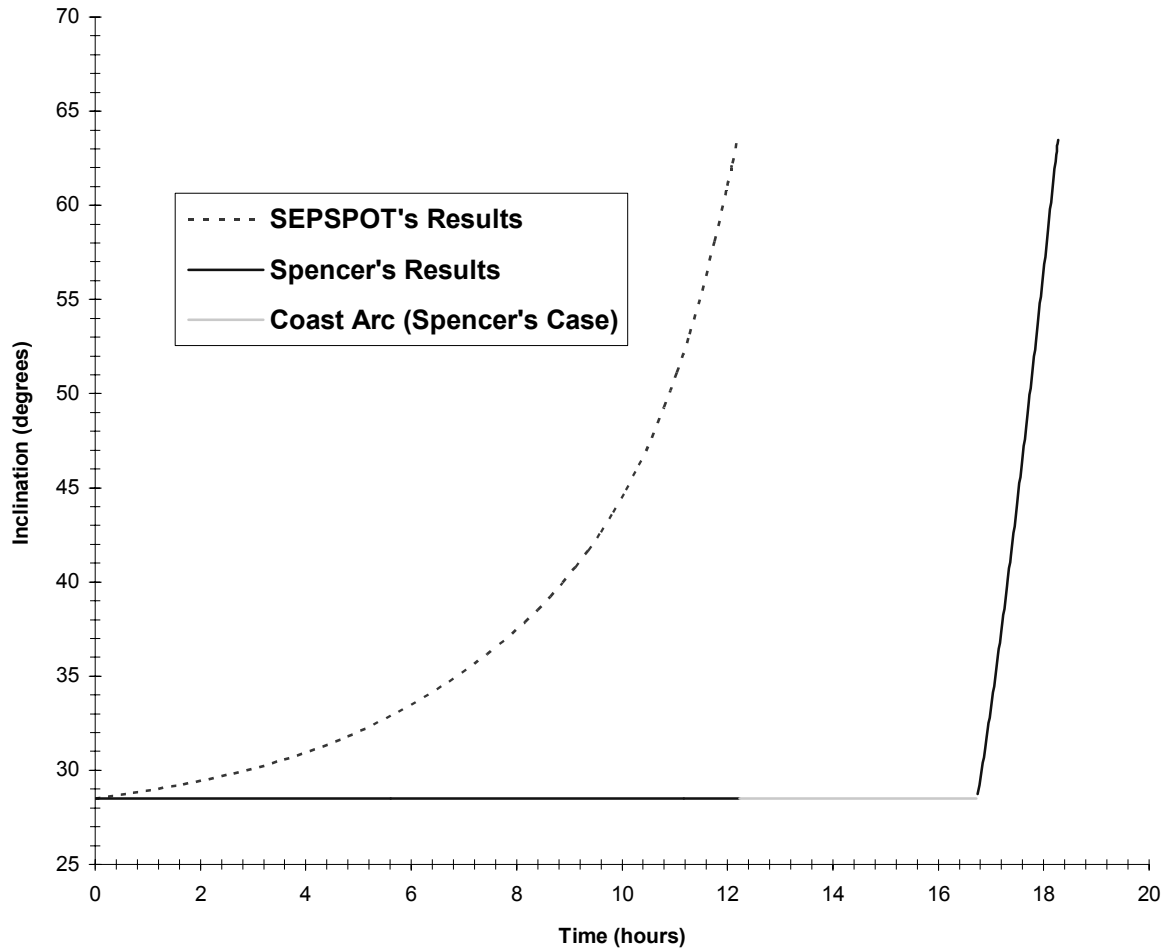
Figure 9: Inclination Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-1}$

Figure 10 shows how SEPSPOT's trajectory achieves both the apogee and perigee radius desired conditions in the amount of time it takes Spencer's trajectory to complete the first burn. An important aspect of this plot is that SEPSPOT keeps the perigee radius free which forms a parabolic type curve while Spencer maintains a constant perigee radius. This is a key difference since it might account for the optimal trajectory found by SEPSPOT.
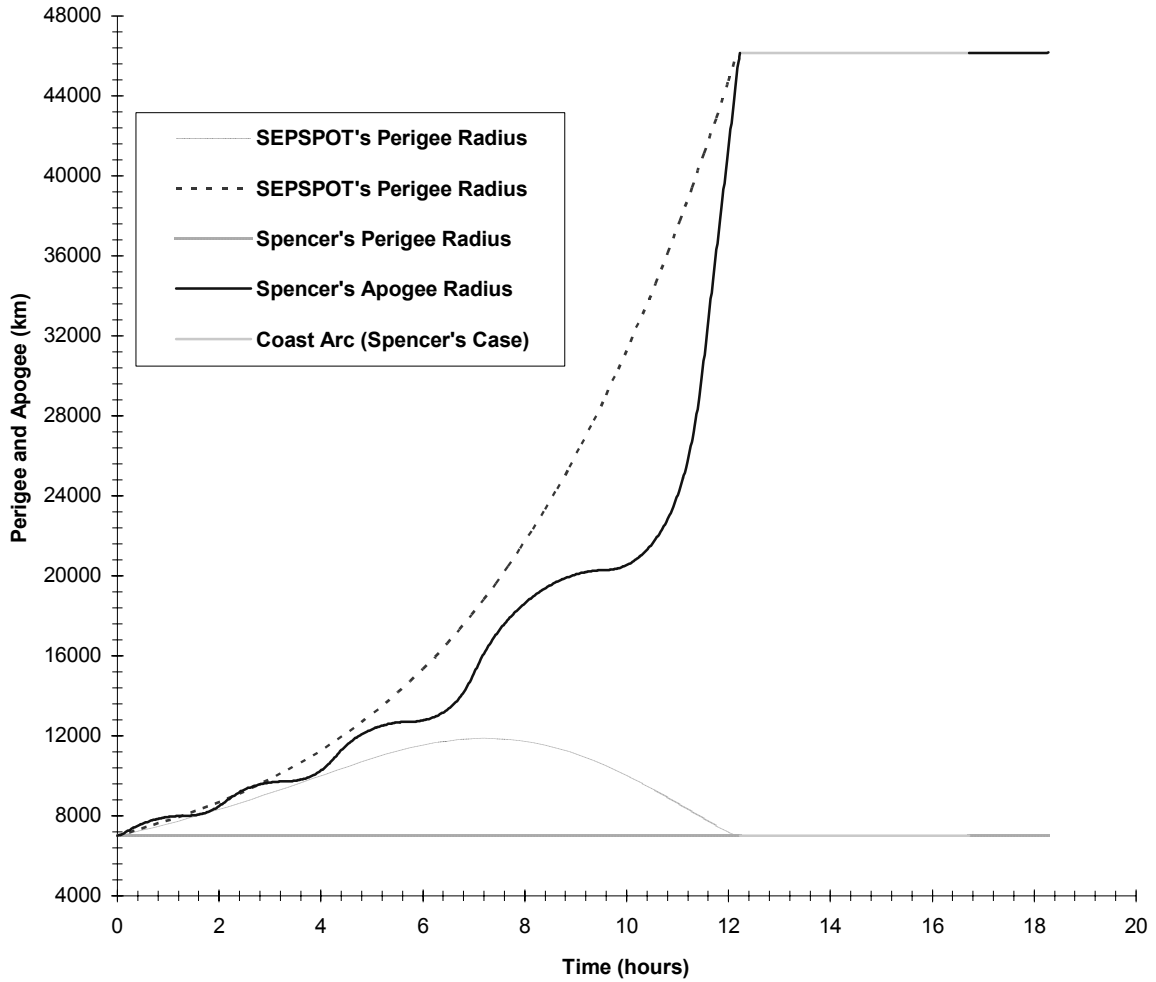
Figure 10: Apogee and Perigee Radius Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-1}$

From Figure 11 one can see that SEPSPOT's trajectory achieves the energy levels required to complete the trajectory in less time. In addition, from the plot Spencer's data indicates that more energy is required to accomplish the desired LEO-Molniya transfer.
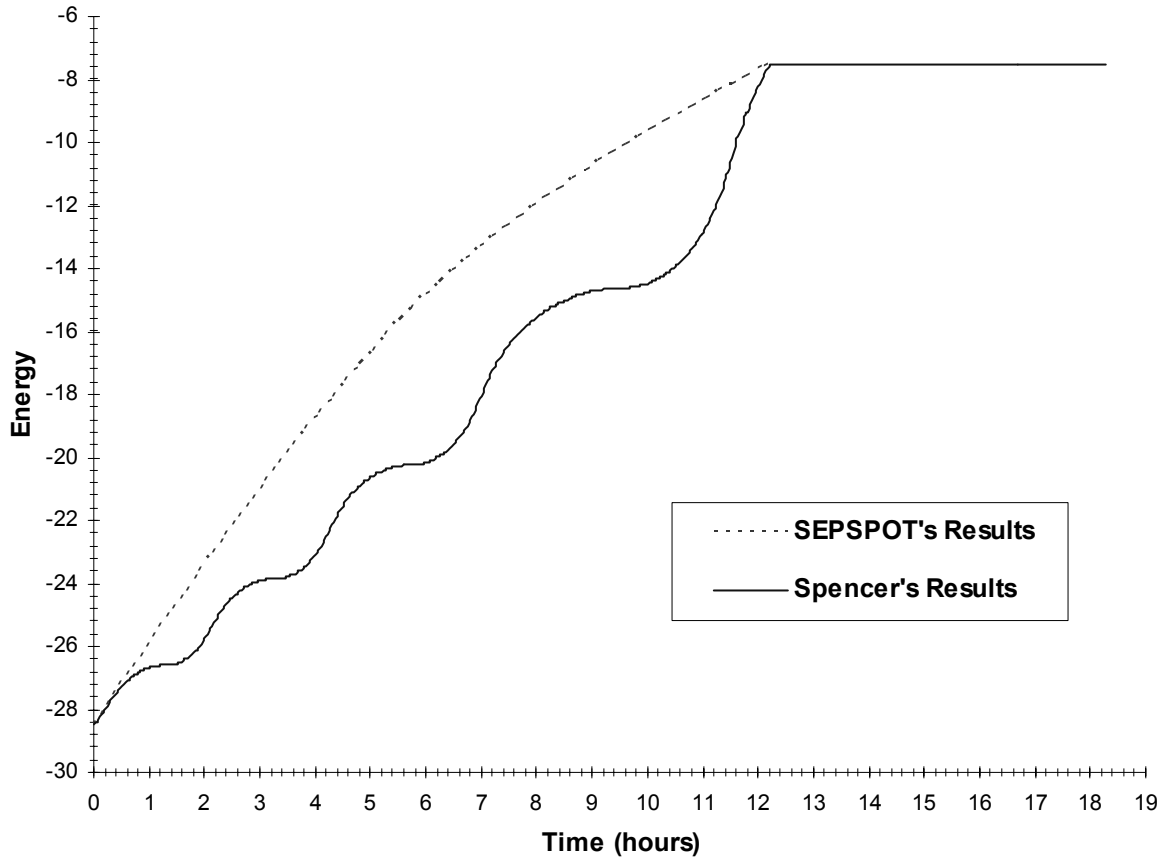
Figure 11: Energy Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-1}$

4.2   LEO-Molniya Transfer: $T/m_0 = 10^{-2}$ N/kg

In the case of an initial thrust acceleration of 10^{-2} N/kg, SEPSPOT completes the transfer in one burn while achieving a ΔV of 5814.69 m/s in 121.77 hours. Spencer's analytical solution achieves the transfer using one burn in 136.83 hours at a ΔV of 6844.00 m/s. Table 9 depicts a comparison of the two solutions in which SEPSPOT's trajectory is shown to be slightly more optimal by completing the transfer in less time. Furthermore, there is a percent error of 12.37% between the numerical solution provided by SEPSPOT and Spencer's analytical solution.

Table 9 Initial Thrust Acceleration of 10^{-2} N/kg

	Overall Effective Change in Velocity (ΔV)	Overall Time	Percent Error (%)
Spencer's Results	**6844.00 m/s**	**136.83 hrs**	**12.37%**
SEPSPOT's Results	**5814.69 m/s**	**121.77 hrs**	

Figures 12-16 illustrate a comparison of the time history of the semimajor axis, eccentricity, inclination, apogee and perigee radius, and energy between SEPSPOT's numerical data and Spencer's analytical data. One should note that unlike the previous case, both SEPSPOT's and Spencer's trajectories manage to achieve the desired orbit using only one burn.

Figure 12 shows that despite the fact that both trajectories only require one burn to complete the desired final conditions, Spencer's results are still taking longer to reach the final conditions. In this case SEPSPOT's semimajor axis curve is smooth and parabolic while Spencer's curve has a similar shape, but it contains small oscillations; these oscillations can account for the additional time required to reach the Molniya orbit.
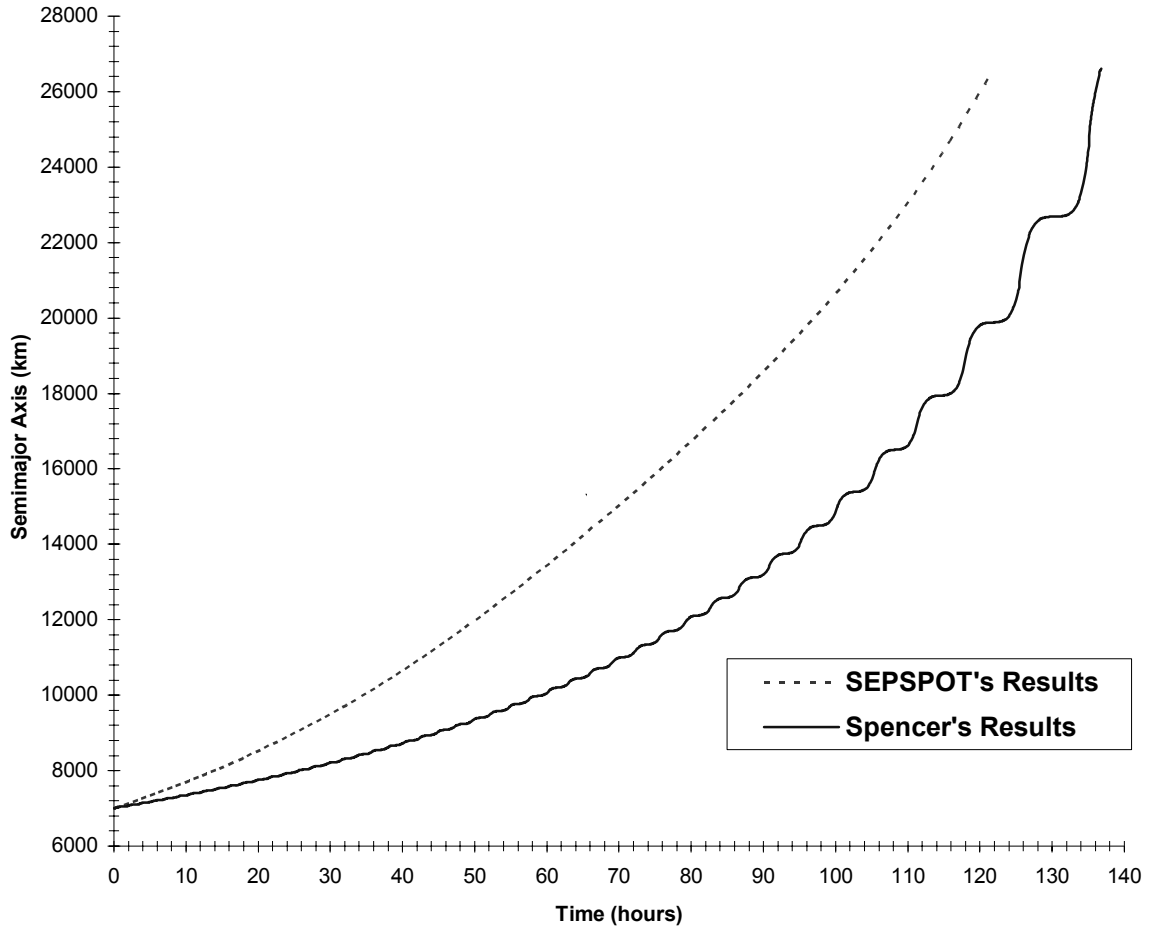
Figure 12: Semimajor Axis Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-2}$

Figure 13 shows how for Spencer's results the eccentricity seems to increase at an almost constant rate with time. Furthermore, for SEPSPOT's results, like in previous initial thrust acceleration case, the curve increases slowly at the beginning and then the eccentricity starts increasing at a faster rate, resulting in a smooth parabolic curve.
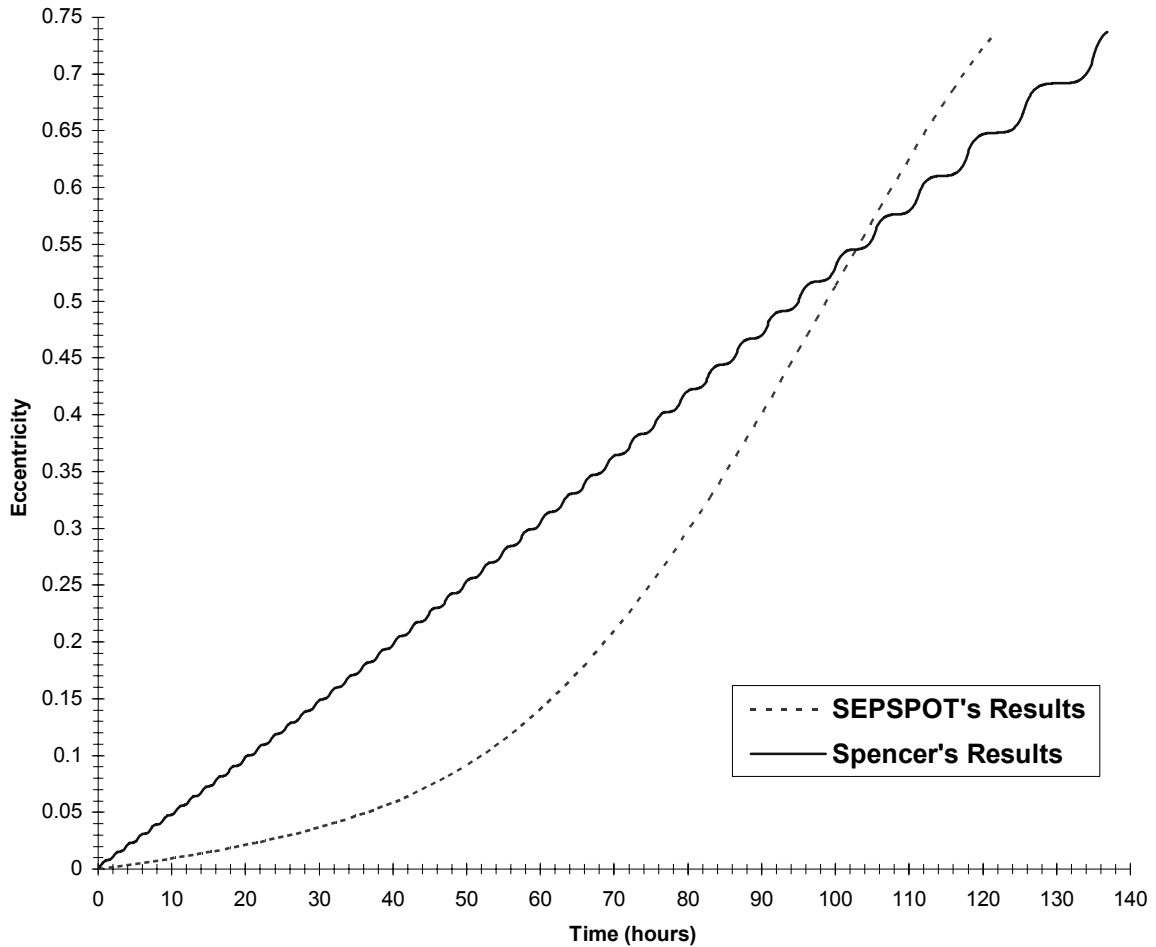
Figure 13: Eccentricity Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-2}$

Figure 14 illustrates how for both Spencer's and SEPSPOT's results, the angle of inclination is increased over time and a parabolic curve is formed. The only difference is that SEPSPOT's curve is smooth while Spencer's curve has oscillations. In addition, the plot depicts how the angle of inclination for SEPSPOT's trajectory increases at a much faster rate than Spencer's trajectory, hence reaching the desired angle of inclination in less time.
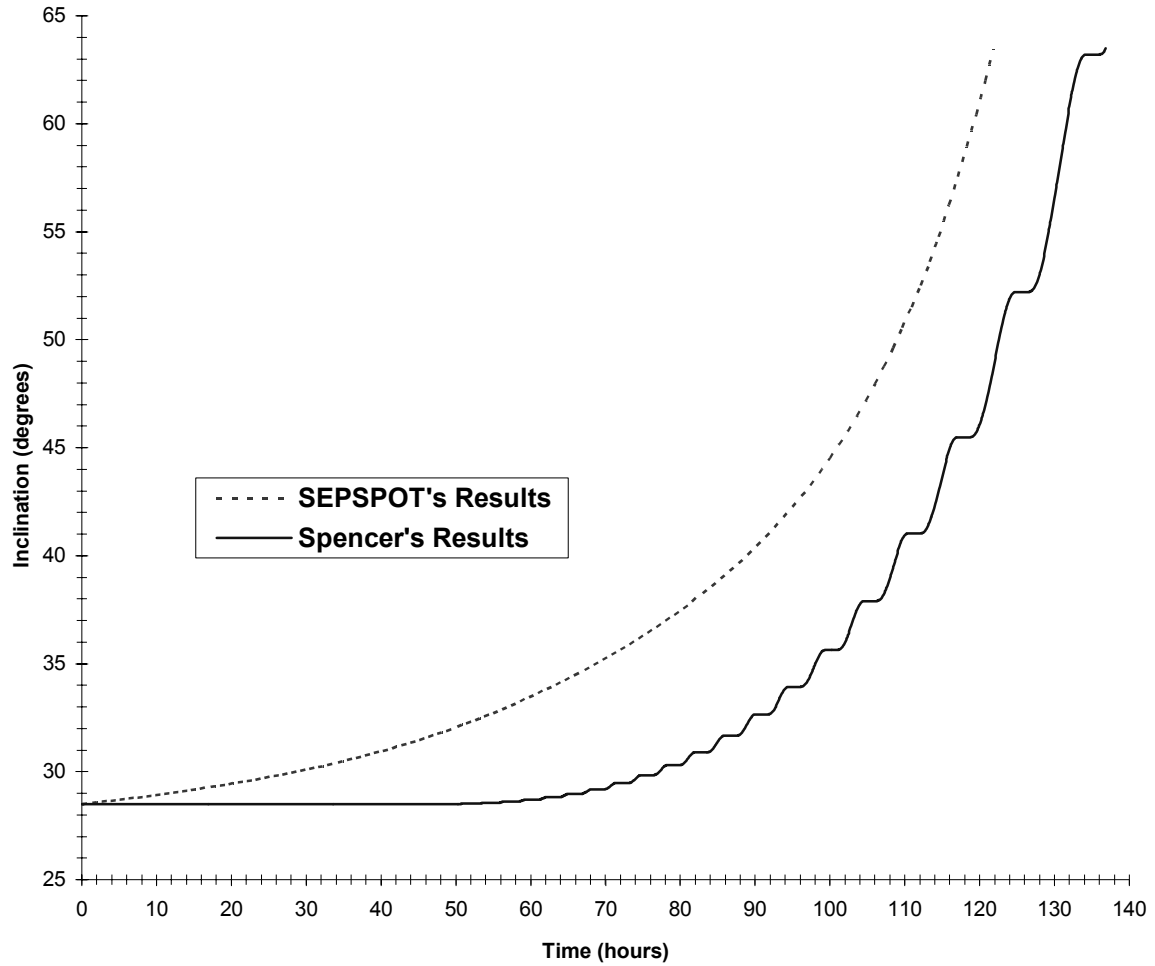
Figure 14: Inclination Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-2}$

Figure 15 illustrates how SEPSPOT keeps the perigee radius free and forms a parabolic type curve while Spencer maintains a constant perigee radius. This is a key difference since it might account for SEPSPOT's more efficient trajectory.
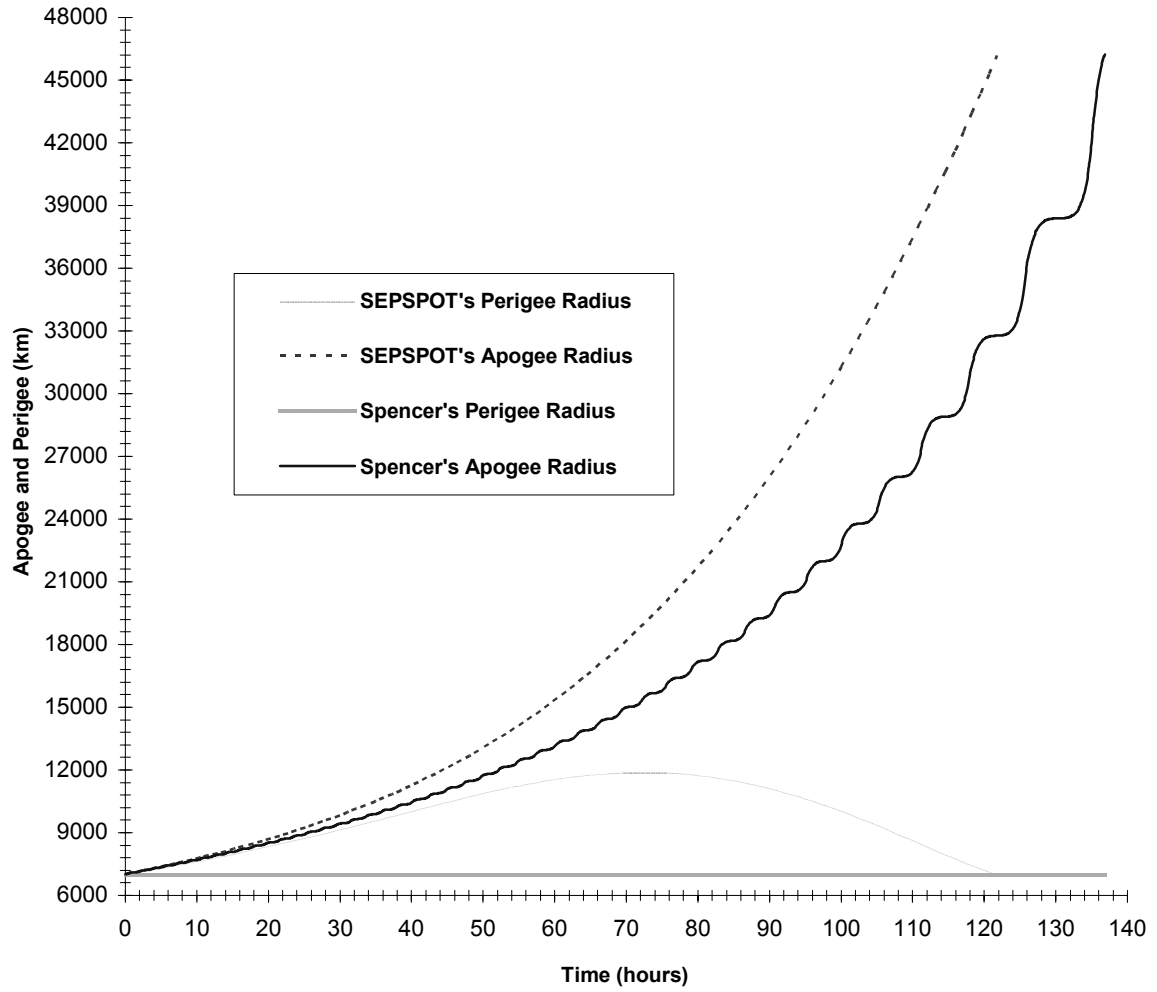
Figure 15: Apogee and Perigee Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-2}$

Figure 16 shows how SEPSPOT's trajectory achieves the energy levels required to complete the trajectory in less time. In addition, from the plot Spencer's data indicates that more energy is required to accomplish the desired LEO-Molniya transfer.
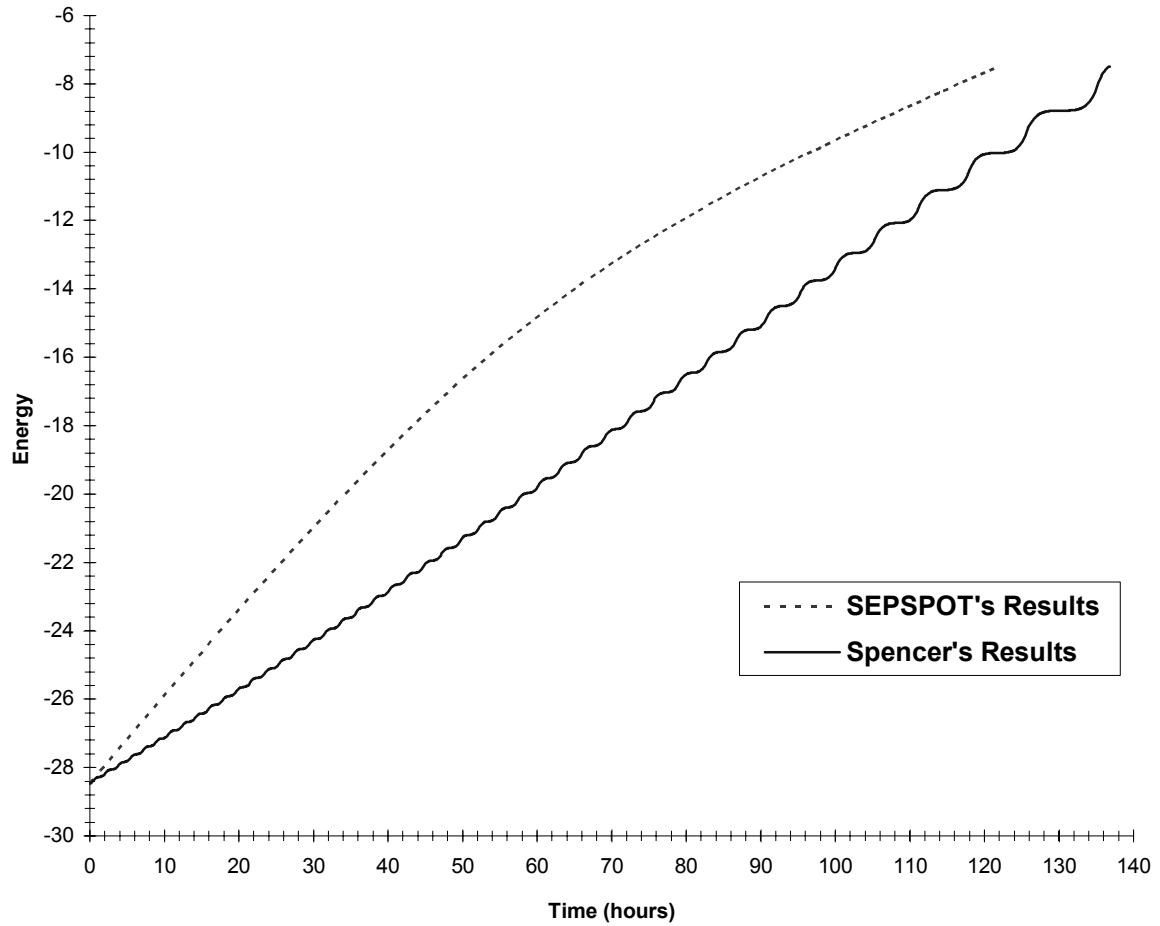
Figure 16: Energy Time History for Three-Dimensional, LEO-Molniya Transfer, $T/m_0 = 10^{-2}$

Chapter 5

CONCLUSIONS

5.1 SEPSPOT Evaluation

SEPSPOT proved to be a very useful and effective tool for the evaluation of Spencer's analytical data. However, one disadvantage is that in order to be able to run the program the user must have knowledge of the costates and of the total transfer time. If the values inputted into SEPSPOT are not close to the actual values then the program will not converge or it will converge to the incorrect solution. This means that SEPSPOT is only useful for evaluating transfers where the user has an idea of how long the transfer should take.

Also, another disadvantage is that the program does not have a graphical interface and the data is inputted and outputted in the form of a data file. SEPSPOT is very particular about the format of the input data, and if the format is incorrect the program will give an error message or it will run and output useless data files (empty).

Despite these disadvantages, one great advantage is that there is a great amount of flexibility in running SEPSPOT since the SSH Secure Shell Program can be run from any Windows based operating system that has internet capabilities. This allows the user to run and execute SEPSPOT from any Windows PC whether it's in a computer lab or the users own home.

5.2 Spencer's Analytical Data Evaluation

A comparison between Spencer's analytical data and SEPSPOT's numerical data showed that the percent error between the two is very small, there is about a 13% percent difference for the initial thrust acceleration of 10^{-1} N/kg and 10^{-2} N/kg. The factor that affected the results of both solutions is that for Spencer's analytical solution the radius of perigee was held constant while with SEPSPOT's numerical solution the radius of perigee was free. This difference indicates that while Spencer's solution was closely related to SEPSPOT's solution it can be improved by doing another analytical analysis with the radius of perigee free.

5.3 Future Work

During this study, only two cases were evaluated using SEPSPOT (TA = 10^{-1} N/kg and TA=10^{-2} N/kg). The cases for initial thrust acceleration of 10^0, 10^1, 10^2, 10^3, 10^4, and 10^5 N/kg will be run in SEPSPOT at a later time to evaluate Spencer's analytical solutions for these cases. After evaluating all these cases another project will be conducted in which new analytical solutions will be computed for the LEO-to-Molniya transfer where the radius of perigee will be free instead of constant.

REFERENCES

1. Herman, Albert L. and Spencer, David B., "Optimal, Low-Thrust Earth-Orbit Transfers Using Higher-Order Collocation Methods," *AIAA Journal of Guidance, Control, And Dynamics*, Vol. 25, No. 1, 2002, pp.40-47.

2. Martinez, Carolina. <u>Ion Engine to Open Up the Solar System.</u> 3 Sep. 2002. National Aeronautic and Space Administration. 23 May 2003. <http://www.jpl.nasa.gov/technology/features/ion.html>.

3. Oglevie, R. E., Andrews, P.D., and Jasper, T.P., "Attitude Control Requirements for an Earth Orbital Solar Electric Propulsion Stage," AIAA Paper 75-353, March 1975.

4. Spencer, David B., "An Analytical Solution Method For Near-Optimal, Continous-Thrust Orbit Transfers." Ph.D. diss., University of Colorado, 1994

5. Meyer, Rudolf X., "Solar-Electric Ion Propulsion: Future Possibilities," *The Journal of the Astronautical Sciences*, Vol. 47, Nos. 1 and 2, 1999, pp. 47-52.

6. Brusch, R.G., and Vincent, T.L., "Low-Thrust, Minimum Fuel, Orbit Transfers," Acta *Astronautica*, Vol. 16, No. 2, 1971, pp.65-73.

7. Edelbaum, T.N., "Propulsion Requirements for Controllable Satellites," *ARS Journal*, Vol. 32, No. 8, 1961, pp.1079-1089.

8. Redding, D. C., "Optimal Low-Thrust Transfers to Geosynchronous Orbit," Stanford Univ. Guidance and Control Lab., Rept. SUDAAR 539, Stanford, CA, Sept. 1983.

9. Herman, Albert L., and Conway, Bruce A., "Optimal, Low-Thrust, Earth-Moon Orbit Transfer," *AIAA Journal of Guidance, Control, And Dynamics*, Vol. 21, No. 1, January-February, 1998, pp. 141-147.

10. Prussing, John E., "A Class of Optimal Two-Impulse Rendezvous Using Multiple-Revolution Lambert Solutions," *The Journal of the Astronautical Sciences*, Vol. 48, Nos. 2 and 3, April-September, 2000, pp. 131-148.

11. Kechichian, Jean Albert, "Minimum-Time Low-Thrust Rendezvous and Transfer Using Epoch Mean Longitude Formulation," *AIAA Journal of Guidance, Control, And Dynamics,* Vol. 22, No. 3, May-June, 1999, pp. 421-432.

12. Coverstone-Carroll, Victoria and Williams, Steven N., "Optimal Low Thrust Trajectories Using Differential Inclusion Concepts," *The Journal of the Astronautical Sciences,* Vol. 42, No. 4, October-December, 1994, pp. 379-393.

13. Betts, John T., "Optimal Interplanetary Orbit Transfers by Direct Transcription," *The Journal of the Astronautical Sciences,* Vol. 42, No. 3, July-September, 1994, pp. 247-268.

14. Kechichian, Jean Albert, "Optimal Low-Earth-Orbit-Geostationary-Earth-Orbit Intermediate Acceleration Orbit Transfer," *AIAA Journal of Guidance, Control, And Dynamics*, Vol. 20, No. 4, July-August, 1997, pp. 803-811.

15. Cnyang, "LEO/MEO/GEO/HEO." 10 Oct. 2001. Security and Algorithm Laboratory. 25 Sept. 2003. <http://www.sna.csie.ndhu.edu.tw/~cnyang/sat/sld008.htm>.

16. Sackett, Lester L., Malchow, Harvey L., and Edelbaum, Theodore N., <u>Solar Electric Geocentric Transfer with Attitude Constraints: Program Manual</u>. Cambridge, Massachusetts: The Charles Stark Draper Laboratory, Inc., August, 1975.

17. Sackett, Lester L., Malchow, Harvey L., and Edelbaum, Theodore N., Solar Electric Geocentric Transfer with Attitude Constraints: Analysis. Cambridge, Massachusetts: The Charles Stark Draper Laboratory, Inc., August, 1975.
18. Jahn, Robert G., Physics of Electric Propulsion. New York: McGraw-Hill Book Company, 1968.

Diversity Blues: The Lack of Racial Justice Within the Race-Conscious Admissions Debate

Christopher S. Arlene, McNair Scholar, Pennsylvania State University

Faculty Research Advisers
Dr. Susan Searls-Giroux
Senior Lecturer in English and Education

Dr. Henry Giroux
Waterbury Chair Professor in Secondary Education
Pennsylvania State University

Introduction

Within the debate on the legality of affirmative action in America, these have indeed been interesting times. The Supreme Court recently found that the affirmative action policies being used in the admissions process by the University of Michigan Law School to be constitutional. In doing so the Court, in a 5-4 decision, gave powerful legal support to the goal of campus diversity that was first made prominent by Justice Lewis Powell in *Regents of the University of California v. Bakke* (1978). And while the Court shot down the mechanical system of giving Black, Latino, and Native American applicants twenty points on a 150-point scale at the undergraduate level, the decisions have largely been viewed as victories for racial inclusion.[1]

But, unfortunately, the decisions handed down by the Court will not have the enormous effects that many supporters of affirmative action believe that they should.[2] The reason for this lays in the fact that, however important the decisions in the Michigan cases may be, the commitment to diversity being promoted at the college level as reasoning for the decisions lacks one critical ingredient: a true commitment to racial justice. The ideals of racial equality, an attempt to legitimately equalize a playing field that heavily favors Whites and strongly discriminates against Blacks, Latinos, and Native Americans, have been almost entirely ignored within the debate concerning race-conscious admissions policies.

The lack of discussion concerning the need for racial justice in this particular debate is extremely troubling for a number of reasons, but most importantly because the ultimate desire for racial equality seems nonexistent within mainstream public debates. And as long as this is the case, then the victory for the University of Michigan is really nothing more than powerless window dressing that allows us to continue our silence concerning the real issue behind the affirmative action debate: past and present racial discrimination. Until we begin to talk about past and present discrimination against people of color, then we will never be able to publicly deal with the deeper issues that require race-conscious admissions policies to be necessary.

---

[1] Schmidt, Peter. "Affirmative Action Survives, and So Does the Debate," *The Chronicle of Higher Education*. July 2003: S1.
[2] Washington, Wayne & Globe Staff. "Supporters of Affirmative Action Are Pleasantly Surprised," *The Boston Globe*. 24 June. 2003: A13.

I want to focus on two specific reasons why the decisions by the Court should not be looked at as historic, or as a step forward in the journey for racial equality, or even as an assurance that minority students will continue to have access to higher education. The reasons for this current state of affairs exemplify a number of different institutional problems that plague most minorities in this country, and they should be examined as being important elements of a social system based on discrimination and disadvantage.

First, supporters of diversity on America's college campuses include some of the most influential political, business, and military leaders in the nation, and one would be hard pressed to find anyone who feels that diversity is not an enviable goal. But, nowhere in the argument for supporting race-conscious admissions policies is the idea that American society's discriminatory tendencies must be talked about and solved.

The diversity rationale is primarily based on the idea that we live in a multiracial, multicultural society, and that because of this we must be able to learn about and work with each other.[3] There is the acknowledgement that Black, Latino, and Native American people have been discriminated against in the past, but such abuses have been rectified. For the purposes of securing another twenty-five years of affirmative action policies, those in the public debate are willing to forget that we not only have to make up for a discriminatory past, but also an extremely discriminatory present.

Second, the celebrated 1954 *Brown v. Board of Education* decision that found segregated education illegal and forced integration in America's public schools has been, for the most part, repealed by a number of Supreme Court decisions in the last thirty years. Due to a number of important decisions made by the Court in the early 1970's and the 1990's, we are currently faced with an American public school system that is once again based on an unequal form of segregation.[4] The current state of public education is eerily reminiscent of a time when "separate but equal" was a legally supported social idea.

While there is obviously truth to the idea that a diverse student body is a worthwhile goal for any institution of higher learning, the segregation of K-12 education in this country has created the unfortunate situation where most students do not have experience going to school with students of different races. And the implications of this situation are quite serious because segregation leaves minority students seriously disadvantaged in comparison to White students in regards to the quality of their education. For example, Black students consistently score lower than White students on assessment tests throughout elementary and secondary school.[5]

Segregation in America's schools forces most Black and Latino students to attend inferior schools because informal segregation is a function of both race and class politics. Since schools that are intensely Black and Latino are fourteen times as likely to be high poverty schools than schools with 90% White students,[6] the overlap between race and class cannot be overlooked. As Gray Orfield, Co-Director of the Civil Rights Project at

---

[3] "Right Ruling, Wrong Reason." Editorial. *The Washington Post* 29 June. 2003: B3.
[4] Orfield, Gary. *Dismantling Desegregation: The Quiet Reversal of Brown v. Board of Education.* New York: The New Press, 1996. 1.
[5] National Center for Education Statistics. "Educational Achievement and Black-White Inequality." July 2001. 60. <http://nces.ed.gov/pubs2001/2001061.PDF>.
[6] Orfield, 55

Harvard University, clearly explains this problem in his book *Dismantling Desegreation* when he writes:

> The intense segregation of minority and low-income students in urban schools is a critical factor in analyzing educational opportunity because it is systematically connected to patterns of low achievement. In virtually every large metropolitan area studied that lacks city-suburban desegregation, low-income minority students and middle-class white students attend schools tat are not only separate, but profoundly unequal.[7]

Tragically, given such conditions, the role of education will only grow more important as we move into an economy based on technological expertise and information. While in the past education certainly did separate those who had and those who had not, there was still an important and meaningful role for workers who were not highly educated, but that is not the case anymore.

The reality of the situation is that earning an undergraduate degree no longer ensures the same amount of material success and opportunity for upward mobility that it once did,[8] and as post-graduate education becomes a necessity poor minority students find themselves at an even deeper disadvantage. For example, in 1999 the average annual income for a college graduate was $45,400, while just $25,900 for high school graduates and $18,900 for high school dropouts.[9] As long as most minority students are subjugated to schools that do not prepare them for academic or economic success many of the current inequalities related educational achievement, employment, and income will continue.

The role of the judiciary also plays an important role within this discussion, because the precedent that has been set by the Supreme Court, and generally followed by the lower courts, is one that has all but eliminated the courts as a place where disadvantaged minorities can attempt to solve problems of educational inequity. The publicity brought about by the Michigan decisions was largely positive for liberal supporters of affirmative action, although some conservatives have claimed the decisions as victories as well.[10] But, however promoted by the media, the Court's decision has to be taken with a grain of salt and examined within the context of its past decisions. The reversal of *Brown* shows the true colors of the very conservative Rhenquist Court, and the decisions in the Michigan cases should be understood within this context.

Ultimately, the decisions will not have any real effect towards the promotion of racial equality and justice as ideals at the university, as well as on the enrollment of minority students at our most prestigious colleges and universities. While the Court gave legal approval for the attainment of a culturally and racially diverse student body through the use of affirmative action policies, the Michigan decisions in no way ensure equality in education, equality in opportunity, or commitment to fixing our deeply rooted racial problems. And on the eve of *Brown's* 50th anniversary, it seems that we are still

---

[7] Orfield, 65

[8] National Alliance of Business. "Workforce Economics Trends." May 2000. Retrieved on September 29, 2003 from <http://www.nab.com/PDF/wft_may2000.pdf>.

[9] U.S. Census Bureau, 2002, retrieved on September 23, 2003 from http://www.csun.edu/~hfoao102/@csun.edu/csun02-03/csun1007_02/census.html.

[10] Kahlenberg, Richard. "The Conservative Victory in *Grutter* and *Gratz*." *Jurist: The Legal Information Network.* 5 Sep. 2003. <http://jurist.law.pitt.edu/forum/symposium-aa/kahlenberg.php>.

searching for answers on how to implement the most fundamental and necessary step in the journey for equality: equal access to a quality education.

Historical Background: The Bakke Decision

The cases against the University of Michigan cannot thoroughly be understood without situating them within the history of affirmative action policies in the realm of college admissions. And there has been no more important case in the legal history of this issue than that of the *Board of Regents of the University of California v. Bakke*. The historical Supreme Court decision dealt with a white student, Allan Bakke, who applied multiple times to the University of California-Davis Medical School and was repeatedly denied admission despite outstanding grades and standardized test scores. He argued that because the University was setting aside a fixed amount of openings for its incoming class for racial minorities that he was being unfairly discriminated against because of his race.[11]

The Court's decision in *Bakke* was not unanimous, decisive, or even clear, and it is because of the lack of clarity in the Court's decision that the issue of using race in the admissions process is once again being brought before the Court. The nine-member Court found itself split, with four Justices finding that race can legally be considered in the admissions process, while four other Justices found that the quota system used by the University violated Title VI of the Civil Rights Act of 1964. So with the Court split, Justice Lewis Powell became the deciding vote, and his interpretation of the legality of the program has set the standard for affirmative action in college admissions up to this point.[12]

Justice Powell believed that the system of set-asides used by the University was indeed a violation of the Civil Rights Act of 1964 and therefore ordered that Bakke be admitted to the University. But, he also found that race can indeed be a factor in the admissions process, officially ruling 5-4 in favor of the University. In his majority opinion Justice Powell wrote that student body diversity "clearly is a constitutionally permissible goal for an institution of higher education,"[13] essentially making it legal for colleges and universities to use affirmative action.

What the decision left the academic world with is the idea that student diversity is an important and worthwhile goal, but at the same time quota systems were found illegal. Therefore race was supposed to be only one of many different factors considered in the admissions process.[14] Race cannot legally be used in the mechanical way that the quota system at the University was used, but when looking at an individual applicant race can be considered amongst other criteria, such as academic potential, economic background, and community service.[15]

It is basically within this context that the Michigan cases are debated, because the *Bakke* decision was the basis for most prestigious institutions of higher education to

---

[11] *Regents of the University of California v. Bakke*, 438 U.S. 265 (1978).

[12] Spann, Girardeau A. *The Law of Affirmative Action: Twenty-Five Years of Supreme Court Decision on Race and Remedies*. New York University Press, 2000. 17.

[13] 438 U.S. 265 (1978).

[14] Wightman, Linda. "Standardized Testing and Equal Access: A Tutorial," *Compelling Interest: Examining Evidence on Racial Dynamics in Colleges and Universities*. Eds. Mitchell J. Chang, et. al. Stanford University Press, 2003.

[15] Girardeau, 16.

implement affirmative action policies. While the decision did not necessarily create a road map or a strict set of guidelines for using affirmative action in the admissions process, it did give colleges and universities across the country the ability to create diversity of in the student population. And ultimately, the basis for the Michigan cases is that the plaintiffs felt that the University's individual system for creating diversity was unfair and illegal.

The Michigan Decisions

The University of Michigan is, without question, one of America's most repubutable and elite institutions of higher education. The elite status of many of its schools and departments, both for undergraduate and graduate programs, has clearly situated the University with other elite public universities, such as the University of California-Berkeley, the University of North Carolina-Chapel Hill, and the University of Virginia. And it is because of the University's growing prominence as a superior academic institution that the complaints filed against the University in the fall of 1997 were so serious in nature.

Barbara Grutter, a White woman who was denied admission to the University's Law School (Law School), filed a lawsuit, *Grutter v. Bollinger, et. al*., against the University that claimed that she was unfairly denied admission because she was White.[16] Jennifer Gratz and Patrick Hamacher, White applicants who were denied admission to the University's flagship undergraduate college, the College of Literature and the Sciences and the Arts (LSA), also filed a suit against the University, *Gratz, et. al. v. Bollinger, et. al*., claiming that they had been denied admission because they were White, which is a violation of the Equal Protection Clause of the Fourteenth Amendment of the United States Constitution.[17]

The basis for the plaintiff's claims in these two cases rested in the University's race-conscious admission policies. The University, as most other national universities do, used affirmative action policies during its admissions process, with the ultimate goal of creating a racially and ethnically diverse student body.[18] Due to the magnitude of applicants the University receives for its undergraduate program, usually over 25,000 applications for 5,000 spaces, it used a 150-point system to judge applicants. After examining characteristics such as grade-point-average, SAT/ACT scores, community service experience, extracurricular activities, and many others, if an applicant scored 100 points he or she was offered admission.

The fatal flaw of the undergraduate program's point system rested in the fact that members of specified underrepresented racial and ethnic groups, in this instance Black, Latino, and Native Americans, were automatically given twenty points. The argument against this aspect of the point system was grounded in the idea that minority applicants were receiving an unfair and undeserved advantage in the admissions process over their White counterparts.[19] And while the Law School did not use a point system for its admissions process, instead being able to utilize a more holistic admissions review, it

---

[16] *Grutter v. Bollinger, et al.* 28 U.S.C. (1997).
[17] *Gratz, et al. v. Bollinger, et al.* 28 U.S.C. (1997).
[18] *Gratz, et al. v. Bollinger, et a*l. U.S. 02-516 (2003).
[19] *Gratz,* U.S. 02-516 (2003).

openly used affirmative action policies to foster the racial and ethnic diversity in its student body that the University believed was a "compelling state interest."[20]

The cases eventually made their way to the Supreme Court, and the legality of student body diversity as a "compelling state interest" is what the Court eventually based their "monumental" decision on. The Court, in what was generally hailed by supporters of affirmative action as a significant victory, ruled in favor of the Law School and against LSA. In a 5-4 decision with Justice Sandra Day O'Connor as the swing vote, the Court found that diversity was indeed a compelling state interest, and that the Law School's admission process, based in a more holistic overview of applicants, did not violate the Equal Protection Clause.[21] In a 6-3 decision, though, the Court ruled against LSA because its mechanical point system was "not narrowly tailored" enough to survive strict scrutiny,[22] meaning that blindly giving any applicant one-fifth of the points needed for admission simply because he or she is an underrepresented minority is unconstitutional.

While the decisions may mean different things depending on whose side one is on, the Court's rulings have some very concrete answers to questions that had been plaguing higher education ever since the *Bakke* decision. That decision, in which Justice Powell stated that the "diversity that furthers a compelling state interest encompasses a far broader array of qualifications and characteristics of which racial or ethnic origin is but a single though important element,"[23] is the basis for the adoption of non-quota affirmative action policies by the University, and higher education in general.

First, the Court's deciding in favor of the Law School gives new legal strength to the *Bakke* decision, and essentially gives legality to fostering diversity at the university level through the use of race-conscious affirmative action policies. The plaintiff's argument in both instances was based on the idea that diversity, certainly a worthwhile goal, is not such an important governmental interest that it warrants the use of, as conservative opponents would argue, "reverse discrimination." But as Justice O'Connor phrased it in her majority opinion, "education. . . must be inclusive of talented and qualified individuals of every race and ethnicity, so that all members of our heterogeneous society may participate in the educational institutions that provide the training and education necessary to succeed in America."[24]

Second, just as Justice Powell dictated in *Bakke*, race can certainly be examined and considered during the admissions process, but it has to be examined along with other important applicant characteristics. This is why the undergraduate program was shot down, because the point system did not look at race as an equal factor in the admissions process. What this means for admissions offices across the country, especially at other flagship state universities such as the University of Texas-Austin and the Pennsylvania State University-University Park, is that they must somehow figure out how to expand their staffs in order to ensure the holistic judgment the Court is requiring.

Third, the Court does not view affirmative action policies as the ultimate answer for creating diversity at the university level, and it set a twenty-five year limit for the continuation of race-conscious policies. As Justice O'Connor wrote, it "has been 25

---

[20] *Grutter v. Bollinger, et al*. U.S. 02-241 (2003).
[21] *Grutter v. Bollinger, et al*. 539 U.S. 02-241 (2003).
[22] *Gratz et al. v. Bollinger, et al*. 539 U.S. 02-516 (2003).
[23] *Regents*, 438 U.S. 265 (1978).
[24] *Grutter*, 539 U.S. 02-241 (2003).

years since Justice Powell first approved the use of race to further an interest in student body diversity in the context of public higher education. . . We expect that 25 years from now, the use of racial preferences will no longer be necessary. . ."[25] With this stipulation attached to its decision, the Court has made it clear that affirmative action policies cannot be counted on in the future to ensure minority access to higher education.

The decisions in the cases against the University of Michigan will undoubtedly be considered in the same company with the likes of *Bakke* and *Brown* in the realm of American legal history. Just as the *Brown* case found separate but equal public education unconstitutional, the Michigan cases will probably be viewed as similarly significant in the effort towards ensuring minority access to important social institutions like higher education, and supporters of affirmative action have already claimed the Court's decision as a decisive victory.[26]

But while the Court's ruling should be seen as good news for supporters of race-conscious policies, it should not be viewed as a colossal victory for minority students in America. While the decision gives the appears to ensure minority access to our nation's prominent colleges and universities, there are two primary reasons why this appearance is little more than an misleading façade: the lack of a true commitment to diversity for the purposes of racial justice and the savagely unequal public education system due to a reversal of the Brown decision.

The Role of Diversity

The most prominent and powerful argument supporting the University of Michigan's race-conscious admissions policies was the importance of campus diversity, the creation of a mix of students with many different backgrounds with emphasis put on, but certainly not limited to, racial and ethnic heritage. The University used this strategy in the lawsuits against the admissions policies at both the Law School and LSA, and in the end the Court agreed with this argument, although not with the "mechanical" point system used in the undergraduate process.

Justice O'Connor wrote in her majority opinion that the University "has a compelling interest in attaining a diverse student body" and that "effective participation by members of all racial and ethnic groups . . . is essential."[27] So the Court clearly believes that America's diversity is not something that should simply be acknowledged, but rather that it should be embraced. And the Court is certainly not alone in supporting the magnitude of diversity as a new direction for our nation. Even President George W. Bush, who vehemently opposes affirmative action and spoke out against the University, stated that "diversity is one of America's great strengths."[28]

Two very powerful briefs were filed in support of the University, one signed by officers of major Fortune 500 companies and the other filed on behalf of a number of prominent retired military officers, that exemplify the genuine desire to incorporate diversity into society's most influential social institutions. Leaders of corporations such

---

[25] *Grutter*, 539 U.S. 02-241 (2003).

[26] Liptak, Adam. "Affirmative Action Proponents Get the Nod in a Split Decision," *New York Times*. 24 June. 2003: A26.

[27] *Grutter*, 539 U.S. 02-241 (2003).

[28] Bush, George W. "Statement by the President." Online Posting. 23 June. 2003. <http://www.whitehouse.gov/news/releases/2003/06/20030623.html>.

as Microsoft, Boeing, General Electric, and Proctor & Gamble endorse the idea that in order for students to "realize their potential as leaders, it is essential that they be educated in an environment where they are exposed to diverse people, ideas, perspectives, and interactions."[29] And former military leaders such as Gen. H. Norman Schwarzkopf argue that "a highly qualified, racially diverse officer corps educated and trained to command our nation's racially diverse enlisted ranks is essential."[30]

Most importantly, though, diversity is clearly and powerfully supported in the academic world of higher education for its personal and societal benefits. The educational benefits of diversity are based in the realization that we live in a country, and more importantly a global economy, that is racially, ethnically, and culturally diverse, and because of this fact the importance for cross-cultural communication skills, the ability to work with different types of people, and the breaking-down of stereotypes is essential for all future employees.[31] And, according to an expert report filed in support of the University by professor Patricia Gurin, diversity benefits all students because they "learn better in a diverse educational environment, and they are better prepared to become active participants in our pluralistic, democratic society once they leave such a setting."[32]

So, it seems that the importance and value of diversity in higher education is not something that can legitimately be argued against. Walter Feinberg explains the overall acceptance of diversity as a national virtue in his book *On Higher Ground* when he argues that "diversity appeals to certain conceptions of fairness and equity without appearing to blame the living members of one group for the historical misfortunes or injustices of another."[33] Diversity, the catch phrase that has no single definition, has therefore become an important aspect of the modern collegiate experience, especially for the white students who, for the overwhelmingly majority, do not have much experience going to class with minorities upon arriving on campus.[34]

But it is here where the first glimpse of the falsehoods of diversity, and its sister catch-phrase "multiculturalism," can be seen: the notion of diversity that supports affirmative action at the university level, an acceptance of multiculturalism as important and necessary, is not genuine at its core and does not care to legitimize the students or the subjects that it should be equalizing. This form of multiculturalism does not, in the words of scholar Henry Giroux, "address how material relations of power work to sustain structures of inequality and exploitation in the current racialization of the social order.[35] It is more centered around the acknowledgement of different races and different cultures, rather than focused on examining social problems of inequity and discrimination.

---

[29] Amicus Brief, "65 Leading Businesses in Support of Respondents," *Gratz, et. al. v. Bollinger, et. al* U.S. 02-241 (2003) and *Grutter v. Bollinger, et. al.* U.S. 02-516 (2003).

[30] Amicus Brief, "Consolidated Brief of Lt. Gen. Julius W. Becton, Jr., et. al. in Support of Respondents" *Gratz, et. al. v. Bollinger, et. al* U.S. 02-241 (2003) and *Grutter v. Bollinger, et. al.* U.S. 02-516 (2003).

[31] Milem, Jeffrey F. "The Educational Benefits of Diversity: Evidence From Multiple Sectors," *Compelling Interest*, Eds. Chang, et. al. 2003. 152.

[32] "The Compelling Need for Diversity in Higher Education," Gratz, et. al. v. Bollinger, et. al. 97-75231 (E.D. Mich.) and Grutter v. Bollinger, et. al. 97-75928 (E.D. Mich.). Jan. 1999. Copyright by the Regents of the University of Michigan.

[33] Feinberg, Walter. *On Higher Ground: Education and the Case for Affirmative Action*. Teachers College Press, Columbia University, 1998. 76.

[34] Frankenburg, Erica, Lee, Chungmei, and Orfield, Gary, "A Multiracial Society with Segregated Schools: Are We Losing the Dream?" The Civil Rights Project, Harvard University. Jan. 2003. 27.

[35] Giroux, Henry. *Impure Acts: The Practical Politics of Cultural Studies*. Routledge, 2000. 74.

The incorporation of multiculturalism does not rest in the inherent desire to realize that Black, Latino, and Native American students, and the cultural legacy associated with their unique ethnic and racial experiences, are equal to their White counterparts. But instead, it rests in the belief that diversity is an essential aspect of the educational experience in order to prepare future employees, as opposed to molding critical citizens who are committed to democratic values of equality and social justice. Because the current role of the university is to prepare students with the tools that the corporate world requires, diversity is now promoted universally as an important part of the college experience. Many universities are very clear about their need to help students by fostering "the type of cross-cultural experiences that will make them attractive to future employers.[36]

In his collection of essays *Dispatches from the Ebony Tower*, scholar Manning Marable explains the problematic state of ethnic studies within the realm of higher education as diversity is promoted for its moneymaking possibilities. Marable writes:

> In this period of globalization corporate capital requires a multicultural, multinational management and labor force. Racialized ethnic consumer markets in the U.S. represent hundreds of billions of dollars; black Americans along spend more than $350 billion annually. To better exploit these vast consumer markets, capital has developed "corporate multiculturalism," the manipulation of cultural diversity for private profit maximization.[37]

As Marable explains, the push for multiculturalism can most succinctly be understood as an exploitative mechanism for business to extract as much as they can from minority communities. The promotion of learning about and incorporating minorities in the corporate world and the university is therefore rooted in monetary goals, as opposed to the democratic and moral ideals that were evident during the student movements of the 1960's. Therefore, corporate multiculturalism works to both use minorities in the workforce and to educate White employees on how to better tap minority resources. Once again Feinberg clearly explains the thinking behind corporate America's strong support for affirmative action when he writes that in "an age when the cultural and ethnic character of American society is undergoing remarkable changes and when the marketplace is the world itself, surely it makes sense in many instances to diversify the work force."[38]

What is lost when diversity is presented in terms of its corporate possibilities is the desire to address our ongoing and deep-rooted racial problems. The general absence in mainstream politics of concerns for racial equality and racial justice shows the true colors of the diversity rationale that has largely been promoted, and the Court's opinions in the Michigan cases exemplify this problem. Only Justice Ruth Bader Ginsburg seems to acknowledge the problems of racial inequality as primarily important, as she writes that "conscious and unconscious race bias, even rank discrimination based on race, remain

---

[36] "A Framework to Foster Diversity at Penn State: 1998 – 2003," The Pennsylvania State University, retrieved on September 25, 2003 from <http//www.equity.psu.edu/Framework/understand.html>.

[37] Marable, Manning. "The Problematics of Ethnic Studies," *Dispatches from the Ebony Tower: Intellectuals Confront the African American Experience*. Ed. Manning Marable. New York: Columbia University Press, 2000. 261.

[38] Feinberg, 77.

alive in our land."[39] While Justice Ginsburg does not argue against the ideals of diversity, she does raise the crucial point that the importance of the decisions should not be wholly seen through the lens of diversity.

It is here where another fundamental flaw of the diversity rationale must be examined: the promotion of colorblindness as the ultimate goal for an American society that since its inception has been plagued by racial discrimination and inequality. With the twenty-five year limit for the continuation of race-conscious admissions policies that it set, the Court made it clear that it believes America is working towards a day when racial preferences will no longer be necessary. And though colorblindness and diversity appear to be different, they work together to hinder the cause of racial justice because both ideologies allow us to leave the problems of deeply rooted institutional racial discrimination unaddressed.

Within the affirmative action debate, the promotion of colorblindness is quite often a powerful tool for many conservatives. Based on the principle that all Americans should not be given any type of preferential treatment based on race, colorblind ideology appeals to the deep-rooted American tradition of equal opportunity (even though that tradition has never been universally applied). It is quite common to see conservative proponents of colorblindness appeal to Martin Luther King, Jr's "I have a dream" speech" and its call for judging all people based on the "content of their character." David Theo Goldberg examines the complexity of colorblindness in his book *The Racial State*, and ultimately he views the colorblind ideology, largely advocated by conservatives, as best serving the interests of white supremacy. While right wing critics such as Curt Levey promote colorblindness as the way to achieve "a fully integrated society in which race does not play a role in a discriminatory way,"[40] Goldberg sees it as a way of maintaining the supremacy of a white culture that is generally thought of as natural and superior to that of any colored person.

First, colorblindness is defined by whiteness in American culture, so just as white Americans are simply known as Americans, as opposed to black Americans being known as African Americans, the state of racelessness that is the eventual goal of colorblind ideology is nothing more than a state of whiteness. As Goldberg puts it, "the racelessness of absorption and transmogrification of the racially differented into a state of values and rationality defined by white standards and norms, ways of knowing and being, thinking and doing."[41]

Second, colorblindness provides a structure for decontextualizing the present state of social inequality faced by brown skinned people throughout American society because it simply refuses to recognize the constant institutional injustices that have created the present state of social, political, and economic inequality. Considering that the net worth of White families is eight times that of Black families and twelve times that of Latino families,[42] that Black Americans owned just 1% of the nation's wealth by 1990,[43] and that in less than twenty years there will be as many Black men in America's jails as were

[39] *Gratz,* 539 U.S. 02-516 (2003).

[40] Levey, Curt. "Racial Preferences in Admissions: Myths, Harms, and Alternatives," *The Albany Law Review* 66.2 (2003): 490.

[41] Goldberg, David Theo. *The Racial State*. Oxford: Blackwell, 2002. 206

[42] Wolff, Edward N. cited in "A Scholar Who Concentrates... on Concentrations of Wealth," *Too Much*, Winter 1999. 8.

[43] Conley, Dalton. *Being Black, Living in Red*. Berkeley: University of California Press, 1999.

enslaved during the peak of slavery in 1860,[44] the social implications of color have to be examined. The very power in the colorblind argument is that it does not address the issues of the past as they impact the present, which is exactly why it can be promoted as a positive goal that we should all one day hope to achieve, but at the same time it absolutely does not attempt in any way to begin the process of racial justice.

As Goldberg succinctly puts it, colorblindness rests in "the failure of whiteness to recognize itself as a racial color, the implication must be that colorblindness concern itself exclusively with being blind to *people* of color."[45] And it is through this mechanism that colorblindness, coupled with appeals to diversity, can be publicly promoted as an end that will benefit all Americans, but they actually work as nothing more than ways out of dealing with a past and a present that have always, and painfully so, recognized how very different our colors are.

The Reversal of Brown and the Unequal State of American Public Education

It has been almost 50 years since Thurgood Marshall, who later became the first black American Supreme Court Justice as well as cultural and legal icon, argued successfully before the Court in *Brown.* In perhaps the most significant legal decision in the nation's history, a unified Court found that "separate-but-equal" segregation in public education unconstitutional, effectively overturning *Plessy v. Ferguson* (1896), the decision that had been the legal backbone for Jim Crown segregation. The Brown decision was based on the Court's findings that "separate educational facilities are inherently unequal,"[46] and therefore that segregation in public education violated the equal protection clause of the Fourteenth Amendment.[47]

The *Brown* decision of 1954 (*Brown I*) was followed by a 1955 decision in *Brown v. Board of Education* (*Brown II*) where the Court ruled on how to implement desegregation in America's public schools. It was in this aspect of the case where Chief Justice Warren stated that desegregation in America's public schools should happen with "all deliberate speed."[48] The Court, however, issued no specifics on how to implement the massive change, and resistance to the Brown decisions effectively nullified the decisions for more than ten years as most southern schools remained segregated.[49]

But by the mid 1960's, after many techniques for resisting integration had been nullified due to strong efforts through the burgeoning civil rights movement, an era of sizable and genuine change began that would last for over twenty years and would see drastic improvement in the effort to integrate America's schools. Practices such as creating inner-city magnet schools, city-suburban district interaction, and busing were used in order to desegregate schools, and these practices eventually became fairly successful.[50] This is not to say that a specific date or place can be seen as the starting point for similar integration efforts across the country, but rather that the first small steps of any desegregation efforts were taken during this time.

---

[44] Boyd, Graham. "Collateral Damage in the War on Drugs." *Villanova Law Review* 2002.

[45] Goldberg, 222-223.

[46] *Brown v. Board of Education*, 347 U.S. 483 (1954).

[47] *Brown*, 347 U.S. 483 (1954).

[48] *Brown v. Board of Education*, 349 U.S. 294 (1955).

[49] Chemerinsky, Erwin. "The Segregation and Resegregation of American Public Education: The Courts' Role." *North Carolina Law Review*, 81 (2003): 1603.

[50] Chemerinsky, 1603.

Though there was progress throughout the 1970's in school integration due to such decisions as *Swann v. Charlotte-Mecklenberg Board of Education* (1971) and *Keyes v. Denver School District No. 1* (1973), there were two decisions that laid the groundwork for the eventual return to segregation: *San Antonio Independent School District v. Rodriguez* (1973) and *Milliken v. Bradley* (1974). In *Rodriguez*, the Court ruled in favor of Texas' school funding schemes based on local property value, and by doing so the Court found that poor students are not a protected class under the Equal Protection Clause.[51] In *Milliken,* the Court overruled a federal district court's ruling that imposed a multi-district strategy to integrate Detroit's mostly Black student population and the almost completely White suburban school districts. The *Milliken* decision made it impossible to create inter-district remedies for segregation unless a constitutional violation in one district affected another district,[52] and considering the very segregated state of most school districts, *Milliken* banned a very powerful strategy for creating integration.

The 1980's saw the pinnacle of success for desegregation efforts, largely due to federally enforced desegregation orders that forced school districts to act. For example, in 1964, 2.3% of Black students in the South attended majority White schools, but by 1988 43.5% did; and by 1988 the average Black student attended a school with 36% White students.[53] But by the 1990's, the progress that had been made in desegregating schools was beginning to fade, largely due to the demographic makeup of Black and Latino inner-city school districts and White suburban districts. The Supreme Court had not heard desegregation cases in over two decades, but when the Court did hear cases in the early 1990's their decisions lessened desegregation orders that had been, up to that point, successful.[54] Basically, the Court ruled that desegregation was a temporary solution, and once a district reached a point where it had integrated enough to reach its federal mandate it no longer has to continue with desegregation efforts.

The decision in *Board of Education v. Dowell* (1991), *Freeman v. Pitts* (1992), and *Missouri v. Jenkins* (1995) have effectively combined to drastically reduce the amount of integration in public schools across the country. In *Dowell*, the Court ruled that once a district that has been under a desegregation order for a lengthy period of time and has taken tangible steps toward remedying its past discrimination, it should be released from its desegregation order.[55] In *Pitts*, the Court ruled that a district court in Georgia could give up control in certain aspects of a desegregation order that was being fulfilled even though other aspects of the order had not yet been achieved.[56] And in the *Jenkins* decision, the Court ended the desegregation order for Kansas City (Mo) public schools even though it had been extremely successful.[57] Once again, Orfield clearly explains the consequences of these decisions:

Under *Dowell*, *Pitts*, and *Jenkins*, school districts need not prove

---

[51] *San Antonio Independent School District v. Rodriguez*, 411 U.S. 1 (1973).

[52] *Milliken v. Bradley*, 418 U.S. 717 (1974).

[53] Frankenburg, Lee, and Orfield, 37.

[54] Orfiled, Gary. "Schools More Separate: Consequences of a Decade of Resegregation." The Civil Rights Project, Harvard University. July 2001. 16.

[55] Davis, Abraham and Graham, Barbara. *The Supreme Court, Race, and Civil Rights*. Sage Publications, 1995. 360.

[56] Davis and Graham, 361.

[57] Chemerinsky, 1617.

actual racial equality, nor a narrowing of academic gaps between the
races. Desegregation remedies can even be removed when achievement
gaps between the races have widened, or even if a district has never
fully implemented an effective desegregation plan. Formalistic
compliance for a time with some limited requirements was enough,
even if the roots of racial inequality were untouched.[58]

All in all, the decisions of the Court both in the 1970's and the 1990's have created a
situation where it is almost impossible to implement effective desegregation plans, and
many districts that had at one time made progress when under federal supervision have
now reverted back to segregation once out of federal control. The lower courts have
followed the lead of the Supreme Court, and desegregation cases across the country are
largely being decided in favor of "local control," aka segregation. And the results of
these decisions has been seen everywhere, in many different ways, but the overall reality
is that most students, white and of color, attend schools where racial and ethnic
segregation are the norm.

One must only look at the statistics to clearly see that while some progress had been
made in efforts to desegregate, the combination of the Court's decisions has now
effectively segregated America's children. For example, the urban school districts in
New York, Los Angeles, Houston, Philadelphia, and Chicago are all over 75%
minority;[59] and on average, White students attend schools where they make up 80% of
the student population and Black and Latino students makeup a combined 16% of the
student body, while most Black and Latino students attend schools that are almost 75%
minority.[60]

In 2000, the average Black and Latino student attended a school where 44% of all
students lived at or below the poverty line, while the average White student attended a
school with 19% poor students.[61] A recent assessment of 4th graders found that 73% of
White students can read at or above the basic level, compared to only 40% of Latino
students and 36% of Black students.[62] And in 1988, in the average SAT score for Whites
was 1036, 189 points higher than the average score of 847 for Blacks. By 2002, the
average score for Blacks had risen all the way to 857, while the White average was
1060.[63]

The reversal of *Brown* has essentially created a situation where most Black, Latino,
and Native American students do not have the same equality of educational opportunity
as White students.   As a 2003 study by the *Journal of Blacks in Higher Education*
explains, "Public schools in many neighborhoods with large black populations are under
funded, inadequately staffed, and ill equipped to provide the same quality of secondary
education as is the case in predominantly white suburban school districts."[64] It is no

---

[58] Orfield, *Dismantling Desegregation*, 4.
[59] Frankenburg, Lee, and Orfield, 54.
[60] Frankenburg, Lee, and Orfield, 27.
[61] Frankenburg, Lee, and Orfield, 35.
[62] Paige, Roderick. "The Back Page: No Child Left Behind," retrieved on July 24, 2003 from
<http://www.carnegie.org/reporter/02/backpage/index_low.html>.
[63] "The Expanding Racial Scoring Gap Between Black and White Test Takers," *The Journal of Blacks in
Higher Education*. Retrieved on July 30, 2003 from <http://www.jbhe.com/latest/37_b&w_sat.html>.
[64] "The Expanding . . ." *The Journal of Blacks in Higher Education*.

wonder why, under these types of conditions that on average dropout rates are much higher for minority students than for whites.

Once again, Justice Ginsburg seems to be the only member of the Court concerned with the current state of public schools, as she mentions the problem of unequal public education in her concurring opinion in *Grutter*, arguing that "however strong the public's desire for improved education systems may be, it remains current reality that many minority students encounter markedly inadequate and unequal educational opportunities."[65] And until K-12 education is improved, and minority students actually have an equal chance for admittance to elite institutions such as the University of Michigan Law School, affirmative action policies will be necessary, but more importantly they will be devoid of any powerful meaning. As CUNY professor Stanley Aronowitz eloquently explains, until we equalize educational equity, "the relative deprivation of resources and equipment. . . the erosion of well-appointed and safe school buildings in urban elementary and secondary schools. . . the lack of enriched cultural programs" forces inequality on poor, minority students.[66] "Unless national policy works to reverse failures at these levels," according to Aronowitz, "the demand to raise standards is tantamount to a policy of wholesale class and racial exclusions."[67]

Conclusion

The decisions handed down by the Supreme Court in the cases against the University of Michigan certainly are important, but the Court's ruling in favor of affirmative action policies in college admissions will not have a very big impact on minority students. Only about 150 out of 1,800 institutions of higher education are competitive enough to warrant the use of affirmative action admissions policies, and 74% of the student populations at these institutions come from families with incomes in the top 25%.[68] When you consider that 50% of Black children and 44% of Latino children under the age of 6 live in poverty,[69] it appears clear that the gap between poor minority students and their middle class white counterparts will not be lessened by these decisions.

Understood in this context, the Michigan decisions are little more than window dressing on the issue of racial justice and equality, covering up enough of the real problems concerning institutional racism so that we do not have to publicly address them. What we need instead, according to writer Salim Muwakkil, is "A more honest reckoning of our history" that "would reveal the difficulty of transcending racial disadvantage without some attempt to repair the damage done to a people victimized by 16 generations of racial slavery and Jim Crow apartheid."[70]

In general, we need to realize that bickering about affirmative action enables us to look past the deeper problems of institutional racism that we must address in order to

---

[65] *Gratz,* 539 U.S. 02-516 (2003).
[66] Aronowitz, Stanley. *The Knowledge Factory: Dismantling the Corporate University and Creating True Higher Learning.* Boston: Beacon Press, 2000. 104.
[67] Aronowitz, 104-105.
[68] Hentoff, Nat. "Sandra Day O'Connor's Elitest Decision." *The Village Voice* 18 July 2003. <http://www. villagevoice.com/issues/0330/hentoff.php>.
[69] Moses, Michele. *Embracing Race: Why We Need race-Conscious Education Policy.* New York: Teachers College Press, Columbia University, 2002. 131.
[70] Muwakkil, Salim. "Affirmative Denial." *In These Times* 15 July 2003. <http://inthesetimes.com/comments.php?id=280_0_3_0_M>.

"find real and lasting solutions to racial disparities in educational opportunity."[71] We need to have an honest public debate in order to realize that, as Princeton professor and celebrated public intellectual Cornel West contends, "Race is not a moral mistake of individuals . . . It is a feature of institutions and structures that insures that one group of people have less access to resources, both material and intangible."[72] Until we look past our personal differences and realize that racial discrimination is most powerful within the institutions that dictate access and opportunity, the Michigan decisions, and others like it, won't mean much of anything to those it alleges to empower.

---

[71] Crenshaw, Kimberle Williams. "Beyond Affirmative Action: The Twenty-Five Year Détente."
TomPaine.com 10 July 2003. Retrieved on July 17, 2003 from <http://www.aapf.org/pages/detente.html>.
[72] West, Cornel. *Beyond Eurocentrism and Multiculturalsim*. Common Courage Press, 1993. 11.

Ethnicity, Politics and Social Conflict: The Quest for Peace in Liberia

Heneryatta Ballah, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Clemente K. Abrokwaa, Senior Lecturer of African Studies
The College of Liberal Arts
Pennsylvania State University

Abstract

This article investigates the fundamental causes of the Liberian Civil War, which has been going on since1989. It examines the social, economic, political, historical, and ethnic concerns that have contributed to the crisis. It argues that the current Liberian Civil War is the consequence of the socio-economic policies and political structures implemented by the freed slaves from America, who settled in the country in 1822. These policies discriminated against the indigenous Liberians, causing discontent among them, and resulting in their determination to win political and economic control of the country. The article suggests that lasting peace can only be achieved in Liberia, when these fundamental issues – including ethnicity – have been carefully studied and effectively addressed.

Introduction

Since 1989, Liberia has been engulfed in a bloody civil war, which has claimed more than 300,000 lives, including women and children. The war has led to increased social and economic decay, ethnic tensions, and political instability. More importantly, it has dislodged thousands of Liberians from their homes, turning them into destitute refugees both in Africa, Europe, and in North America. According to Dolo (1996), Liberian refugees are living in distressing conditions in neighboring Guinea, Cote d'Ivoire, Sierra Leone, Ghana and in other refugee camps where they are plagued by "high unemployment, increasing malnourishment and other conditions with terrible emotional consequences" (p.3). Attempts initiated by the Economic Community of West African States (ECOWAS), to resolve the conflict have all failed. And, despite the recent negotiations between the ECOWAS and the former President, Charles Taylor, which sent the latter into exile in Nigeria, the fact remains that the conflict continues to rage beneath the seeming calm surface of "temporary peace" in the country. Ethnic politics and the determination by the indigenous groups to prevent the government from falling into the hands of the Americo-Liberians, once again, could easily disrupt the temporary fragile peace in the country.

Ethnicity and Politics in Africa

Ethnicity constitutes the foundations of the African society, for it shapes communities, cultures, economies and the political structures of the peoples. More

importantly, it shapes the perceptions of the African, defines his universe, and provides him with meaning, understanding and the power to interpret the world around him. It is therefore an integral part of every African, despite the deepening influences of westernization and increasing cultural adulteration, since colonial times.

Ethnicity is important to Africans in many ways. First, it provides security both to the group, as a whole, as well as to the individuals constituting the group. The sense of belonging to an ethnic group means the members are safe together as one people, and ready to defend themselves against any external attacks on their existence and sovereignty. This notion of security also provides the groups with a sense of direction in their lives.

Second, ethnicity provides each group with a common ancestry and history, which is an important aspect of the African peoples: they desire to know who or what gave birth to their ancestors and where they are destined, following their departure from this earth. This knowledge of a common ancestry creates a strong bond within members of the group, for they realize that without the ability to support each other the entire group is doomed to die off or conquered by other groups around them. They are therefore ready to support their representatives in government at any cost- including a civil war.

Third, ethnicity also identifies each group by providing its members with a common language. Language defines a people giving them the power to think and reason logically based on their created world. Through language communication is possible among members, making it easier to share ideas and make any necessary changes required to benefit the people. It is through the distinct ethnic language that the knowledge, skills, values, taboos and other cultural beliefs and customs are passed onto the succeeding generations, in attempts to keep the group from dying off. Language also assists the groups in keeping their secrets from each other, for it is through language that they derive their power to rule and to exist.

Finally, ethnicity serves as an organizing force, which assists in bringing the people together to fight or seek a common goal (Okwudiba, 1998). This creates a sense of communalism, family, and togetherness, which also deepens the sense of belonging. Thus an entire community could belong to one major ethnic group, providing them with the opportunity to do things together as one family. In short, without ethnicity life is meaningless to the African.

It is important to note that members of the same ethnic group not only have a lot in common and share things that are unique to them, but they also live together in a specific region of the country involved. Thus it is common to find that the Ashantis of Ghana, the Yoruba of Nigeria, the Kpelle of Liberia are congregated in a specific region of those countries. They claim to own that region hence fight to keep "intruders" from taking over their land. Thus ethnicity has created the notion of regionalism, which is also counter to nation building in the modern world. Regional boundaries also dictate ethnic cultural practices, shape their belief systems and customs, and their perceptions of the world around them. For example, the Kpelle ethnic group of Liberia occupies the coastal area of the country and, therefore, holds sacred the ocean since it provides them with their sustenance. They also believe that the ocean is the abode of supernatural entities, such as witchcraft and other harmful beings. Based on this belief, it is therefore taboo for any member of the group to swim or bathe in the rivers and the ocean at night, as it

portends bad luck or may cause the death of a family member. They also believe that the person swimming or taking a bath may be captured and enslaved forever by one of the spirits.

Ethnicity, therefore, has been extremely important in the African's life and affairs, as it provided the people with a way of life, until the arrival of the European colonialists in the late 19th century. Colonialism came to alter the African society by undermining the ethnic groups in their ability to live together as one people - as it occurred in Liberia - with the arrival of the freed slaves from America.

Colonial Rule, Politics, and Ethnicity in Africa

At the Berlin Conference of 1884, Europeans divided the continent of Africa among them, and adopted various methods of ruling to govern their new colonies. These methods included the Direct, and Indirect rule. Inevitably, at independence, the legacies of these policies would become impediments to nation building across the continent.

The Direct Rule was mostly practiced by France, Germany, Portugal and Belgium. Under this political philosophy, colonies were ruled directly, meaning the colonizers brought in their respective colonial administration, including the police, clerks, governors, and other staff personnel directly from their respective countries in Europe and transplanted them in their colonies. Africans were excluded from participating in the government and were required to report to their new rulers, something they had never done before. This method of ruling also introduced the one-state government in Africa, which destroyed the African traditional system of ethnic autonomy of government and sovereignty. For centuries, Africans lived and worked in communities ruled by their respective ethnic leaders and councils of elders. Thus the one-state government forced Africans into abandoning their indigenous political systems, by adopting the new European political systems.

Indirect rule, on the other hand, was practiced by the British. This policy allowed them to rule their territories through the local ethnic leaders who, in turn, reported to the British officials. Judging from the surface, one could argue that this method of government was, perhaps, better than the Direct Rule, since it left Africa's traditional systems of government and cultures intact. However, this was not case, for it compelled African rulers to assist the European colonizers in exploiting their own peoples. The duties of the African chiefs were to collect taxes on behalf of the British, as well as implementing other policies, which greatly helped the colonizer to achieve his economic goals. Chiefs forcefully collected taxes, such as the hut tax, from their people and those who could not afford to pay were prosecuted. This led to divisions among the ethnic groups, for some of the leaders were seen as collaborators of the colonizers, and would later be removed from office.

Colonialism was also based on ethnic favoritism, which caused askew development strategies in the various colonies. The ethnic regions settled by the colonialists became the "preferred" regions, which received European development projects such as hospitals, schools, infrastructure, housing, road networks, and the creation of a modern sector economy, which relegated the rural economy to the fringes of the new African state. The ethnic groups in the areas settled by the colonizers, therefore, had access to public sector jobs, while their children were given the opportunity of acquiring European education. These political and economic policies created

fragmentation and uneven development within the African society, for many parts of the continent were left undeveloped. More importantly, it created animosity between the ethnic groups, since the majority of the ethnic groups had no access to education and jobs, and lived in poverty. The politicizing of ethnicity, by the Europeans, has been one of the fundamental causes of many of the civil wars in post- independence Africa. The arrival of the freed slaves from America illustrates this point in the Liberian case.

Arrival of the Freed Slaves in Liberia
Following the abolition of the Atlantic Slave Trade in the United States, in 1808, many former slave owners and politicians, particularly those of the North, began to fear the growing population of the freed slaves in the American society. As noted by Omonijo (1990), "They feared the possibility that they might lose their distinctive cultural identity, if the Negro remained longer in the society" (p. 10). Out of this fear, the American Colonization Society (ACS), a Christian organization, in 1816, began to advocate the resettlement of Negroes in Africa. Although a Christian organization, many of its members and executives were slave owners, including Robert Harper and Charles Fenton Mercer of the Virginia Legislature, who became active fundraisers for the resettling of the freed slaves in Africa (Dolo, 1998).

Some scholars, including Dolo (1998), argue that Southerners supported the deportation of freed blacks back to Africa as a method of preserving slavery in America, for they feared that the growing number of freed blacks in the North would influence the slaves of the South to rebel against their owners, by demanding their freedom. It was believed that by resettling the freed blacks in Africa, would limit the support for the liberation of enslaved blacks in the South. It is also argued that American missionaries, both from the North and the South, supported the resettlement program because they saw it as a means of spreading Christianity to the Africans whom they considered heathen and barbaric (Dolo, 1998). These hidden agenda, therefore, led one to question the original claim of Americans, concerned about the problem of losing their "distinct cultural identity", since their real motives were mainly economic and religious.

After raising enough money, and with the support of the American government, the American Colonization Society began looking for a place in Africa, which was going to become the new home for the freed blacks. Liberia, then known as the Malaquette Coast, in present-day West Africa, due to the abundance of the Malaquette Pepper in the region, was chosen as a suitable place for the resettlement of the freed slaves. In mid 1821, the first group of freed slaves from the United States, boarded the Ship, the Elizabeth for the journey and arrived on the Malaquette Coast, which they named Liberia, on January 1, 1822. The word Liberia, which comes from the Latin word, Libel means free. Between 1822 and 1867, more than 13,000 freed slaves were sent to Liberia (http://www.loc.gov). Although some discrepancy exists concerning the total number of freed slaves who settled in Liberia, Liebenow (1969), puts the total number at 18, 958. More importantly, the arrival of the freed slaves marked the beginning of a new era in Liberia's history, particularly for the ethnic groups that inhabited the region.

Foundations of the Conflict

Long before the arrival of the freed slaves from America, and who later became known as the Americo- Liberians, sixteen indigenous ethnic groups inhabited the Malaquette region of modern day West Africa. They included the Kpelle, Lorma, Kru, Gissi, Bozzi, Vai, Gola, Grebo, Mano, Bassa, Bandi, Sapo, Krahn, Geo, Mandi, and the Dei. These ethnic groups were governed by their ethnic leaders under their indigenous political systems and lived in communities based on their ethnicity. However, it was also possible for more than one ethnic group to inhabit one region. For example, the Lorma and the Bandi ethnic groups have occupied the northern part of the country (Lofa) for generations and long before the arrival of the freed slaves. Although two distinct groups, and with different languages, cultures, and traditions, they managed to live in peace and harmony, often referring to each other as cousins. This example could also be applied to the other ethnic groups, and despite the occasional ethnic conflicts, which occurred between the groups, everyone was considered equal. It is important to mention that politically, the region was not under any central government; hence each ethnic group governed itself based on its cultural traditions, religious, economic and social practices. However, the arrival of the Americo-Liberians would change the existing social structures, replacing them with western socio-political and economic institutions.

Upon their arrival, the Americo-Liberians segregated themselves from the indigenous Liberians and colonized them. They believed that they were more civilized than the indigenous Liberians, since they had experienced western civilization and acquired western cultural values, skills and attitudes. According to Nass (2000), the freed slaves that settled in the colony saw themselves as a distinctly enlightened group in comparison with the Africans they met on the land, whom they often referred to as "heathens and savages" (p.9). This attitude, which generated the notions of the "superior and the civilized" freed slaves, and the "inferior and backward" native Liberians soured the socio- political, and economic relations of the two groups from the beginning. Similarly, Omonijo (1990) notes:

> Right from the beginning therefore, the seeds of discontent were sown. The newly freed Negro slaves, rather than see themselves as Africans who were lucky to have been brought back to their roots, merely transferred the oppression they suffered in the United States of America to the Native population; they became the new lords of the settlement (p.11).

The Americo-Liberians, believing in the racist notions of the west at the time, which labeled Africans as inferior peoples, took full advantage of the indigenous Liberians, viewing them as "unfit" human beings to live with; hence they segregated themselves from them by establishing their own communities. This action laid the foundation for the policy of political and economic exclusion of native Liberians from the affairs of the country, a policy that would later result in civil conflict.

In terms of settlement, it should be pointed out that much of the land they settled on was forcefully taken away from the indigenous Liberians. Representatives from the settler communities often met with the local chiefs and demanded the sale of certain portions of their land for little or nothing, or should be given for free. Those indigenous Liberians who refused to give up the land that they had inherited for centuries faced severe consequences, including death. The dispossession of the native Liberians of their lands formed the prelude to the many injustices they suffered at the hands of their new colonizers. The Americo-Liberians later ensured that their new constitution gave them

the legal control over the affairs of the country, while denying even basic rights to the indigenous populations- the same rights, which had been denied them as slaves in America. The constitutional policies led to frustration, anger and animosity towards the settlers, thus laying the foundations for a future civil war.

The Americo-Liberian Constitution and the New Oppressive Rule

As the freed slaves began to settle in Liberia, the American Colonization Society took over the governing of the colony. In 1825, the ACS drafted a constitution that gave the Society the full responsibility for governing the colony and elected Elijah Johnson as its first agent (Lowenkopf, 1976; Omonijo, 1990). This action on the part of the Society contradicted their purpose for the resettlement of the freed slaves: to give the freed blacks the right and the freedom of self-government. The new constitution also affirmed the application of American law for the governing of the new territory. Interestingly, all the appointed administrators of the colony were white and many of them ruled with an iron fist. It was not until 1841, after the death of Thomas Buchanan, brother of the United States president that freed slaves became governors of the settlement (Lowenkopf, 1976).

The Americo-Liberians, who had earlier been denied all social, political, and economic freedom and other civil rights in the United States, became unhappy with the new political arrangements and wanted to ensure that their new freedom remained secure and unchallenged by any internal and external forces. Thus they pressured the ACS to give up control of the settlement and on July 26, 1847, they declared Liberia an independent nation and drafted a new constitution for the country. John Doe an Americo-Liberian drafted the new constitution, a replica of the American Constitution.

However, the new constitution excluded the indigenous groups from participating in the economic and political affairs of the country, giving all rights to the Americo-Liberians. The country's natural resources including gold, diamonds, rubber and timber now belonged to the Americo- Liberians and their families. The profits gained from these resources were used for the development of the Americo-Liberian communities, in the building of schools, churches, hospitals and other socio-economic facilities. Above all, the hinterland of the country, where a vast majority of the indigenous population lived was underrepresented in the National Legislature until 1964. In fact, prior to 1964, indigenous Liberians had to pay money to the government of the Americo-Liberians, if they wanted to observe the proceedings of the legislature (Alao et al, 1999). To the indigenous Liberians, therefore, the new constitution was illegal and so lacked legitimacy. Seyon (1995), points out that the Americo-Liberians "were forced to rely on coercion, not consent of the governed, to rule" (p.22).

Indigenous Liberians, both men and women, were also denied the right to vote under the new constitution and could not be citizens of the new republic. They were viewed by the Americo-Liberians as "unimportant" and faced constant discrimination.

According to Nass (2000):

> Citizenship was restricted to the settlers and their descendants. The Africans were required to pay taxes such as the obnoxious hut tax. It took quite a long time before the Africans had the right to send representatives to the government in Monrovia, initially only as non-voting observes. For several decades, only those Africans who were "civilized" with western cultural values were granted citizenship. The Aboriginal people were powerless. As second-class people, they were required by law and custom to adopt the western way of life before becoming full citizens (p.9).

Furthermore, in addition to acquiring western cultural values and practices, indigenous Liberians were also required to convert to Christianity and had to denounce their traditional religious beliefs and practices for three years, before they could become citizens. However, the fulfillment of these criteria did not guarantee them social equality with the Americo-Liberians; the racial segregation policies implemented by the settlers remained. For instance, converted indigenous Christian Liberians had to enter the home of an Americo- Liberian through the back door (Alao et al, 1999).

The exclusion from participating in the new government, the right to vote, and the denial of citizenship were not the only forms of discrimination suffered by the indigenous Liberians: they were also subjected to forced labor under inhuman conditions. In some cases, they were recruited to work on governmental projects as well as those of foreign companies with whom the government had business agreements. On September 16, 1925, the government of the Americo- Liberians signed a contract with the American Rubber Company, Firestone. Under the contract, the government supplied 50,000 laborers annually, to work on the Firestone plantation (Nass, 2000). To come up with the required amount of workers, the government created a Labor Bureau, which was responsible for recruitment. Often, the task of finding workers rested on the shoulders of indigenous chiefs and kings, and those leaders who refused to comply with the government's request were dealt with ruthlessly, by the new Liberian military - the Liberian Frontier Force.

Indigenous laborers worked long hours without pay and yet those who benefited from their hard work included the government. Nass (2000) remarks:

> Thus, the advantages of the contract labor generally benefited only a minute percentage of the population in government positions and their proxies. From the point of view of the contract laborers, the contract was only in favor of those who sent them and those who employed them, with total disregard for the laborers. The terms of these contracts made one wonder if they were not worse than the slave trade (p.12).

The environment, in which the workers lived and worked were unfit for animals let alone humans. Nass, (2000) describes the living conditions of workers on the Firestone rubber plantation:

> Conditions of the laborers in the farm could best be described as very appalling and at worst as inhuman and satanic. The workers were quartered in dingy shanties with little or no regard for the provision of most basic social amenities. The Corporation was only concerned with their labor output and had no commitment to their welfare and survival. The Corporation only dealt with the Liberian government. Their only concern was to ensure that any short fall in the stock of fit laborers resulting from sickness and death was replaced by the Liberian government. The contract laborer was never paid nor given medical care (p.13).

The forms of oppressive conditions under which, the indigenous Liberians worked were not that different from those experienced by the slaves in America. It appears as though, the indigenous Liberians paid for the transgressions committed by white Americans, who had previously enslaved the Black Americans. However, it was also obvious that the once oppressed slaves had indeed, become the new oppressors in Liberia.

Prior to 1925, many European countries had also taken advantage of the "abundance of labor" in Liberia and treated the indigenous laborers with the same

inhuman conditions. For example, in 1890, the French recruited Liberians to work on the Panama Canal and to serve in their colonial army. Again, in 1897, the Legislature of Liberia granted a German firm a labor recruiting concession. By 1925, Liberia had become the breeding ground for modern day slavery. The situation of forced labor- and even slavery- began to gain international attention, with constant reports coming into America about the working conditions of indigenous Liberians. In 1930, the League of Nations decided to investigate these reports. After the investigation, it was uncovered that the government had, indeed, been engaging in forced labor and slavery. Unfortunately, this discovery failed to end the practice. In fact, many indigenous Liberians were murdered by the government for complying with the League of Nations during their investigation.

Indigenous Resistance Against the Americo-Liberian Oppression

For over one hundred years the Americo-Liberian political party, the True Whig Party, ruled Liberia with an iron fist and kept the indigenous populations isolated from political participation, and socio- economic growth. The True Whig Party was the only legalized political party in the country and as such, faced no opposition. Nevertheless, it is important to note that the indigenous Liberians fiercely resisted their oppressors in the same way as the freed slaves had resisted slavery in America. As noted by Omonijo (1990), "History has been largely silent on the heroic resistance by the local population to the new system of overlordism. At various times, the indigenous Liberians fought their new rulers who had succeeded largely, in robbing them of their political rights" (p.12). Much of the indigenous resistance came from the Kru, the Gola, and the Grebo ethnic groups, which fought vigorously against the expansion of the Americo- Liberian territory in the country. As the Americo-Liberians began to expand their communities, many indigenous Liberians lost their homes and properties. To add insult to injury, the government passed the Hut Tax law, which required the indigenous people to pay tax on whatever properties they owned, a law, which only led to, increased resistance against the government.

In 1943, William V.S. Tubman, an Americo- Liberian, became president of the Republic and was considered a friend of the indigenous Liberians. While in power, he built schools, clinics and other facilities in the tribal hinterland of the country. In 1948, the right to vote was extended to indigenous Liberians. However, many of them could not vote due to the Constitutional Land Clause, which required that in order to vote an individual had to own a certain amount of land. Tubman also introduced the Open Door Policy, which allowed the free movement of trade between the tribal hinterland and the developed part of the country.

Although, Tubman made some positive contributions to the welfare of the indigenous Liberians, his government consisted almost entirely of Americo- Liberians, many of whom were his relatives. It is argued that the attempts made by Tubman to "equalize" Liberians were a political strategy for maintaining power. With regards to his open door policy, Liebenow (1987) argues that this policy was, at base, "a calculated strategy of economic development, which was designed not only to enhance the foundation of privilege for the Americo-Liberian elites but also to give them the revenue for maintaining a more modern and efficient system of control over the tribal majority"

(p.59). It is impossible to know the main objective of Tubman; however, what is certain is that he made some positive contributions to the well being of the indigenous Liberians: something that previous leaders had failed to accomplish. After the death of Tubman in 1971, vice president William .R. Tolbert became president. Tolbert's coming into office sparked a new wave of activism amongst students and others against the Americo-Liberian rule, which inevitably would lead to its overthrow in 1979.

The Tolbert era (1971-79), was marked by further exclusion of indigenous representation within the government and was stamped with nepotism, like his predecessors. For example, he made his brother, Stephen Tolbert, his Minister of Finance, while his daughter, Willie Mae Tolbert, became a Board member of the Bank of Liberia. Again, his son-in-law, Tonieh King, was appointed Commissioner for Immigration, and another son-in-law, Capt. Jehu Richardson, also became Captain of Air Liberia and Member of the Board (Liebenow, 1987). President Tolbert, in effect, ran the country as if it was his private property.

As the Liberian economy continued to decline, and with further isolation of the indigenous populations, various activist groups were formed to raise national and international awareness of the injustices many Liberians faced. In 1973, The Movement for Justice in Africa (MOJA) was founded by students and professors at the University of Liberia, whose leaders included Togba-Nah Tipoteh, an indigenous Liberian, and Dr. Amos Sawyer an Americo-Liberian, and Dean of the College of Liberal Arts and Humanities. The movement included Labor Organizations, Marketers, and other Student Movements, with the aim of raising public awareness about the ills of the Americo-Liberian government, while calling for change. Its program called for the nationalization of major economic enterprises, the confiscation of the illegal land holdings of the Whig Party, and the punishment of government corruption (Adebajo, 2002). The Organization also had branches in Algeria, Ghana, Nigeria, Gambia, Zambia, Mauritius, and Kenya to also raise public awareness of other injustices that were happening in Africa, such as the Apartheid system of South Africa (Dolo, 1996).

Another organization that was created to combat the settler domination in Liberia was the Progressive Alliance of Liberians (PAL). Liberian students who were studying in America, and was headed by Gabriel Bacchus Matthews created the organization. Their goal was to challenge the Americo- Liberian government, through the legal process, and to appeal to young Liberians (Dolo, 1996). To create awareness and recruit new members, several demonstrations against the government were held, with the most notable being the "rice riot" of 1979.

In early April 1979, the government announced that the price of a bag of rice, Liberia's staple food was going to be increased from $22.00 to $30.00 dollars. This huge increase in the cost of rice upset many Liberians, since many of them lived in poverty hence could not afford the new price. Moreover, the sudden increase was to benefit Tolbert himself and some members of his cabinet financially (Adebajo, 2002; Dolo, 1996). To voice their grievances over the government's decision, The Progressive Alliance of Liberians organized a demonstration. On April 14, 1979, two thousand Liberians took to the streets of Monrovia, the nation's capital, in protest. The group consisted of students from the University of Liberia, market women, social workers and ordinary citizens. As the protestors demonstrated, the police shot into the crowd, turning the march into a riot. At the end of the riot, more than 40 students lay dead, with 400

others wounded. The police arrested thirty-three of the demonstration organizers upon the orders of the President, who alleged that the demonstration was an attempt to overthrow his government. He charged them with treason, but later granted them general amnesty (Adebajo, 2002). This incidence set the stage for the final show down between the Americo-Liberians and the indigenous Liberians.

On April 12, 1980, a group calling itself The People's Revolutionary Council (PRC), attacked the president's mansion, and killed him and the key members of his cabinet. The group, led by Master Sergeant Samuel K. Doe, went on the air and announced that a coup had taken place and that President Tolbert had been killed. Prior to the coup not much was known about its participants, except that they were all soldiers in the Armed Forces of Liberia. Interestingly, all 17 members on the Council were indigenous Liberians, mostly from the Krahn ethnic group. This ethnic composition of the PRC later engendered a new conflict for Liberia: an ethnic/ civil war. However, April 12, 1980, ended 132 years of settler oppression and began the era of indigenous rule" (Dolo, 1996). After the coup, many indigenous Liberians rejoiced, believing that years of oppression had finally, ended, and now they would be able to enjoy the freedom and liberty they had been denied for so long, by the Americo-Liberians.

The Doe Regime: 1980-1990

The new government, headed by Master Sergeant Doe, and who had declared himself president, promised to return the government of Liberia over to its citizens. Again, he pledged that he would "curb the alarming rate of corruption in the country, and halt the insensitivity of government to the plight of the poor". He also vowed to build a new society in which there was justice, human dignity, equal opportunities and fair treatment for all before the law. Unfortunately, Doe did not live up to his promises. Instead he became the epitome of what he claimed to be against: like the Americo-Liberians, he, too, became an oppressor. He began his oppressive rule by retaliating against the Americo-Liberians. Dolo (1996) remarks:

> Hostility toward settlers was horrifying. Many members of the Tolbert cabinet were arrested, assaulted, jailed, tried, and 13 of them were executed. Settlers watched their properties get ransacked. And even more painfully, they watched powerlessly and in horror as their wives and children were treated with every level of discontent that the soldiers and their supporters could supply (p. 53).

Doe did not only murder members of the Americo-Liberian elite class: he also fired all of those members from the government whom he did not execute, replacing them with indigenous Liberians, many of whom had very little education. Interestingly, many of those he placed in high offices belonged to his ethnic group, the Krahn, thereby isolating the rest of the indigenous groups. Alao (1998) observes: "The new lease of life which many indigenous Liberians anticipated, did not materialize, as Doe, in whom they placed this expectation, was more interested in entrenching himself in power"(p.10). He took refuge in his Krahn constituency to seek support and ethnic solidarity: he wanted the Krahn to rule Liberia. This political strategy, based on ethnic loyalty, later caused divisions both in the government and within the Liberian society. However, his realization that the Krahn constituted merely 5 percent of the Liberian population, and the

reality that members of this ethnic group were not known to have occupied any positions of importance, caused him to embark on the rapid promotion of the Krahn people (Alao, 1998). On the other hand, it is also important to note that the Krahn ethnic group was not the only group isolated during the 132 years of the Americo-Liberian rule: all the indigenous groups were isolated from the economic and political sectors of the country during this period.

It seemed, at first, that the Doe administration would be free of corruption, and that the government was going to use its funds to improve the standard of living for Liberians, and invest in the development of human capital. However, Doe did not honor his promise of ending corruption; his government was infested with it. Within six weeks, following his inaugural speech, he collected a $600,000 "loan" from the International Trust Company of Liberia (Nass, 2000), to build a private house in his village; the "loan" was never repaid. Other officials in his government also got money to build private mansions in their hometowns and owned several properties in the city, all bought with the stolen money of the Liberian people. Adebajo (2002) writes: "Doe and his officials illegally acquired wealth and land as blatantly as the True Whigs once did. Revenue from logging concessions and fuel went straight to Doe's private funds; even U. S. food assistance was diverted into private pockets." By the end of his rule, Doe and his cronies had stolen a reported $300 million in public funds. Doe's dishonesty and corruption led to the further decay of the Liberian society and its economy.

Doe's promise of returning the Liberian government to its people was also never implemented. Instead, he became a dictator and banned all political parties from assembling, including The Progressive Alliance of Liberia, MOJA, and other student organizations, while strikes and demonstrations were also banned (Nnoli, 1998). Democracy was no longer a concern for him, thus from the beginning of his rule, to the end, his administration was accused of gross human rights violations. Those who did not support his ideologies were often beaten, jailed or killed; and many of his victims were students and lecturers from the University of Liberia.

Although all forms of opposition to Doe's government were banned, student activists and others were relentless in their pursuit for democracy and justice, many of whom would pay the ultimate price. On August 22, 1984, the Armed Forces of Liberia, upon orders from Doe, raided the University of Liberia, after the student union had questioned Doe's government about a timetable for the return to democratic rule. As the soldiers raided the campus, hundreds of students and staff members were beaten, wounded, killed, and many female students were raped (Nass, 2000). After the incident, Doe dismissed the University Senate and administration, and replaced them with staff loyal to him.

Prior to the 1984 incident, five students from the University of Liberia were sentenced to death after being tried by a military tribunal for questioning a policy that had abolished competitive student politics on the campus. After the verdict, students and other citizens strongly expressed their disapproval with the decision and protested against it, and a day before the execution was to take place; Doe granted the students executive pardon (Nass, 2000). The press of Liberia also faced constant persecution from Doe. It is said that the press had suffered under Doe's administration more than any past government in the history of Liberia. Journalists were often beaten, jailed and some killed

for writing the truth, and in some instances, newspaper houses were closed down: freedom of speech did not exist in Doe's Liberia.

To prevent further opposition to his government and to remain in power, Doe turned against those who had helped put him in power. Alao (1998), writes, "As a means to this end, he soon began to eliminate his former associates in the coup plot, so that within three years, all 16 colleagues who plotted the coup with him had either been killed or fled to neighboring countries" (p.10). One of such colleagues was Thomas Quiwonkpa, who belonged to the Gio and the Mano ethnic groups. Before overthrowing the Americo- Liberian government, he was a senior officer in the military of the Tolbert government and was well educated. However, as Doe continued to squander the resources of the country and oppressed the citizens, Quiwonkpa grew unhappy with his administration, arguing that the latter had failed to turn power over to the people of Liberia. Consequently, following the presidential elections, of November 1985, in which Doe declared himself the winner, Quiwonkpa, citing among other reasons, the "blatant rigging" of the elections, staged a coup against the Doe government. The coup, however, was unsuccessful and Quiwonkpa was brutally murdered (Alao, 1998). Following the failure of the coup, Doe deemed all members of the Gio and the Mano ethnic groups as his enemies, due to the fact that Quiwonkpa belonged to these groups. Severe punishment was inflicted on these groups by the government and it is believed that countless numbers of women and children were brutally beaten and murdered. This incident contributed to the immediate reasons for the current Liberian civil war.

The Beginning of the Liberian Civil War

On December 24, 1989, a group of 15,000 rebels calling themselves the National Patriotic Front of Liberia (NPFL) invaded Liberia from the neighboring Cote d'Ivoire. The rebel group, headed by Charles Taylor, an Americo-Liberian descendant who had served as the Director-General, of the General Services Agency (GSA) when Doe took over power in 1980. In 1984, as Doe's repressive rule continued, Taylor fled the country and came to the United States. Following his departure, he was accused by the government of embezzling $1 million in public funds, thus Doe requested his arrest and extradition to Liberia to face charges. Taylor was arrested and imprisoned in Boston, while the Liberian government worked on his extradition. As the extradition time approached, in 1985, it was alleged that he had escaped from prison and his whereabouts remained unknown until his invasion of the country in 1989 (Aboagye, 1999). Over the years, many Liberians and some scholars have questioned how Taylor escaped from an American prison; it has since been believed that he might have been deliberately released from prison by the American government.

The stated objective of the NPFL was to remove Doe from office, arguing that Liberians had suffered too long under Doe's regime (Omonijo, 1990). After entering the country from the Ivory Coast, the rebels made their way to Nimba County, the home of Quiwonkpa, whom Doe had murdered, following his failed coup. They encountered no difficulties in recruiting members from the Gio and Mano ethnic groups to join their cause. The willingness of these two ethnic groups to join the rebels was the result of the atrocities they had suffered under the Doe regime. Thus they sought revenge on Doe and

his Krahn group. To help further trust and support for his cause, Charles Taylor, deceitfully, told the Gios and the Manos that his effort was a continuation of Quiwonkpa's failed coup in 1985. Rienner (2002), notes that "This manipulation of ethnic differences predictably led to the NPFL attacks on the Krahns and Mandingos in its advance to the Monrovian capital, and many Krahn civilians were killed in the earlier stages of the war" (p.42). By 1990, thousands of Liberians had been killed, because of their ethnicity.

The atrocities that many Liberians faced during the early stages of the war were not only experienced at the hands of the rebels; the Armed Forces of Liberia, which consisted of mostly Krahns, were just as guilty as the rebels, since they, too, upon the orders of Doe, embarked on a massacre of the Gios and the Manos and those who opposed him. Innocent men, women and children were brutally murdered, and entire villages were set ablaze. Equally disturbing were the reports of hundreds of babies and children that were thrown into wells to drown after their parents had been killed (William, 2002). The senseless massacre of hundreds of innocent Liberians in Nimba county sparked off national and international criticism of the Doe government, and several groups appealed to the government to end the hostilities. One of those voices was the then U.S. Ambassador to Liberia, James Bishop. He called for an end to the hostilities and asked the government to declare Nimba a disaster zone for relief aid to be sent to the region, but the government declined his request (William, 2002). Doe continued to stand his ground and refused to resign and threatened that more lives would be destroyed, if the NPFL did not surrender to the government.

By April 1990, the structure of the NPFL was beginning fracture. Commander Price Yormie Johnson had split from the NPFL and formed the Independent National Patriotic Front of Liberia, after Taylor had executed some of his soldiers for their defeat by government forces in Ganta (Aboagye, 1999). The split between Taylor and Johnson now created a second war front, as both fought against each other, as well as the government. However, the rebel groups made their way to the capital, Monrovia, and the stronghold, which Doe had had on the country for years, began to weaken. The West African states, at this time, began to make constructive efforts to bring peace into the country.

ECOMOG and Peace Efforts in Liberia

In 1975, the Economic Community of West African States (ECOWAS), was created to promote regional economic integration and friendship among its member states. In 1978, the Committee Protocol on Non- Aggression, which was established in April 1976, was expanded when the protocol on Mutual Assistance on Defense was signed. Under the protocol all acts of hostility or aggression on a member state constituted a threat against the entire community (Aloa, 1998). Therefore, as innocent Liberians continued to lose their lives, leaders of the ECOWAS met in Banjul, the Gambia, from August 6-7, 1990, to discuss the Liberian crisis. Under the auspices of the ECOWAS Peace Mediation Committee, and with the approval of the members' heads of state, the ECOWAS concluded that something needed to be done quickly to resolve and to restore peace in Liberia. Nigeria's Head of State, Ibrahim Babangida, was at the fore front of the campaign. Consequently, the ECOWAS Cease-Fire Monitoring Group

(ECOMOG) was sent to Liberia in late 1990, consisting of soldiers sent from the various member states.

Although it was agreed that a peace keeping force should be sent to Liberia, some member states , including Burkina Faso and the Ivory Coast, opposed it arguing that such force would only prevent an imminent victory for Charles Taylor's NPFL, which cause they supported (Alao, 1998). Mali and Togo also refused to contribute troops to the peace force. In fact, Burkina Faso had provided some of its prisoners to join the NPFL and had hosted some of its training camps. Some members were also concerned about Nigeria's dominance in the peace process and saw it as a mechanism for further dominance in the region. Also opposing the ECOWAS intervention was Charles Taylor, who believed that Nigeria could not be trusted, especially since its leader, Babangida and Doe were close friends. Babangida was also criticized by his people, for many Nigerians saw the dispatching of ECOMOG, on the part of Babangida, as one of his preservation strategies (Alao, 1998). The Liberian population, on the other hand, embraced the idea and saw the ECOMOG as the answer to their prayers. President Doe and the INPFL leader, Prince Johnson, also supported it.

On August 24, 1990, three thousand five hundred ECOMOG soldiers arrived in Monrovia. Their specified mandate was to, maintain law and order, protect life and property, maintain essential services, provide security to the interim administration, observe elections and to conduct normal police duties in Liberia (Aboagye, 1999). As the naval ships made their way into the seaport, missiles fired by the NPFL, which opposed their intervention, hit them. On the other hand, the peacekeepers were welcomed by Prince Johnson's INPFL rebel group, which now controlled the seaport. There were also tremendous amounts of celebration from the Liberian civilian population. Upon their arrival, the ECOMOG took over the roles of a humanitarian group, as many starving civilians made their way to their camps to receive food.   The soldiers were very generous and shared their food, and the sick and wounded also received free medical treatment from the doctors, while refugees fleeing the country were evacuated to neighboring countries by the ECOMOG Navy, and the corpses that lay in the streets of the city were also removed and buried. It is important to note that the ECOMOG was the only symbol of order and help in Liberia at the time (Nass, 2000).

After a week of arriving in Monrovia, the peace keeping force successfully set up an Interim Government of National Unity (IGNU), which was headed by Dr. Amos Sawyer, Professor and Dean of the College of Liberal Arts at the University of Liberia. The Interim government was set up in accordance with the Banjul Resolution of 1990, which made it clear that whatever temporary government that was going to be set up in Liberia, should exclude all three warring factions. However, the war between Doe's AFL forces, the NPFL, and the INPFL continued.

The capture and the inhuman execution of Doe left many wondering how effective the ECOMOG was, as they were unable to prevent Prince Johnson from capturing him. Doe's death caused his Krahn constituents of the AFL and other supporters to battle Prince Johnson's forces, as well as those of Charles Taylor. However, Taylor saw Doe's death as an opportunity for him to claim presidency of the country.

The ECOWAS, realizing what had happened with the death of Doe, changed the ECOMOG mandate from peacekeepers to peace enforcers and the ECOMOG was given orders to enforce an immediate cease-fire on September 12, 1990 (Alao, 1998).

Unfortunately, those who had come to preserve life and to bring peace to Liberians now had to destroy some of these lives in the quest for peace and stability. More troops were brought in from Nigeria, Ghana, and Senegal to help in enforcing the peace. The ECOMOG wasted no time in enforcing the new peace mandate, and with more troops and weapons, they were able to successfully get all the warring factions to sign a cease-fire agreement by November 1990. However, throughout the 1990s, there was constant fighting between the warring factions, as the ECOWAS continued with its search for peace in Liberia.

In August 1996, the Heads of State of the ECOWAS member countries met in Abuja, Nigeria to review the Liberian situation. This meeting was a follow up to the 1995 Accord. The three main points highlighted at the meeting were: (1) when and how to reschedule elections to seek a legitimate government for Liberia; (2) when and how to revive disarmament and demobilization; and (3) whether and how to change the structure and membership of the Council of States (Nass, 2000). As part of the Accord, leaders of the three major warring factions were to oversee the disarmament of their fighters for free and fair elections to be held by 1997.

All the major warring factions agreed to the 1996 Abuja Accord and disarmament began in November 1996, as scheduled. Although, the pace of disarmament was slow, it proved to be successful and on July 23, 1997, elections were held in Liberia, and Charles Taylor won thus becoming the 22nd president of the country. However, his regime was never stable due to his oppressive policies and human rights abuses of those who opposed his government. In addition, he has been named as suspect in the trafficking of arms to help other dissidents of the region to destabilize their governments, and has also used the country's diamonds to buy weapons. Recently, the United Nations declared him a war criminal and has been seeking his arrest and trial.

Conclusion

As stated earlier, the Liberian civil war has claimed thousands of innocent Liberian lives, including women and children. It has not only created social, economic and political decay and ethnic tensions, but sadly, thousands of children have also been turned into soldiers, who have no understanding of the many grizzly crimes they have committed against their own people. The war has been the result of the economic and political exclusion policies implemented by the Americo-Liberians, since their arrival in the region in 1822. This political strategy has now turned the ethnic groups against each other, whilst the aspiring political leaders are capitalizing on the issue to their selfish advantage. In August 2003, Charles Taylor was forced into exile in Nigeria, to assist the ECOWAS to achieve temporary peace in the country. However, his exile does not rule out another flare up of the conflict, if the fundamental issues of economic and political exclusion of the ethnic groups are not addressed. It is, therefore, important for the peace efforts to consider that Liberia needs a democratic government, which gives equal representation and opportunities to all Liberians. This policy would ensure equal representation and freedom, and would remove unnecessary tensions among the various groups in the country.

It is also suggested that the United Nations enacts an international law that would stop those western nations involved in buying illegal diamonds from Liberian officials,

who use such profits to finance their selfish political goals and the bloody civil wars. It should also put an embargo on arms sales to Liberia, until the conflict has been resolved.

  Finally, African states and their regional organizations, particularly the ECOWAS, and the new African Union, should be more active in pursuing peace both in Liberia and across the African continent.

References

Alao, Abiodun, (1998). <u>The Burden of Collective Goodwill.</u> England: Ashgate

----------- (1999). <u>Peace Keepers, Politicians, and Warlords: the Liberian Peace Process</u>.
New York: United Nations University Press.

Adebajo, Adekeye. (2002). <u>Liberia's Civil War</u>. Boulder, CO: Lynne Rienner.

Brehun, Leovard. (1991). <u>Liberia: The War of Horror</u>. Accra, Ghana: Adwinsa

Dolo, Emmanuel. (1996). <u>Democracy versus Dictatorship</u>. New York: University Press
of America

Forsythe, P. David. (2000). <u>Human Rights and Comparative Foreign Policy</u>. New York:
United Nations University Press

Konah, Augustine. (1996). <u>Religion, Commerce, and the Integration of the Mandingo in
Liberia</u>. New York: University Press of America.

Liebenow, J.Gus. (1969). <u>Liberia : The Evolution of Privilege.</u> Ithaca: Cornell University
Press.

----------- (1987). <u>Liberia: The Quest for Democracy.</u> Bloomington: Indiana University
Press.

Nass, A. I. (2000). <u>A Study in Internal Conflicts</u>. Nigeria: Fourth Dimension

Okwudiba, N. (1998). <u>Ethnic Conflicts in Africa.</u> Dakar: CODESRIA

Omonijo, Mobolade. (1990). <u>Doe: The Liberian Tragedy</u>. Oregon: Sahel

Osaghae, Eghosa. (1994). <u>Ethnicity and its Management in Africa: the Democratization
Link</u>. Lagos, Nigeria: Mathouse Press.

Rupesinghe, Kumar. (1998). <u>Civil Wars, Civil Peace</u>. London: Pluto.

Seyon, L. N. Patrick. (1998). <u>Quick-Fixing the State in Africa: The Liberian Case</u>.

Boston: Boston University African Studies Center.

Timing of Fertility Transition in Kinshasa

Hae Chang Chung, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
David Shapiro, PhD.
Professor of Economics, Demography and Women's Studies
The College of Liberal Arts
Pennsylvania State University

INTRODUCTION

Purpose

The purpose of this study is to analyze the history of fertility transition in sub-Saharan Africa, specifically, in the city of Kinshasa (Democratic Republic of the Congo). Before fertility transition began, ethnic group differences in fertility were very large in Kinshasa; now, fertility differences by women's education have become more important. The study will explore the timing and the age pattern of the onset of transition by looking at age-specific fertility rates (ASFRs) over time.

Background

Demographic transition is something that every industrial country has experienced. It can be defined as the transition from high birth and death rates to low birth and death rates. Fertility transition is that part of the demographic transition which focuses on birth rates. Europe and North America faced fertility transition early, in the 19th and early 20th century. However, developing countries in Asia and Latin America did not experience fertility transition until the latter half of the 20th century. Many recent studies involving fertility transition use the Demographic and Health Surveys (DHS). DHS surveys, which have been carried out since 1985, are nationally representative household surveys with large sample sizes of between 5,000 and 30,000 households, typically. DHS surveys provide data for a wide range of monitoring and impact evaluation indicators in the areas of population, health, and nutrition. And the surveys are gathered from women ages 15-49. Part of the surveys covers reproductive behavior and intentions, contraception, antenatal, delivery, postpartum care, and children's health, among other topics (Demographic and Health Surveys, 2003). Because of these contents, DHS surveys help researchers to study fertility transition in developing countries.

According to Bongaarts (2002), who studied global fertility transition in developing countries after World War II, there are three general stages of fertility transition. In the first stage, fertility is high until the transition begins. In the second stage, once the transition gets underway fertility declines rapidly and tends to continue to decline. And in the last stage, the pace of decline decelerates as countries reach the later stages of the transition. Shapiro and Tambashe, looking specifically at sub-Saharan Africa, argue that there is a three-stage fertility transition by urban-rural changes. The

three-stage transition starts with fertility declining in urban areas while in rural areas staying at the same rate, then fertility declining in both urban and rural places but urban areas decline more rapidly, and at last with fertility declining more in rural than in urban areas (Shapiro and Tambashe 2002).

Significance of the Study

Sub-Saharan Africa is the last major region of the world to experience fertility transition, and DHS data shows that it is going through the early stages of fertility transition. Unlike other regions in the world, it has different characteristics in fertility transition that make it hard to predict what is going to happen in the future. However, some of the known factors that affect fertility rates are women's schooling, age at marriage, contraceptive use, and infant and child mortality (Shapiro and Tambashe, 2002). Another important aspect of this study is the direct relationship between the fertility rate and national population. So knowing future fertility rates will help sub-Saharan African countries to predict their populations in the future. Since the future is uncertain, this study will help sub-Saharan African countries to plan sufficient educational and health care systems.

Shapiro has already studied fertility decline in the city of Kinshasa (Shapiro, 1996). He analyzed data from a 1975 survey done in the city and data from a 1990 survey that he carried out, and he looked at fertility during the previous five years from both of the surveys and compared them. I will be looking at the fertility history data from 1990 and going farther back in time, so that I can look more closely at the changes and study the timing of fertility transition.

Research Questions

1. When did fertility transition begin in Kinshasa?
2. What is the age pattern of declines in fertility? That is, did older women initiate fertility transition only or by older and younger women?

Assumption

Sub-Saharan Africa is the last region to face fertility transition. Previous fertility transitions in Asia and Latin America had three stages of fertility transition beginning after WWII, with high fertility rates that decreased rapidly. The factor promoting fertility transition in Asia and Latin America was mostly economic development (Bongaarts, 2002). However, the rate at which fertility changes in sub-Saharan Africa will be different from any other region. And one of the most important factors that will influence the fertility rate will be the education of women.

Limitation

The 1990 data set has a sample size of only 2,450 childbearing-age women. Especially, the sample sizes of older women are quite small so data might not be

sufficient for the study. Also, recall error, where women are omitting or inaccurately dating their children's births, is a potential limitation on the study.

Review of literature

Fertility decline in the world

In the decades after World War II, developing countries in Asia and Latin America experienced fertility transition. As noted above, these fertility transitions had three stages. First, fertility is high until the transition begins, and once the transition gets underway fertility declines fairly rapidly and tends to continue. Finally the pace of decline decelerates as countries reach the late stage of the transition (Bongaarts, 2002, 5). The fertility transition in developing countries in Asia as well as fertility transitions in Europe began with older women, and only later did those transitions spread to younger women (Knodel, 1977, 231).

Fertility decline in sub-Saharan Africa

"Up until the latter part of the 1980s, sub-Saharan Africa was the only major world region in which fertility transition at the level of national populations had not yet begun" (Shapiro and Tambashe, 2002, 103). However, Garenne and Joseph argue that in general, fertility transition in urban areas began even before 1975, and fertility transition in rural areas began about 10 years later (Garenne and Joseph, 2002, 1841).

Also, Shapiro and Tambashe describe a "three-stage transition pattern in which overall fertility decline begins slowly and then accelerates. In the initial stage, with overall fertility quite high, rural fertility is more or less stable while urban fertility declines. Subsequently, both urban and rural fertility decline, with the decreases in the former being clearly larger than those in the latter. And finally, when a comparatively low overall level of fertility has been reached, the pace of continued fertility decline appears to be as great or greater in rural than in urban areas" (Shapiro and Tambashe 2002, 114).

"In general, urban women have more schooling, are considerably more likely to use modern contraception, are less likely to be in union, and experience distinctly lower infant and child mortality as compared to rural women" (Shapiro and Tambashe, 2002, 120). In most cases, schooling delays women from getting married or getting involved in a sexual union. And educated women use modern contraception to prevent unwanted pregnancy. Also, in urban areas the cost of living is much higher than in rural areas. And it is more likely that urban children get more education than rural children so it can be of a greater burden for urban families to have many children. However, rural families require hands to help out with their farming. So children can be more beneficial to rural families than to urban families. Therefore, urban areas have lower fertility than rural areas in sub-Saharan Africa.

Fertility Decline in Kinshasa

In mid-1955, Kinshasa had considerably higher fertility than the rest of the Congo. One of the reasons was that in the northern parts of country about 35 to 40 per cent of women over the age of 15 were childless. This was significant compared with a national average of 28 per cent and a figure for Kinshasa of 26 per cent. "Marriage occurred early in Kinshasa, with 62 per cent of women aged 15-19 being married and more than 90 per cent of those aged 20-24 living in sexual union. This reflected the low levels of women's schooling: fewer than ten percent of young women aged 15-19 was enrolled in school, with enrollment negligible at higher ages. Fully 64 per cent of those aged 15-19 had not received any schooling and the same is true of between 78 and 95 per cent of those in older cohorts" (Shapiro, 1996, 91). By 1975, the population had more than quintupled to a little over 1.6 million: an average annual rate of growth in excess of eight per cent. In Kinshasa, there was a slight increase in the crude birth rate compared to 1955 and a decline of the order of five per cent in the general fertility rate (Shapiro, 1996, 92). By 1990 there was evidence of further fertility decline among women aged 20-34, compared to those in the same age group in 1975. By contrast, numbers of children ever born in 1990 for women aged 35-49 were distinctly higher than in 1975.

"The single most important factor contributing to fertility decline in Kinshasa during this time period has been the tremendous increase in educational attainment of the adult female population" (Shapiro, 1996, 90). The majority of women of reproductive age in Kinshasa have now completed at least some secondary schooling. Increased schooling of women has contributed to increasing age at marriage and, for schooling at the secondary level and beyond, ultimately to reductions in fertility.

"Fertility differentials by schooling in Kinshasa were evident in 1975; however, numbers who had attended secondary school and higher education where these differentials are important were relatively small in the 1970s, and consequently had little impact on aggregate fertility behavior. By 1990, however, these groups comprised nearly two-thirds of the female population of reproductive age, and their behavior was an important component of overall fertility" (Shapiro, 1996, 90)

Methodology

This study will examine fertility decline in Kinshasa, capital of the Democratic Republic of the Congo. The study uses data from a 1990 survey carried out by Dr. Shapiro in Kinshasa. The survey covered a sample of 2,450 childbearing-age women. For each woman, the data include the age of the woman and the years when her children were born. With the data from the survey, we can calculate age-specific fertility rates (ASFRs) and the total fertility rate (TFR). From an earlier survey carried out in 1975, the TFR was estimated at 7.2, while for the 1990 survey it was estimated at less than 5.7. The TFR represents the average number of children women would bear in their lifetime, based on the current age-specific fertility rates. From this information, we can see the clear indication of fertility transition between 1975 and 1990. This study will use the data to

73

calculate ASFRs for various periods prior to the 1990 survey going back to the 1970s, and find out when and for what age groups fertility decline started in Kinshasa.

ASFR can be calculated by dividing the number of live births in each age group by the total female population in thousands in each age group. For example, let B_i be births to women age i during a particular year and let W_i be the number of women age i in that year. With this notation, the ASFR for women age i is calculated as a thousand times B_i divided by W_i. That is, the ASFR shows the probability of women having a baby in a single year. Multiplying that ASFR by five shows the expected number of children that women will bear while in that five-year age group. And adding up across all of the age groups tells the average number of children that women would bear during the course of their reproductive years if those ASFRs remained unchanged, i.e., the TFR.

In calculating TFRs, demographers often use a "window" or observation period of five years. It is conventional in demography to use 5-year age groups to cover the age range of childbearing women. That is, the youngest women are 15-19, 20-24 and the ASFR is calculated up through ages 45-49. In addition, TFR is defined as the sum of ASFRs from youngest to oldest multiplied by five and divided by 1000. Applying this equation, the TFR estimates are 7.2 for the 1975 data and 5.7 for the 1990 data, as noted above. These data show that there was fertility decline and we can assume that fertility transition occurred some time between 1975 and 1990.

I will be following demographic convention of looking at five-year age groups with five-year periods preceding the survey. The table below is an example of trends in age-specific fertility rates of Ghana in 1998. Ghana is a good example of a country that is experiencing fertility transition.

Age-specific fertility rates for Ghana, 1998

Age group	Number of years preceding the survey			
	0-4	5-9	10-14	15-19
15-19	90	104	117	123
20-24	192	213	242	265
25-29	206	247	257	255
30-34	183	216	241	[267]
35-39	143	162	[182]	-
40-44	79	[97]	-	-
45-49	[16]	-	-	-

Source: GSS and MI, 1999

We can see the declines in ASFRs for all age groups in the years preceding the survey, except for 10-14 years before the survey for women age 25-29.

I will analyze the Kinshasa data with five-year and two-year windows. This will detect evidence of fertility decline and indicate when fertility transition began in each age group. And I will be able to see if the fertility transition was initiated by older women, similar to what happened in the earlier European fertility transition and in the developing

74

countries of Asia in the 1960s and 1970s, or if the fertility transition was initiated by older and younger women, as Caldwell et al. (1992) suggested would be the case for sub-Saharan Africa.

Summary of Results

If we look at Figure 1, showing ASFRs for five-year periods, for women age 15-19 in the period from 1970 to 1974 the ASFR was 147, then 119.8 in 1975 to 1979, 94 in 1980 to 1984, and 96.9 in 1985 to 1989. From 1970-1974 to 1985-1989, the total decline in the ASFR is 50.1, i.e., about one third of the initial level. And for age group 20-24 there is a similar pattern, with the total decline in the ASFR being 111.9, again almost one third of the initial level. For the age group 25-29 declines are also present, with the total decline in the ASFR being 87.6, about one fourth of its initial level.

The situation was a bit different for the age group 30-34, with a hike between 1970-1974 and 1975-1979, but the ASFR declined after that. And for older age groups (not shown) the data sometimes show increases in ASFRs. We can explain this by small sample size. For 1985-1989 women in the age group 45-49 had a sample size of only 115. And 113 was the sample size for age group 40-44 in 1980-1984, age group 35-39 in 1975-1979, and 35-39 in 1970-1974. By comparison, the sample size was 2702 for age group 20-24 for the five-year period of 1984-1989. Since the sample size for older age groups are small, there is a greater likelihood the observations may not be accurate. The same issue is found in two-year period ASFRs.

Another explanation for lack of trend for age group 30-34 is the increase in educational attainment. As discussed in the review of literature, in most cases, schooling delays women from getting married or getting involved in a sexual union. As shown in Figure 3, not only is fertility falling but also the age pattern is shifting to later fertility. There is a big hike in the ASFR between age group 15-19 and age group 20-24 for both the early 1970s and the late 1980s. But in the early 1970s, there is a significant drop in the ASFR between age group 25-29 and age group 30-34. However, in the late 1980s, ASFRs stay close to each other from age group 20-24 to 30-34, which can be seen between age groups 20-24 and 25-29 in the early 1970s.

For Figure 2, the two-year periods, the transition isn't as smooth as for the five-year periods, but we can still see the declines in fertility for the younger age groups. For age group 15-19 there are two significant inclines from 1970-1971 to 1972-1973 and again from 1980-81 to 1982-83, but in general, the ASFR declined. In fact, between 1970-1971 and 1988-1989, the total decline of ASFR was 32.5. For age group 20-24, there is a clear indication of ASFR decline for most of the period. Age group 25-29 has more variability, but still shows a general pattern of decline.

Conclusion

We can conclude from these data that fertility transition was already under way in Kinshasa by 1975 (apparent from the early 1970s on). We can tell this from Table 1. Table 1 shows ASFRs for five-year periods preceding the 1990 survey. As discussed in

the results, for younger age groups we see pretty much continuous decline in ASFRs starting from 15-19 years preceding the survey, which is the early 1970s. Therefore, even though the data did not contain sufficient sample sizes for older women it is safe to say that younger women (15-29) clearly played an important role in Kinshasa's fertility transition.

Future studies

Because of limited time, I haven't had a chance to study the differences in the onset of fertility transition according to ethnic group and education. This is the area in which I plan to pursue this research.

Table 1
Age-specific fertility rates for city of Kinshasa, 1990

Age group	Number of years preceding the survey			
	0-4	5-9	10-14	15-19
15-19	96.9	94	119.8	147
20-24	240.8	263	298.6	352.7
25-29	274.6	275.8	315.2	360.2
30-34	241.1	264.4	276.1	[237.5]
35-39	194.4	168.1	[166]	-
40-44	79.8	[55.9]	-	-
45-49	[2.0]	-	-	-

Calculated from birth history data as of 1990 survey.

Figure 1 Age-specific fertility rates, five-year periods

ASFR 15-19

ASFR 20-24

ASFR 25-29

ASFR 30-34

78
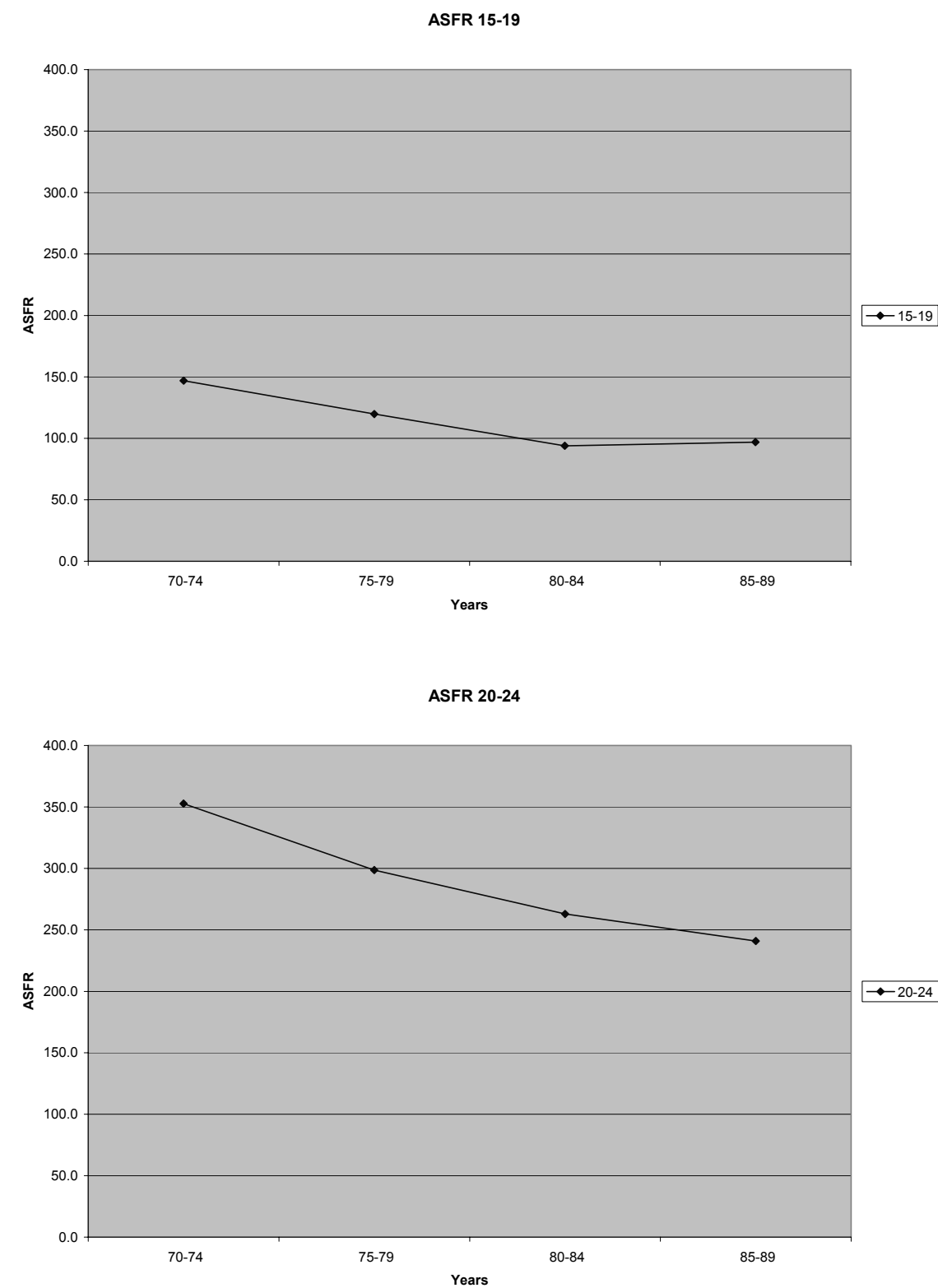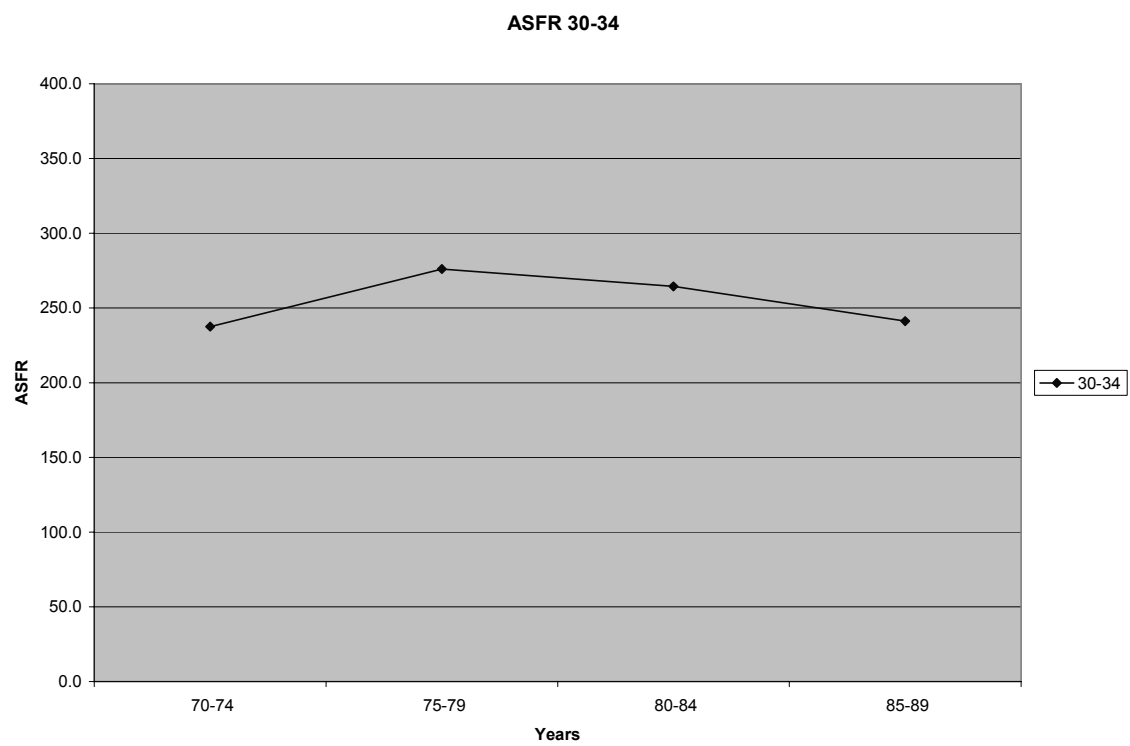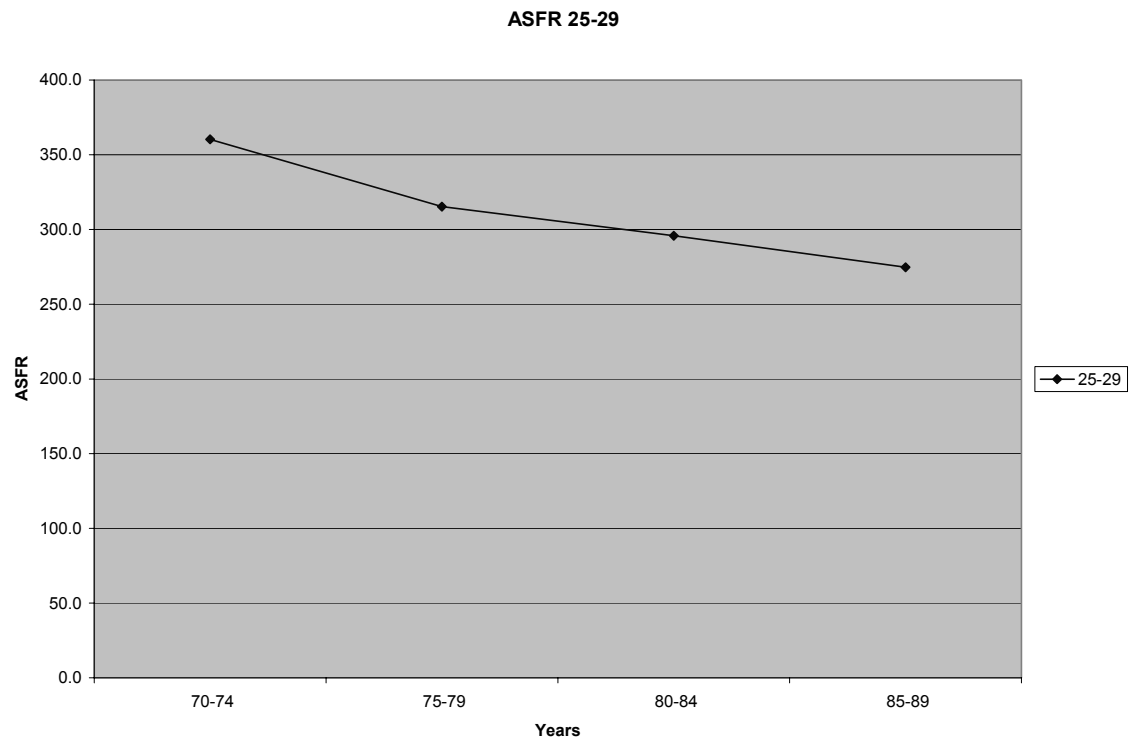
Figure 2 Age-specific fertility rates, two-year periods

ASFR 15-19

ASFR 20-24

ASFR 25-29

ASFR 30-34

80
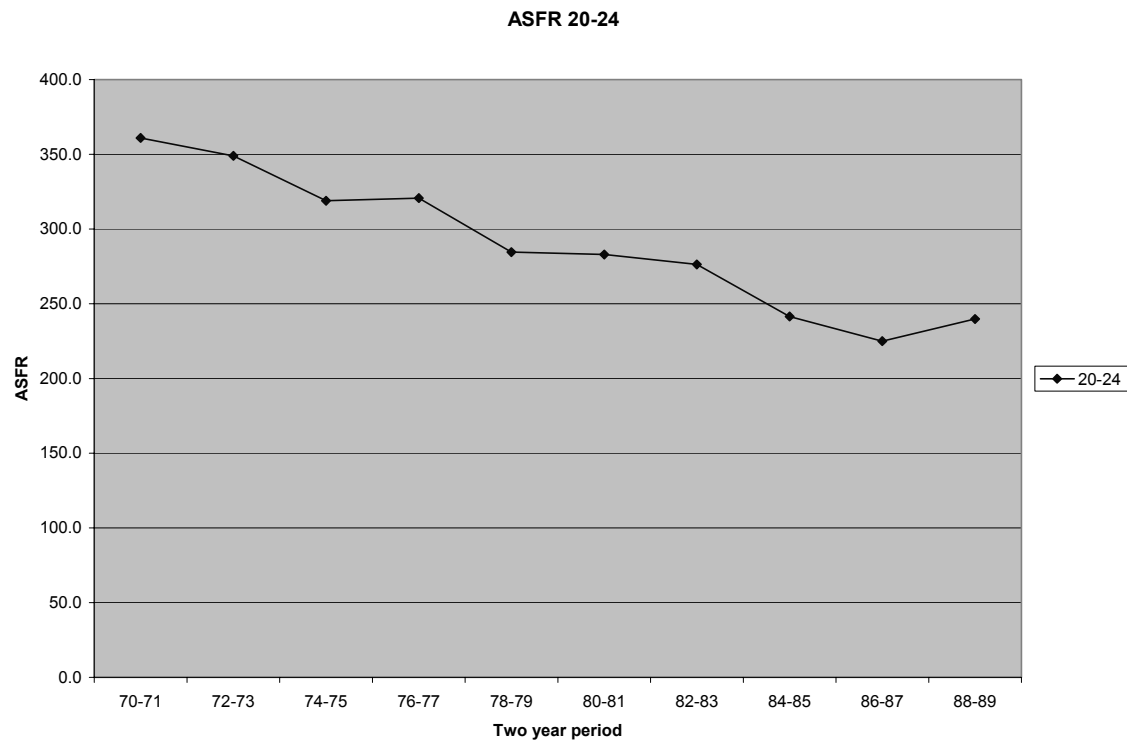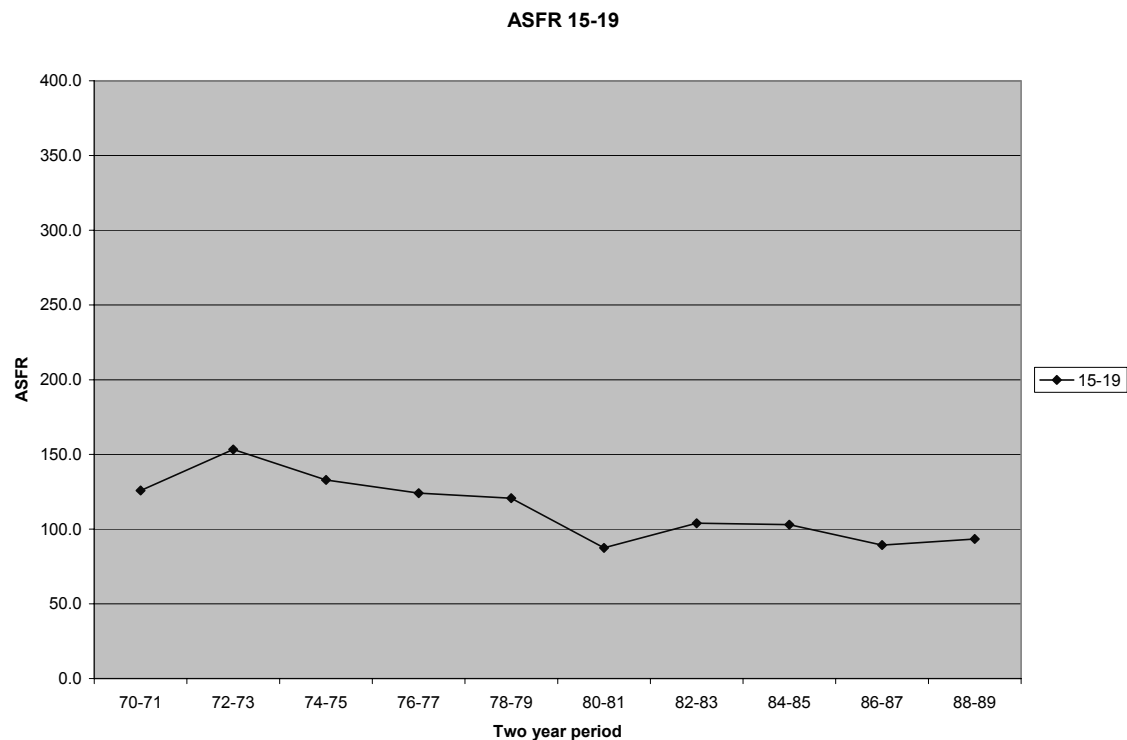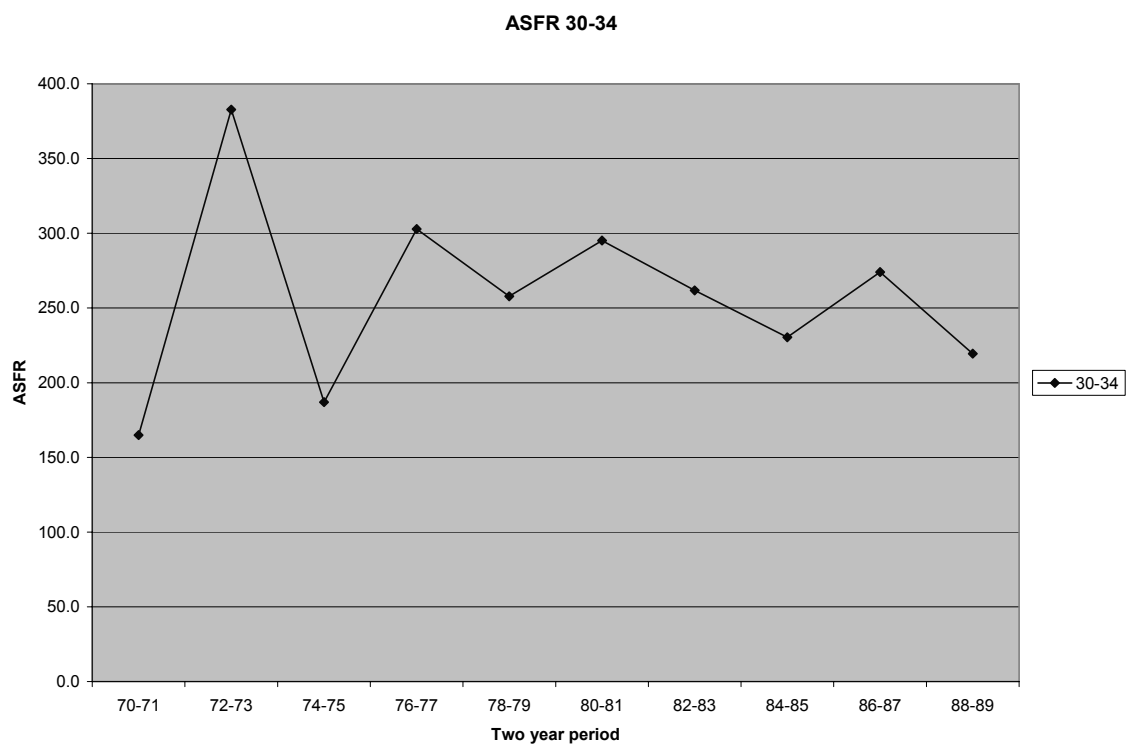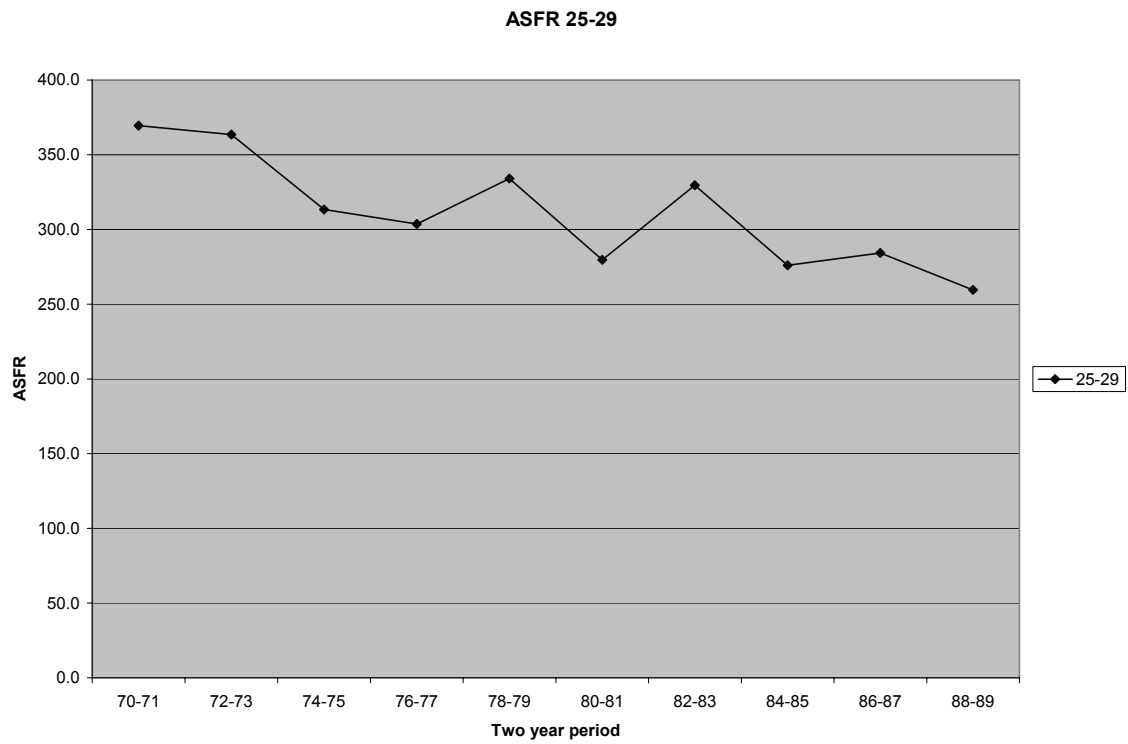
Figure 3 Fertility comparisons in early 1970s and late 1980s.

Fertility comparison
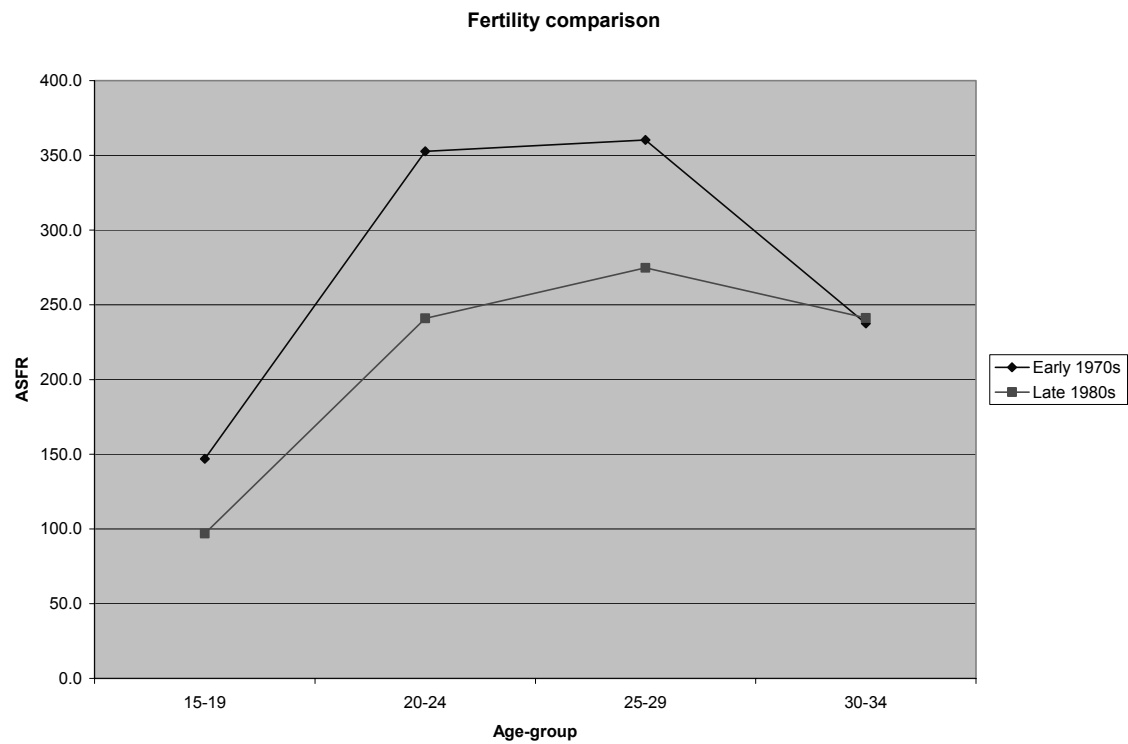
REFERENCES

Bongaarts, John. "The End of the Fertility Transition in the Developing World." Working Paper. New York: The Population Council, 2002.

Caldwell, John C., I.O. Orubuloye, and Pat Caldwell. 1992 "Fertility Decline in Africa: A New Type of Transition?" *Population and Development Review*, Vol. 18, pp. 211-242

Demographic and Health Survey. <http://www.measuredhs.com/about/types_of_surveys.cfm?CFID=634628&CFTOKEN=58582332> (5 June 2003)

Garenne, Michel and Veronique Joseph. "The Timing of the Fertility Transition in Sub-Saharan Africa." *World Development.* Vol. 30, (10), pp. 1835-1843, 2002.

Ghana Statistical Service (GSS) and Macro International Inc. (MI), 1999. Ghana Demographic and Health Survey, 1998. Calverton, Maryland.

Knodel, John. "Family Limitation and the Fertility Transition: Evidence from the Age Patterns of Fertility in Europe and Asia." *Population Studies.* Vol. 31, (2), pp. 219-249, 1977.

Shapiro, David, "Fertility Decline in Kinshasa." *Population Studies.* Vol. 50, (1), pp. 89-103, 1996.

Shapiro, David, and B. Oleko Tambashe, *Kinshasa in Transition: Women's Education, Employment, and Fertility.* Chicago: University of Chicago Press, 2003.

Shapiro, David and B. Oleko Tambashe. "Fertility Transition in Urban and Rural Sub-Saharan Africa: Preliminary Evidence of a Three-Stage Process." *The Journal of Africa Policy Studies*. Vol. 8, (2&3), pp. 104-127, 2002.

Shapiro, David, Basile O.Tambashe, and Tesfayi Gebreselassie. "The Fertility Transition

in Sub-Saharan Africa: Evidence from the Demographic and Health Surveys."

Working Paper. University Park, PA: Department of Economics, Pennsylvania

State University, 2003.

Validation of a Protocol for Motion Analysis

Michael M. Collins, McNair Scholar, Virginia State University

Faculty Research Adviser
Stephen Piazza, PhD, Assistant Professor
Departments of Kinesiology, Mechanical Engineering, Bioengineering,
Orthopedics and Rehabilitation
Pennsylvania State University

Research Assistant
Prashant N. Bansal
Pennsylvania State University

1. Introduction

Gait analysis is the systematic measurement, description, and assessment of those quantities thought to characterize human locomotion (3). Through gait analysis, kinematic, kinetic, electromyographic, and spatio-temporal data are acquired and analyzed to provide information that is ultimately interpreted by clinicians to form an assessment or used by researchers to develop new treatments and expand the knowledge base. Current examples of gait analysis include: (i) the assessment of cerebral palsy locomotion to assist in the surgical or orthotic intervention, (ii) the progressive examination of neuromuscular diseases such as Parkinson's or muscular dystrophy, (iii) the quantification of the effects of orthopaedic surgery through the comparison of pre- and post-operative patterns, (iv) and virtual reality, movies, and video games.

Quantitative gait analysis using three-dimensional motion analysis systems is becoming common practice in many research laboratories. Reliability is of the utmost importance, especially when clinical decisions are made, or in research such as product design and development. For the results of any motion analysis to be valid and widely accepted, thorough examination of reliability and error associated with the measurement procedure are required (6). The purpose of this research was to first standardize the gait analysis protocol for the Center for Locomotion Studies at the Pennsylvania State University. Subsequently, we investigated the intra-rater and inter-rater repeatability of kinematic data utilizing the VICON 370 (version 2.5) 3D motion analysis system. Having a reliable and valid protocol will be the groundwork for many subsequent studies that will be dependable and more widely accepted.

2. Literature Review

There have been several significant studies that have investigated the reliability of different motion analysis systems. However, the reported results have not been altogether consistent.

M.P. Kadaba et. al. (2) investigated the repeatability of gait variables including kinematic, kinetic, and electromyographic data waveforms and spatiotemporal parameters. Forty normal subjects were evaluated three times a day on three separate test days while walking at their preferred or normal walking speeds. Three-dimensional trajectories of body surface markers for computing joint angle motion were acquired using a computer-aided motion analysis system (VICON, Oxford Metrics Ltd., Oxford, England). While the subject walked in the positive X direction on a 6 m walkway, at least four of the five infrared cameras recorded the trajectories of markers on one side of the body. The opposite side of the body was subsequently recorded as the subject walked in the negative X direction

Retroreflective markers were applied to the shoulders (acromion process) and to the anterior superior iliac spines (ASIS). Key locations on the lower extremities include the lateral aspect of the greater trochanter, the knee joint line (posterior to the lateral femoral condyle), the lateral malleolus, and the dorsum of the foot between the second and third metatarsals. Attached to the pelvis was a posterior sacral wand (8 cm long) to measure the orientation the pelvic tilt. Two lateral wands (7 cm long) were attached to the thigh, midway between the hip and knee joints, and the shank, midway between the knee and ankle joints. The purpose of this was to hopefully measure the rotation angles more accurately. The hip joint center was estimated using regression equations with the leg length as the independent variable. The knee center was assumed to lie in a plane defined by the HJC, thigh wand marker, and knee marker, halfway between the femoral condyles. The ankle joint center was assumed to be in a plane defined by the knee joint center, shank wand marker, and ankle marker one-half the distance between the malleoli. Euler angle definitions were used to compute three-dimensional rotations of the pelvis, hip, knee, and ankle.

The intrarater repeatability was excellent for kinematic data in the sagittal plane both within a test day as well as between test days. In the frontal and transverse planes, joint angle motion yielded good repeatability within a test day but poor between test days. M. P. Kadaba et. al. attributed the poor between-day repeatability of joint angle motion in the frontal and transverse planes partly to variation in the alignment of markers. However M. P. Kadaba et. al. concluded that, in general, the results demonstrate that with the subjects walking at their normal speed, the gait variables are quiet repeatable. Thus, suggesting that it may be reasonable to base significant clinical decisions on the results of a single gait evaluation.

Another study assessed the reliability of gait measurements, and quite interestingly resulted in different findings from the reported reliable M. P. Kadaba study. V. Maynard et. al. (6) investigated the intra-rater and inter-rater reliability of kinematic data using the CODA mpx30 (Charnwood Dunamics, Barrow on Soar, Leicestershire, England) motion analysis system. Using very similar methods to M. P. Kadaba et. al. to define the anatomical co-ordinate system, amongst other

measurements, V. Maynard studied kinematic variables including hip, knee and ankle angles on initial contact, mid stance and mid swing.

Gait analysis was conducted on ten subjects twice each day, morning and afternoon, and once more a week later while keeping the investigators blind to the previous measurements. V. Maynard et. al. (6) used a standard protocol for marker placement and data collection with the hopes of reducing error. Their findings suggested a better inter-rater than intra-rater reliability for most of the gait parameters measured. Test-retest repeatability of measurements of joint kinematics was best for the knee angles and poorest for the hip angles. The findings reported do not demonstrate complete reproducibility of the gait analysis data when measurements were made with the CODA mpx30. The researchers attributed the poor reproducibility of kinematic data to the inaccurate placement of markers on the surface anatomical landmarks.

A similar study conducted by Cowman et al. (5) used the same CODA mpx30 motion analysis system, and obtained similar results. Their aims were to asses the degree of normal variation during the walking cycle and also the degree of error associated with marker placement as measured using the system above. They obtained two normal subjects (9 and 21 years) and assess them three times each by two chartered physiotherapists that were experienced in gait analysis. The subjects walked at an imposed speed, until four clean trials were collected. They observed high percentage error within the inter-rater data and it was believed that this may be attributed to the number of changing factors during an interpreter analysis. Furthermore, they interpreted this to mean that kinematic measurements at specific points in the cycle are less reliable that temporal and spatial data. A possible reason they offer for the poor reliability is the error associated with marker placement.

These studies have reported compromised kinematic data due partly to the misplacement of anatomical landmarks. A study headed by Ugo Della Croce et al. (4) investigated the reliability of the pelvis and lower limb anatomical landmark identification. The two healthy subjects investigated wore four skin marker cluster: on the pelvis, on the left thigh, shank and foot, each conveniently located in front of two cameras of a stereo-photogrammetric system (ELITE, B.T.S. Milan). To assess the intra- examiner reliability, the examiner (the Gait Laboratory physical therapist) was asked to identify the following sequence of: Left and Right, Anterior and Posterior, Superior Iliac Spines (LASIS, RASIS, LPSIS, LPSIS), Greater Trochanter (GT), Medial and Lateral Femoral Epicondyles (ME, LE), Tibial Tuberosity (TT), Head of the Fibula (HF), Medial and Lateral Malleoli (MM, LM), Calcaneous posterior surface (CA), dorsal aspects of First, Second and fifth Metatarsal head (FM, SM, VM). Additionally, the position of the Femur Head (FH) was assessed referring to the acetabulum center in the femur reference frame during the standing. According to the CAST protocol proposed by Cappozzo et al. (1995), a stick supporting two markers was used to point at each anatomical landmark, and a short static acquisition was performed. This anatomical landmark pointing procedure (calibration) was done six times consecutively. To assess the inter-examiner reliability, six registered physical therapists conducted the anatomical landmark calibration once on two subjects.

They reported that the anatomical landmark identification error is greater than the other sources of error. The inter-rater examiner test showed greater error than those

obtained in the intra-rater examiner. Among the body segment anatomical reference frames, the one of the foot is the most difficult to locate. The researchers say that this is likely to cause very low reliability in assessing joint offsets.

In hopes of increasing the accuracy and reliability of kinematic data, T. F. Besier et al (1) aimed their experiment at investigating numerical method used to define joint centers and axes of rotation independent of anatomical landmarks (Als). To do this, they compare the repeatability of gait data obtained from two models, one base on Als, and the other incorporating a functional method to define hip joint centers and a mean helical axes to define knee joint flexion/extension axes (FUN model). They also developed a foot calibration rig to define the foot segment independent of Als. The results indicated that the FUN model produce slightly more repeatable hip and knee joint kinematic data than the AL model, with the advantage of not having to accurately locate Als. This is especially repeatable for subject populations where location of Als is difficult. Repeatability of the models was similar comparing within-tester sessions to between-tester sessions. The foot calibration rig employed in both the AL and FUN model provided an easy alternative to define the foot segment and obtain repeatable data, again without having to accurately locating Als on the foot.

3. Methods

3.1 Gait analysis protocol

Eight able-bodied subjects (18-35y; mean 24y) free of gait altering injuries participated in this study. Gait analysis was performed on each subject twice a day on two separate days. The two raters performed separate gait analysis each day for the inter-rater examination. The same was repeated for the second day for the intra-rater examination.

A seven-camera VICON 370 motion analysis system (Oxford Metrics, Oxford, UK) was used in conjunction with two force-plates to collect motion data. A standing trial followed by a subject calibration trial was collected to locate the Als and the axes of rotation for the knee and ankle. During each marker application session, the subject walked at their preferred or natural speed. A minimum of five successful data collection trials were captured. To determine the hip-joint center, a functional method similar to that used by Piazza et al. was employed, whereby the subject was required move the right, followed by the left, thigh through four circumductions, 2 flexion/extensions, and 2 ad- abductions.

3.2 Marker set and definitions of segment and joint coordinate systems

To determine the three-dimensional position and orientation of each lower limb segment, cluster of four retro-reflective markers were firmly adhered to the subjects sacrum, thighs, shank, and feet. A technical coordinate system (TCS) was defined using each thigh, shank, and foot segment clusters such that the anatomical coordinate system (ACS) and joint centers were defined relative to these TCSs. Markers were placed on the following anatomical landmarks: the lateral and medial malleolus, lateral and medial femoral epicondyles, and the left and right ASIS and PSIS.

The foot segment was defined by the subject aligning the 2nd metatarsal heads and the heels of each foot on top of cardboard having 2nd metatarsal head and the heel markers on it.

3.4 Statistics

Intra-rater and inter-rater reliability was assessed with the interclass correlation coefficient (ICC) method. This approach is an appropriate statistical method for studying agreement between sets of interval data. An ICC coefficient of greater than 0.75 was accepted as evidence of good agreement. ICCs of less than 0.75 were considered less than convincing or not very reliable.

4. Results

4.1 Intra-rater reliability

The intra-rater portion of the study analyzed the variation between one rater's results on any given subject. The results of the intrasubject repeatability are given in table 1. The ICC's for the hip flexion/extension were the best. Rater 1's ICC's were slightly better than rater 2's with averages of 0.64 and 0.56 respectively. The highest ICC for the intrasubject was the ankle minimum for rater 2 at 0.70. Interestingly, the ankle minimum ICC for rater 1 was –0.42. The knee minimum and maximums had the lowest averaged ICC at 0.13. This is probably a result of the subjects not having their knees locked during the standing trial.

4.2 Inter-rater reliability

The inter-rater portion of the study analyzed the variation between raters measurements with only one subject. Results for the inter-rater repeatability are given in Table 1. The single highest ICC was the knee minimum on day 1 at 0.83. However, on the same day the maximum ICC was calculated to be 0.27. The lowest ICC calculated was for day 2's knee maximum at –0.29. Again, this is more than likely due to subjects not being reminded to lock their knees. The ankle ICCs were the highest followed by the hip and then the knee. In terms of averaged ICC's, Day 1 was more successful than Day 2.

Table 1:ICCs

Table of Interclass Correlation Coefficients				
Parameters	*Day 1* *Rtr 1 vs Rtr 2*	*Day 2* *Rtr1 vs Rtr 2*	*Rater 1* *Day 1 vs Day 2*	*Rater 2* *Day 1 vs Day 2*
Pk. Knee Flexion	0.27	-0.29	0.05	0.00
Pk. Knee Extension	0.83	0.19	0.36	0.13
Pk. Ankle Flexion	0.65	0.33	0.34	0.42
Pk. Ankle Extension	0.58	0.24	-0.42	0.70
Pk. Hip Flexion	0.70	0.27	0.69	0.58
Pk. Hip Extension	0.61	-0.04	0.59	0.54

5. Discussion

The purpose of this study was to standardize the motion analysis protocol and determine its reliability by analyzing the inter- and intrarater repeatability. We assumed that a standardized protocol for marker placement and data collection would likely minimize errors due to the rater. Interestingly, our findings suggest a better intrarater that interrater repeatability for kinematic measurement using the VICON 370 system.

Inconsistent with some previous studies, we found better repeatability for the hip and worst for the knee. As stated earlier, this is probably a side effect from the subjects not having their knees locked. In future protocols, subjects will be reminded repeatedly to lock their knees when collecting the subject calibration trials. This would also increase the repeatability of the hip angles. However, the ICCs calculated are fair given a few limitations. One being the time allotted, the study was conducted in a fairly short amount of time. We believe that if the study were extended, there would have been better preparation. Also, there would have been fewer subjects tested per day which is believed would reduce fatigue in raters and subjects. Which brings us to the next limitation, which is the natural variation in individual's gait patterns. Considering the time it took to complete a subjects testing, 3 to 4 hrs, and subjects were affected by fatigue that in turn introduces increased gait variation.

Overall, we consider this to be a fairly successful study, given the limitations, and accept the proposed protocol as reliable, pending minor changes.

References

1. Besier, T.F., Sturniekss D.L., Alderson, J.A., Lloyd, D.G. Repeatability of gait data using a functional hip joint centre and a mean helical axis. Journal of Biomechanics 36, 1159-68.

2. Kadaba, M. P., Ramakrishnan H.K., Wootten M.E., Gainey J.,Gorton G., Cochran G.V.B.. Repeatability of kinematic, kinetic, and electromyography data in normal adult gait. J Orthop Res 1989; 849-60

3. Davis, R.B. III, S. Ounpuu, D. Tyburski, J. R. Gage, 1991. A gait analysis technique data collection and reduction. Human Movement Science, 575-87.4

4. Della Croce, U., Cappozzo, A., Kerrigan, D.C., Lucchetti, L., 1997. Bone position and orientation errors: pelvis and lower limb anatomical landmark indentification reliability. Gait and Posture 5, 156-57.

5. Cowman, J.P., Jenkinson A., O'Connell P., O'Brien T. A model for establishing reliability and quantifying error associated with routine gait analysis (abstract). Gait Posture 191198;8:79.

6. Maynard, V., Bakheit, A.M.O., Oldham, J., Freeman, J., 2002. Intra-rater and inter-rater reliability of gait measurements with CODA mpx30 motion analysis system. Gait and Posture 17, 59-67.

A Study of a Low-Power Microwave Arcjet Thruster Using Ammonia Propellant

Jorge Delgado, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Michael M. Micci, Professor of Aerospace Engineering
Pennsylvania State University

Research Advisers
Dr. Sven G. Bilén, Assistant Professor of Electrical Engineering
Benjamin Welander, Masters Candidate
Pennsylvania State University

Abstract

A low-power microwave arjet thruster was studied using ammonia gas as propellant. The purpose of the project was to generate ammonia plasma using a 7.5 GHz magnetron and a cylindrical cavity thruster, which is resonant in the TM_{001} mode [1]. The desired pressure was atmospheric or above while generating a thrust of roughly 20mN and an I_{sp} of about 500s operating at low power (<100 kW).

Introduction/Overview of Electric Propulsion

Two main parameters used to describe a rocket engine are the thrust and the specific impulse (I_{sp}) generated by the rocket. Thrust is an exchange of momentum, momentum from the exhaust is transferred to the space vehicle which allows is to propel through space. Below is the derivation of the basic thrust equation where $\mathbf{P}$ = momentum, $\mathbf{m}_{dot}$ = mass flow, $\mathbf{v}_e$ = exhaust velocity, and $\mathbf{F}_{THRUST}$ = thrust force.

$$\mathbf{P} = \mathbf{mv}$$
$$d\mathbf{P} = d\mathbf{m}\ \mathbf{v}_e$$
$$d\mathbf{P}/d\mathbf{t} = d\mathbf{m}/d\mathbf{t}\ \ \mathbf{v}_e$$
$$\mathbf{F}_{THRUST} = \mathbf{m}_{dot}\ \mathbf{v}_e$$

Specific impulse is a performance parameter used for rocket engines, which compares the thrust to the amount of propellant used. The equation for the I_{sp} is shown below where $\mathbf{F}_{THRUST}$ = thrust force, $\mathbf{m}_{dot}$ = mass flow, and g_0 = gravitational constant.

$$I_{sp} = F_{thrust} / m_{dot} * g_0$$

Though most space vehicles currently use chemical propulsive devices, the substantial growth of Earth-orbiting satellites for communication and surveillance has sparked a new and intense interest in electric propulsive (EP) devices [2]. Chemical thrusters generate a greater thrust than EP devices but lose when it comes to I_{sp}. Below in Table 1 you can see a comparison of different propulsion systems and the respective specific impulse they produce.

Table 1 : Comparison of I_{sp} for different propulsion systems

Type	Isp (s)	Thrust duration
Chemical	200-465	minutes
Nuclear	750-1500	hours
Electrothermal	300-1500	years/months
Electromagnetic	1000-10000	years/months
Electrostatic	2000-100000+	months/years

The above table clearly shows that both propulsion and thrust duration of electrical propulsion systems is much better than those of chemical and nuclear thrusters. Aside from their higher specific impulse relative to other propulsive devices Electric Propulsion systems can also be very compact in size and may have low-power capabilities. This makes them suitable for micro satellites, deep space, and low drag missions. Electric propulsion devices are divided into three groups electrothermal, electromagnetic, and electrostatic. All three sub groups will be discussed below.

Electromagnetic Thrusters [2, 3, 4]

Electromagnetic thrusters use the electromagnetic force shown below to accelerate the propellant downstream.

$$F_m = j \times B$$

F_m = electromagnetic force per unit volume of gas (N/m^3)
j = electric current density passing through the gas (A/m^3)
B = magnetic field in gas (T)

In a simplified electromagnetic thruster current flows through a propellant gas from an anode to a cathode. A magnet provides a magnetic field perpendicular to the current and propellant flow (permanent magnet, electromagnet, or a solenoid may provide the

magnetic field). The resulting magnetic force accelerates the propellant down stream of the thruster.

A few examples of electromagnetic thrusters are the pulsed plasma thruster (PPT) and the magnetoplasmadynamic thruster (MPD). Most all PPTs use a solid propellant and achieve an I_{sp} between 1000-1500s. PPTs are compact in size and have high power efficiency, they operate in short impulses (~μs) thus making them very suitable for attitude control. MPDs produce a thrust that is proportional to the magnetic pressure inside the cavity and seem to be a promising thruster for the future as they are being heavily researched. Unfortunately no MPDs have achieved efficiencies higher than 35%.

Electrostatic Thrusters [2, 3]

The basic concept behind the electrostatic thruster is that electrical charges attract or repel each other. A source supplies charged particles of either sign into a cavity in which there is an electrostatic field and then they are passed out to a region in which the overall flow is neutralized. Common forms of the electrostatic thruster are the ion engine and the Hall thruster.

The ion engine electrons are produce by a cathode and sent into a chamber along with the propellant gas. The gas is then ionized and an optical grid is used to control the potential difference and accelerate the ions downstream. A Hall thruster is very similar to the ion engine. In a Hall thruster a propellant gas is ionized by counter flowing electrons. These ions are then accelerated by an electrostatic field generated by a negative cathode. The electrons are strongly magnetized and are forced to execute an azimuthal drift known as the Hall current.

Electrothermal Thruster [2, 3, 5]

Electrothermal devices are the most basic type of electric propulsion. Electrothermal thrusters use electrical energy to heat a working gas and then a conventional nozzle is used to accelerate the gas and produce a thrust. Such thrusters include resitojets and arcjets. The current project is a study of an arcjet thruster using ammonia propellant, which is brought to a plasma state by introducing microwaves and operating at low power. Arcjet thrusters can be made in a compact size and with advances in power subsystems they may operate at low-power making them very useful for micro satellites. Figure one below shows the microwave thruster firing. In this figure you can easily see the plasma inside the chamber and the resulting plume leaving the nozzle.
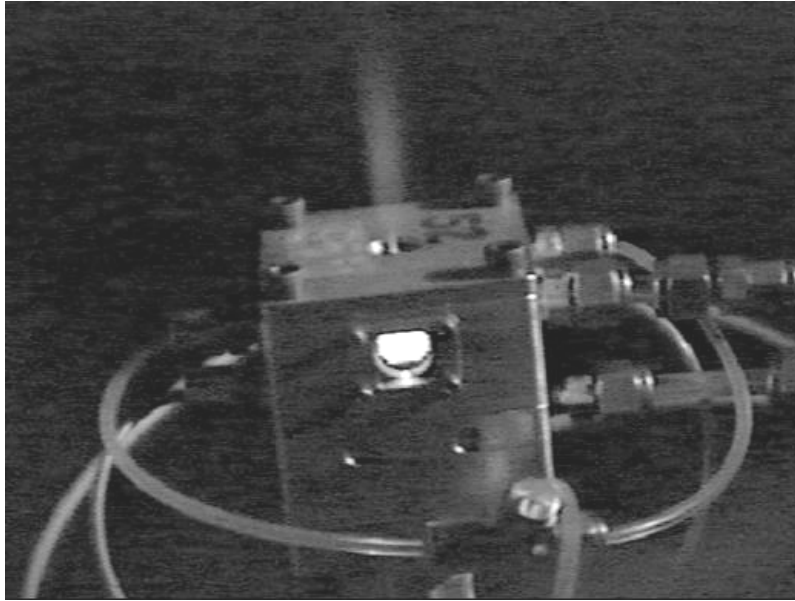
Figure 1 : 7.5 GHz thruster firing helium in vacuum conditions

For the missions in which the microwave thruster is intended for, we need to operate at very low power. Thus for the current experiment low power is defined at less than 1kW. Low-power thrusters were researched and compared to our theoretical assumptions of the microwave thruster performance. Table 1 shows all the thrusters that were researched and compares thrust, I_{sp}, input power, mass flow, efficiency, specific power, and specific thrust (the current study is listed last as NH3 microwave). We have estimated an I_{sp} of 500s and a thrust of 20mN. Figure 2 shows a graph of all the thrusters shown in Table 2, it is a plot of specific impulse versus efficiency. Here one can see that the microwave thruster has a high efficiency for the desired I_{sp}, only three other thruster have a higher efficiency but with lower I_{sp}. For this study and the future use of the microwave thruster an I_{sp} of 500s or greater is desired thus making this device the optimum thruster for the micro satellite missions it was designed for.

Table 2 : Comparison of low-power thrusters

	Thrust	Isp	Input Power	Mass Flow	Efficiency	Specific Power	Specific Thrust
	mN	sec	W	mg/s	P[I] / P[J]	J/mg	mN/W
Mark-IV Resistojet H2O	50.0	180	100	66.000	0.441	1.515	0.500
Pulsed Plasma Thruster	1.0	1500	100	0.068	0.074	1470.000	0.010
Helium Pulsed Arcjet	28.7	290	68	10.100	0.600	6.733	0.422
Teflon Pulsed Plasma	2.9	745	100	0.392	0.104	255.280	0.029
CIT 3-cm Ion Thruster	0.5	3703	24	0.014	0.377	1746.579	0.021
RUS 5-cm Ion Thruster	1.6	2900	72	0.056	0.316	1278.900	0.022
GRC 8-cm Ion Thruster	3.6	1760	100	0.150	0.310	666.667	0.036
GRC Colloid Thruster	0.20	390	0.5	0.051	0.700	10.706	0.366
Busek Hall Thruster	12.4	1346	207	0.940	0.395	220.213	0.060
Stanford Hall Thruster	11.0	544	277	2.063	0.106	134.249	0.040
Fakel Hall Thruster	4.7	1000	94.5	0.480	0.244	197.043	0.050
KeRC Hall Thruster	5.7	895	109	0.600	0.229	181.667	0.052
PPL Annul Hall Thruster	3.5	1086	98	0.400	0.190	245.000	0.036
PPL Cylind Hall Thruster	3.7	1136	103	0.400	0.200	257.500	0.036
NH3 Microwave (Est)	20.0	500	120	4.082	0.408	29.400	0.167
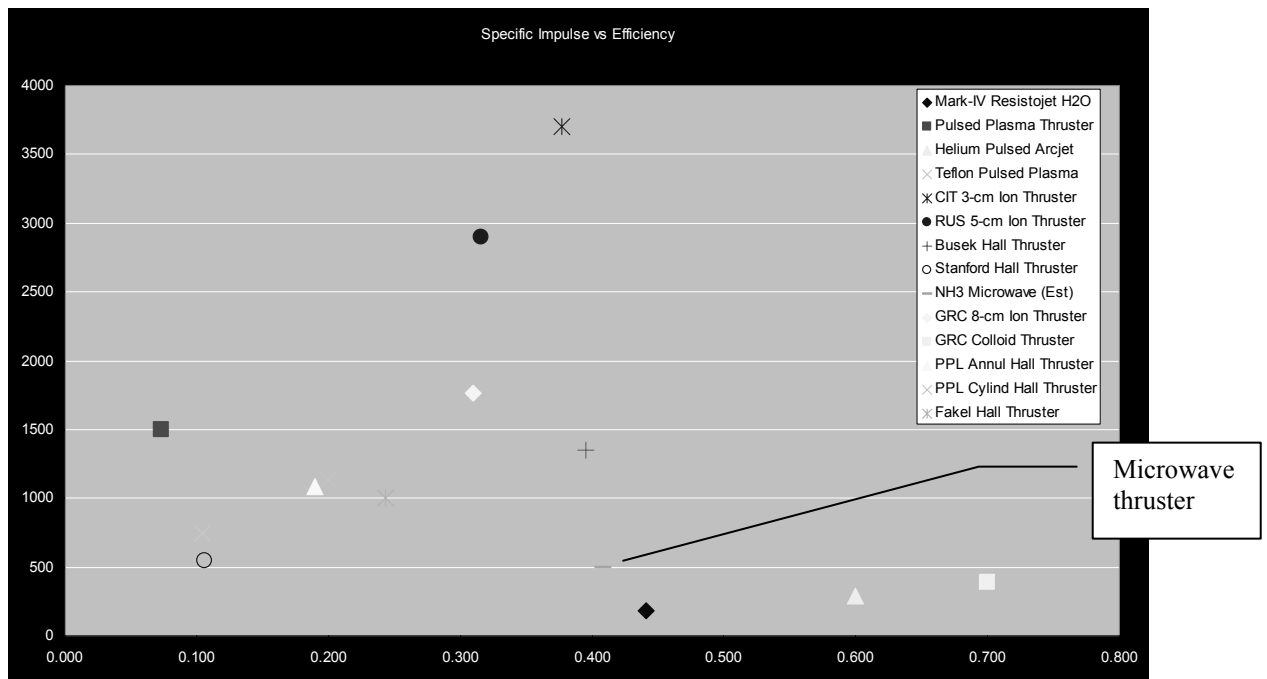
Figure 2 : plot of I$_{sp}$ vs. efficiency of thruster from table 1

Experiment

The thruster is made up of a cylindrical cavity shown in Figure 3. A 7.5 GHz tunable magnetron is used to introduce microwaves into the cylindrical cavity which is resonant in the TM_{001}. The chamber is divided by a quartz plate for its dielectric properties and the propellant is injected tangentially. The propellant is introduced into the plasma chamber tangentially for two reasons; one is to cool the chamber walls and help maintain the temperature at the proper range, this also helps to maintain the radial stability of the plasma [5].

Tangential injection Ports

Quartz Plate

Exhaust Plume
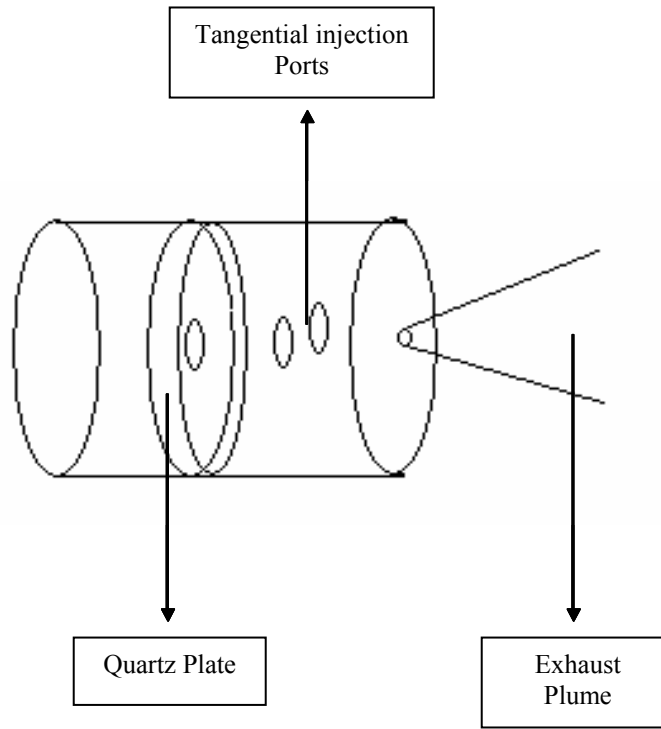
Figure 3 : Thruster Cavity

Fields inside a microwave cavity are governed by Maxwell's equations; a microwave cavity is resonant at many frequencies where there is a solution to Maxwell's Equations [1]. The frequencies at which a solution is met are the modes of the cavity resonator. The microwave cavity used in this experiment is resonant in the TM_{001} mode (first transverse magnetic mode). The geometry of a cavity with radius a, and length h is described by the following equation.

$$(fr)_{001}{}^{TM} = \tfrac{1}{2}\pi \left[(\mu\varepsilon) *\sqrt{(x_{001}/a)^2 + (\pi/h)^2}\right]^{1/2}$$

The experimental setup of the thruster on the thrust stand is shown in Figure 4. Here the thruster is hanging freely connected to the Narda dual-directional power coupler at the thruster cavity antenna which in turn is connected to the magnetron tube antenna. The power coupler is connected to two Hewlett-Packard 432A power meters by attenuator cables in order to measure both the incident and reflected power, which is being put into the system. The magnetron is a 7.5GHz 100 W tunable magnetron by Mictron and is powered by a Mictron power-conditioning unit. Figure 4 also shows the LVDT force transducer which will be used in future works. Pressure was read by a transducer which was affixed to the chamber cavity via a porthole.

Modeled after Counter Balance
- A = Magnetron
- B = Coupler
- C = Thruster
- D = Force Transducer
- E = Counter Weight
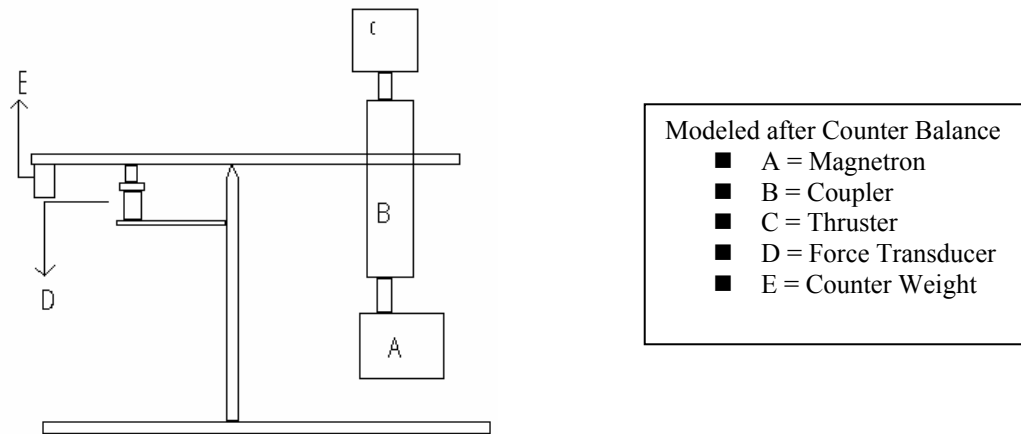
Figure 4 : Thrust stand setup

Testing and Results

For testing we first ignited helium gas in the thruster and then would close off the flow of helium and switch over to ammonia. Ammonia was able to be ignited but was unable to be maintained at desired pressures over the course of the project. Below in tables 4 and 5 are the results for both helium and ammonia gas. Shown in the tables is both forward (incident) and reflected power, anode voltage, anode current, filament voltage, filament current, control voltage, pressure, and mass flow (as a %).

Table 3 : Results for helium gas

arcet thruster	Helium							
Forward Pwr.	Reflected Pwr	VA	IA	VF	IF	VC	psi	mdot
0.52	0	4474	27	2.2		3515	1.03	4
0.5	0	4472	27	22		3573	1.27	5.5
0.5	0	4473	27	2.2		3574	0.98	4
0.5	0	4473	27	2.21		3572	1	4
0.66	0	4360	38	2.2		3442	15.9	70.1
0.66	0	4359	38	2.2		3442	16.05	70.1
0.62	0	4393		2.21		3467	15.96	70.1
0.68	0	4310	36	2.2	5	3392	14.98	70

Table 4 : Results for ammonia gas

Arcjet Thruster	NH_3							
Forward Pwr.	Reflected Pwr	VA	IA	VF	IF	VC	psi	mdot
0.48		4336	28	2.2		946	0.55	4
0.485		4484	29	2.2		948	0.57	4.1
0.46		4420		2.21		943	1.2	6.6
0.4		4410	32	2.21		950	1.76	8.4
0.4		4648	43	2.21		950	0.84	4.8
0.34		4628	38.5	2.2		962	1.02	4.8
0.385		4557	36	2.21		945	2.02	7.6
0.38		4556	39	2.19		951	2.03	7.7
0.38		4561	38	2.21		957	2.02	7.6
0.66	0	4386	32.7	2.2		3463	1.7	3.8
0.68	0.01	4373	34	2.2		3454	2.49	7.8
0.52	0.01	4453	28.4	2.2		3514	2.06	6.5
0.66	0.01	4330	36	2.2	5	3422	1.5	6
0.68	0	4346	35.5	2.21	5	3420	0.71	2.9
0.7	0	4330	36.7	2.2	5	3410	0.68	3.1
0.6	0	4392	32.6	2.2	5	3466	0.58	3.1
0.67	0	4322	37.5	2.2	5	3410	0.7	3.1
0.52	0.03	4418	38	2.2	5	3477	2.63	9.1

Atmospheric pressure is roughly about 14.30 psi, Table 3 shows that helium plasma was able to be maintained until atmospheric pressure conditions where met. Unfortunately ammonia plasma was not, with the highest pressure achieved being 2.63 psi. Ammonia plasma was also never cold started. Where as helium plasma could be ignited by itself regularly ammonia was not, helium plasma had to be ignited and then the propellant would be switched over to ammonia. Figure 5 shows the chamber in the horizontal position with the chamber filled with helium plasma.
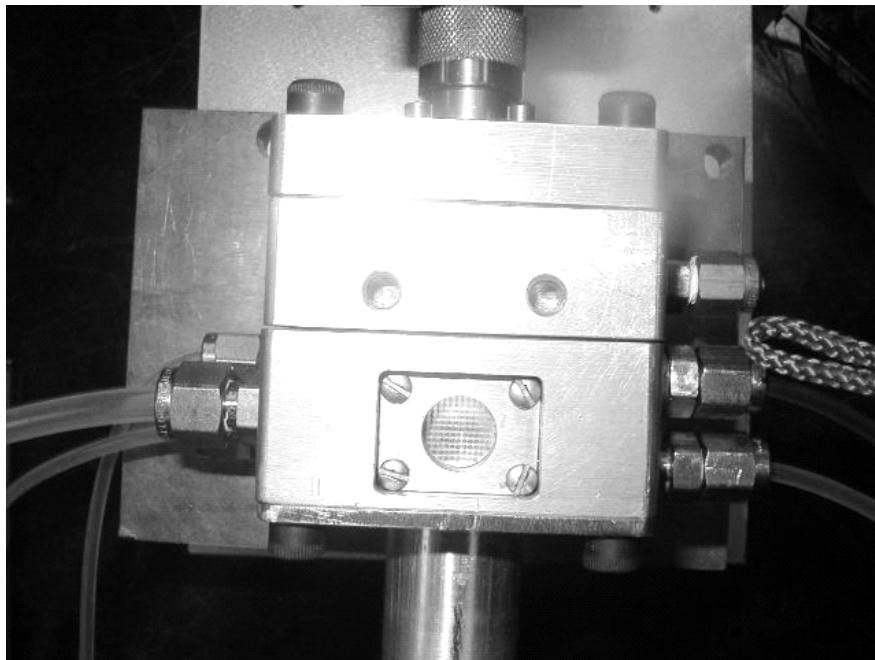
Figure 5: Thruster in horizontal position with helium plasma

Due to the time constraints and technicalities in the laboratory a run with ammonia plasma was never able to meet atmospheric pressure conditions. Also thrust measurements where not able to be taken and the project was not able to move into the vacuum chamber. Fortunately the project will continue through academic year. The working propellant has been switched to nitrogen, and a new in vacuum thrust stand is under works. The ultimate goal is to be able to ignite a helium-ammonia plasma mix and move into the vacuum chamber for further testing.

Acknowledgements

References

1.   *Roos, C. J. A., Vertical-Deflection Thrust Stand and Measurements of a Low-Power Microwave Arcjet Thruster, MS. Thesis, Aerospace Engineering Dept., Pennsylvania State University., University Park, PA, Aug., 2001.*

2.   *Humble, R. W., Henry, G. N., and Larson, W. J., Space Propulsion Analysis and Design, McGraw-Hill Co. Inc., New York, 1995*

3.   *Martinez-Sanchez, M., and Pollard, J. E., "Spacecraft Electrical Propulsion – An Overview," Journal of Propulsion and Power, Vol. 14, No. 5, 1998, pp. 688-698*

4.   Ziemer, J. K., and Choueiri, E. Y., "Scaling Laws for Electromagnetic Pulsed Plasama Thrusters," *Plasma Sources Science Technology*, Vol. 10, 2001, pp. 395-404

5.   Souliez, F. J., Chianese S. G., Dizac, G. H., and Micci, M. M., "Low-Power Microwave Arcjet Testing: Plasma and Plume Diagnostics and Performance Evaluation,"

6.   Funaki, I., Kuninaka, H., Toki, K., Shimizu, Y., Nishiyama, K., and Horiuchi, Y., "Verification Tests of Carbon-Carbon Composite Grids for Microwave Discharge Ion Thruster," *Journal of Propulsion and Power*, Vol. 18, No. 1, 2002, pp. 169-175

7.   Fujita, K., and Arakawa Y., "Performance Computation of a Low-Power Hydrogen Arcjet," *Journal of Propulsion and Power*, Vol. 15, No. 1, 1999, pp. 144-150

8.   Schmidt, D. P., Meezan, N. B., Hargus Jr, W. A., and Cappelli, M. A., "A Low-Power, Linear-Geometry Hall Plasma Source with an Open Electron-Drift," *Plasma Source Science Technology*, Vol. 9, No. 1, 2000, pp. 68-76

9.   Bushman, S. S., and Burton, R. L., "Heating and Plsama Properties in a Coaxial Gasdynamic Pulsed Plasma Thruster," *Journal of Propulsion and Power*, Vol. 17, No. 5, 2001, pp. 959-966

10.   Wilmes, G. F., and Burton, R. L., "Low- Power Helium Pulsed Arcjet," *Journal of Propulsion and Power*, Aerospace Engineering Dept., Pennsylvania State University., University Park, PA, Vol. 15, No. 5, 1999, pp. 440-445

11.   Sweetin, M. N., Lawrence T., and Leduc, J., "Low-cost manoeuvres for minisatellites using novel resistojet thrusters," *Proceedings Institution of Mechanical Engineers*, Vol. 213, part G, 1999, pp 223-231

12.   Spanjers, G. G., Birkan, M., and Lawrence, Major T. J., "The USAF Electric Propulsion Research Program," AIAA Paper 2000-3146, 2000

13.   Rayburn, C., Cambell, M., Hoskins, W. A., and Cassady, R. J., " Development Of A Micro Pulsed Plasma Thruster for the DAWGSTAR Nanasatellite," AIAA Paper 2000-3256, 2000

14.   Gorshkov, O. A., Grigo'yan, V. G., and Muravlev, V. A., "Development of the Low-Power Xenon Ion Thruster for Lightweight Satellites," *Journal of Propulsion and Power*, Vol. 16, No. 6, 2000, pp. 1083-1085

15.   Souliez, F. J., *Low-Power Microwave Arcjet Spectroscopic Diagnostics and Performance Evaluation*, MS. Thesis, Aerospace Engineering Dept., Pennsylvania State University., University Park, PA, Aug., 1999.

16.   Burton, R. L., and Turchi, P.J., "Pulsed Plasma Thruster," *Journal of Propulsion and Power*, Vol. 14, No. 5, 1998, pp. 716-733

Synthesis and Purification of Various Acyl-ACP Derivatives

Kathy Goodson, McNair Scholar, Virginia State University

Faculty Research Adviser
Dr. Squire J. Booker
Assistant Professor of Biochemistry and Molecular Biology
Pennsylvania State University

I. **Specific Aim**

Lipoyl synthase is the primary enzyme responsible for the synthesis of lipoic acid, a unique biological cofactor that is involved in cellular respiration. The substrate for lipoyl synthase is the acyl carrier protein (ACP) containing an octanoyl chain tethered to its phosphopantetheine prosthetic group. In order to study the lipoyl synthase reaction, efficient methods for preparing the substrate must be established. Octanoyl-ACP can be synthesized using acyl-ACP synthetases from *Vibrio harveyi* or *Escherichia coli*, or by a chemical method. The objective of this research is to establish an efficient means of preparing octanoyl-ACP using *Escherichia coli* acyl-ACP synthetase.

II. **Background and Significance**

Lipoic acid is an eight-carbon fatty acid with a unique 1,2-dithiolane ring that comprises carbons 6, 7 and 8 (numbering from the carboxylate), and can exist in the oxidized or reduced state (dihydrolipoic acid) (Figure 1). It is found covalently attached *via* an amide linkage with conserved lysine residues on enzymes in which it functions. Lipoic acid is a key cofactor in a number of multienzyme complexes responsible for the oxidative decarboxylation of α-keto acids, including the pyruvate dehydrogenase complex, the branched-chain oxo-acid dehydrogenase complex, the α-ketoglutarate dehydrogenase complex, and the glycine cleavage system.[1]

These complexes play an essential role in cellular energy metabolism; therefore, genetic defects in them are known to produce severe disease states, such as maple urine disease, which results from defects in the branched-chain oxo-acid dehydrogenase complex[8]. The inability to incorporate lipoic acid in any of these complexes would result in non-viable offspring. Lipoic acid is also known to be an important free radical scavenger in the cell. It is used to treat a number of diseases, including heavy metal poisoning, liver disease, and diabetes.[2]
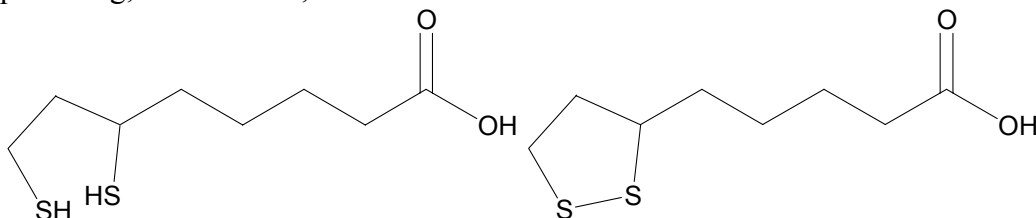
Figure 1: The oxidized and reduced forms of lipoic acid.

III. **Preliminary Studies**

There are several means by which octanoyl-ACP, the substrate for lipoyl synthase, can be generated. The enzyme, holo-ACP synthase, is used to append the phosphopantetheine prosthetic group on to apo-ACP, and acyl-ACP synthetases from either *E. coli* or *Vibrio harveyii* are used to attach an octanoyl group on to the phosphopantetheine group. The goals of the research described herein were to establish a method for isolating the *E. coli* acyl-ACP synthetase and for using it to synthesize octanoyl-ACP. The desire to develop methods to use the *E. coli* acyl-ACP synthetase was fueled by shortcomings associated with employing the enzyme from *Vibrio harveyii*, as well as shortcomings associated with the chemical procedure. A major drawback associated with the *Vibrio harveyi* enzyme is that the enzyme is not cloned and overexpressed, making its purification to homogeneity tenuous and tedious. In addition, the *Vibrio harveyi* enzyme copurifies with contaminating thioesterase activity, which degrades the product as it is being synthesized. The difficulty associated with the chemical synthesis is one of specification. It is a complicated task to place a group in a specific location on a large protein using chemical methods, because of the large number of side reactions that invariably take place.
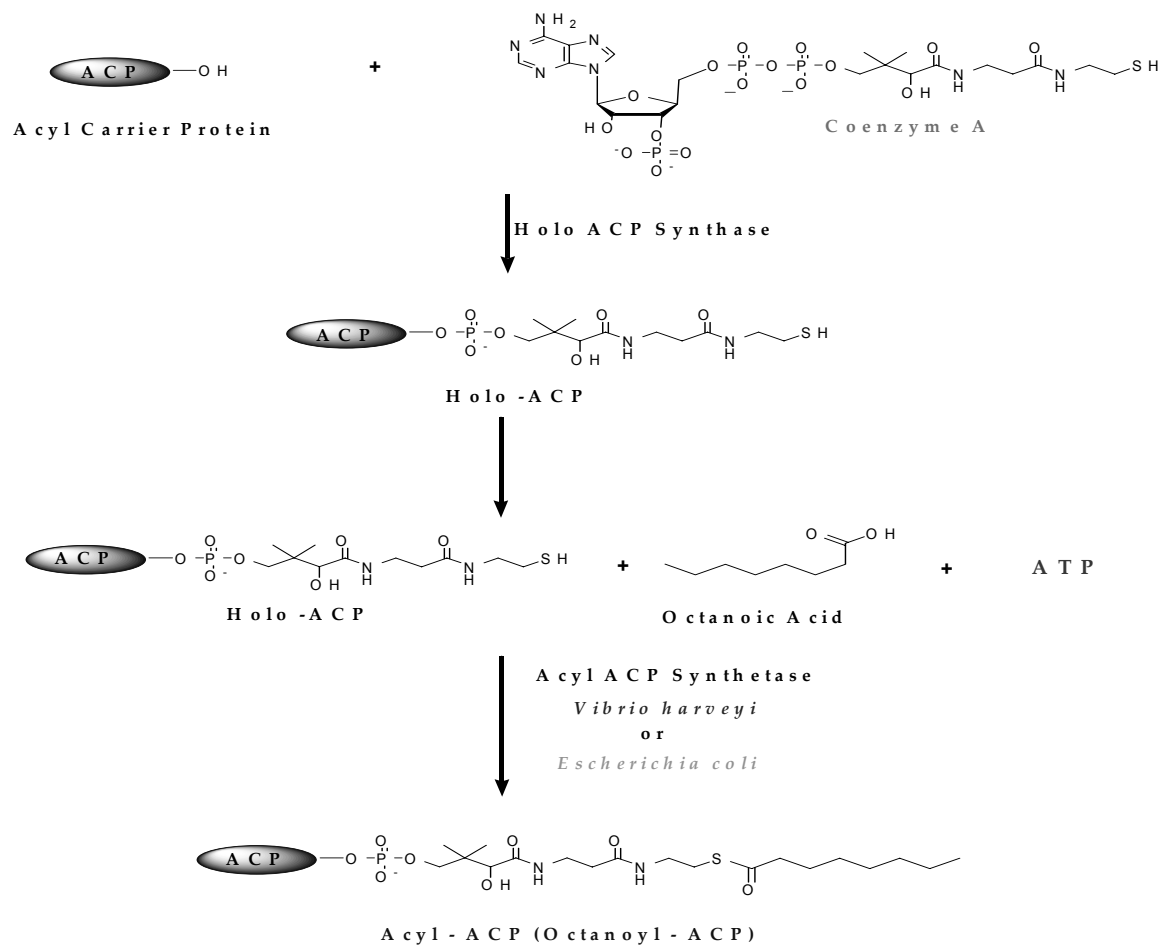
Figure 2: Octanoyl ACP synthesis

The purification and separation of various acyl-ACP derivatives may be achieved by more than one method of experimentation. Our particular research design was included both gel filtration and gel electrophoresis as means of separating reactants from products. The gel electrophoresis system of purification included the use of the Model 491 Prep Cell by Biorad (Hercules, MA). The Model 491 Prep Cell uses continuous-elution electrophoresis to purify proteins from heterogeneous mixtures.[3] The gel filtration method of purification employs size exclusion chromatography under partially denaturing conditions. This method of chromatography separates molecules based on their Stokes radii.

IV. **Materials and Methods**

The purification and handling of all proteins was carried out in a 4ºC cold room, unless otherwise stated.

Overexpression and Purification of Escherichia coli ACP

ACP was overexpressed and purified as described by Haas[4] with slight modifications. A single colony of *E. coli* BL21(DE3)pLysS / pBHF-5 was used to inoculate 200 mL of Luria-Bertani (LB) media containing 50 µg mL^{-1} kanamycin. The culture was incubated at 37 °C for 7-8 hours with shaking at 200 rpm. This culture was used to inoculate 16 L of LB media containing 50 µg mL^{-1} of kanamycin. The culture was grown at 37°C and shaken at 200 rpm until an optical density at 600 nm (OD_{600}) of ~0.6 was attained. Expression was induced by the addition of 500 µM isopropyl-β-D-thiogalactopyranoside (IPTG), 2g L^{-1} casamino acids, and 0.05 g L^{-1} pantothenic acid. Expression was carried out for three hours at 30°C, upon which cells were harvested at 10,000 x g for 6 min and the cell paste was frozen in liquid nitrogen and stored at -80 °C. Frozen cells were resuspended in 100 ml of 100 mM Tris, pH 8.0, containing 0.15 mg mL^{-1} of lysozyme, DNase, and RNase. The mixture was sonicated on ice in twelve 30-second bursts and centrifuged for one hour at 39,000 x g. After centrifugation the supernatant was diluted with one volume of 25 mM 4-morpholineethanesulfonic acid monohydrate (MES), pH 6.1 (buffer A). The mixture was then loaded onto a diethylaminoethyl cellulose (DE52) (Whatman, International Ltd Mainstone, England) column (50 x 300 mm) equilibrated in buffer A, and the column was washed with 750 ml of buffer A. The protein was eluted from the column in 15-ml fractions with a 1.6-L linear gradient from 0 to 0.85 M sodium chloride (NaCl) in buffer A. The fractions were analyzed using Ultraviolet (UV) visible spectroscopy (monitoring the absorbance at 280 nm) using a Cary 50 Bio spectrophotometer (Varian, Walnut Creek, CA) and sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Fractions containing the protein were concentrated by ultrafiltration using an Amicon stir cell (Millipore Bedford, MA) fitted with a membrane of 3000 kDa pore size. The concentrated fractions were loaded onto a 30 x 1200 mm Sephacryl S-300 (Amersham Biosciences Corp Piscataway, NJ) column equilibrated in buffer A, containing 100 mM NaCl. The apo-ACP was eluted in the same buffer and analyzed as previously described. Fractions containing apo-ACP were concentrated as previously described. Aliquots of 100 µl were frozen on liquid nitrogen and stored at -80ºC. A DTNB (Dithio-bis(2-Nitrobenzoic Acid)) assay for free sulfhydryl groups was preformed to determine how much ACP was in the holo form.[5]

Preparation of ACP Synthase

Using the method of Lambalot[5], a single colony of *E. coli* containing pACPSI was grown in 25 ml of LB containing 50-µg mL^{-1} of ampicillin at 37 °C for 7-8 hours. This culture was harvested at 3,000 x g for ten minutes and the pellet was resuspended in fresh LB media containing 50 µg mL^{-1} of ampicillin. The culture was then used to inoculate 4

L of LB media (50 µg mL^{-1} ampcillin), which was incubated at 37ºC with shaking at 200 rpm and grown to an OD$_{600}$ of ~0.5-0.6. Expression was carried out by the addition of 1 mM IPTG followed by incubation for 3 hours at 30ºC. The cells were harvested and stored as described for *E. coli ACP*. Harvesting took place at 10,000g for 6 minutes. The cell paste was frozen in liquid nitrogen and stored at -80°C.

The cell paste was resuspended in 3 mL of 50 mM Tris(hydroxymethylaminomethane) (Tris), pH 8.1, 10 mM magnesium chloride (MgCl$_2$), 2 mM dithiothreitol (DTT) and 5% glycerol per gram of cell paste, sonicated on ice in four 1-minute bursts, and centrifuged at 20,000 x g for 30 minutes. The supernatant was decanted into pre-chilled 50-ml tubes. DE-52 slurry equilibrated in Tris buffer (1 mL per mL of extract) was added to the clarified extract and mixed gently for 15 minutes followed by centrifugation at 5,000 x g for 15 minutes. The supernatant was decanted into pre-chilled 50-mL tubes and the DE-52 extraction was repeated. After the second extraction, the supernatant was decanted and the pH was adjusted to 6.5 using 1 M MES. The mixture was centrifuged at 20,000 x g for 15 minutes and the supernatant was applied to a 30 x 30 cm SP-Sepharose (Amersham Biosciences AB Uppsala, Sweden) column equilibrated in MES buffer (50 mM MES, 10 mM MgCl$_2$, 2 mM DTT, 5% glycerol, pH 6.). The column was washed in 250 mL of MES buffer and the protein was eluted using a 500-mL linear gradient of 0 to 1M NaCl in MES buffer.[4]

Synthesis of holo-ACP from apo-ACP

A reaction mixture consisting of a final concentration of 50 mM Tris-HCl, pH 8.8, 1.0 mM Coenzyme A, 5.0 mM DTT, 10 mM MgCl$_2$, and 207 µM apo-ACP was pre-incubated in a 37°C water bath. A final concentration of 200 nM of holo-ACP Synthase was then added to initiate the reaction, which was run for two hours at 37°C. The reaction was diluted to one liter with 25 mM MES, pH 6.1 (buffer A) and loaded onto a DE-52 column (2.5 x 20 cm) equilibrated in buffer A. The column was washed with 300 ml of buffer A, and holo-ACP was eluted with 25 mM MES, pH 6.1, containing 0.85 M NaCl. Fractions were analyzed by for protein using the Coomassie Blue dye procedure of Bradford. Fractions containing the protein were pooled and concentrated by Centriplus (MW 3,000) to approximately 1-2 ml. The protein was loaded onto a Sephadex G-25 (Amersham Biosciences AB) column (2 x 15 cm) equilibrated in buffer A. The holo-ACP synthase was eluted from the column with 100 mM Tris pH 8.0. Fractions were collected, analyzed, and concentrated using the methods previously described.[6]

Purification of Acyl-ACP Synthetase from Escherichia coli

The expression of Acyl-ACP Synthetase was carried out by growing cells containing pAaSH with vigorous aeration at 275 rpm in LB media as described.[7] The cells were grown to an OD$_{600}$ of 0.5, and induced by addition of IPTG to 0.4 mM. Expression was carried out for 4.0 hours at 37˚C before harvesting. After harvesting the cell were frozen in liquid nitrogen and stored at -80°C.

In 60 ml of 50 mM Tris-HCl pH 8.0, 26.1g of cell paste was resuspended, followed by the addition of 0.2 mg mL^{-1} of lysozyme. The mixture was allowed to stir and then the bacteria were ruptured using a French press. The supernatant was diluted with an equal volume of 2X column buffer (100 mM Tris, pH 8.0, 40 mM MgCl$_2$, 4% Triton), and was loaded on a Nickel-Nitrilotriacetic acid (Ni-NTA) (Qiagen Valencia, CA) column (equilibrated in 1X column buffer. The column was washed with 1X column buffer and the protein was eluted using Ni-NTA elution buffer (1X column buffer containing 250 mM imidazole). The protein was concentrated using an Amicon (30,000 membrane) and dialyzed against column buffer containing 5 mM ATP to remove excess imidazole and to stabilize the protein. The protein was aliquoted and stored at -80C.

Synthesis of Octanoyl-ACP

The reaction mixture, which consisted of 50 mM [4-(2-hydroxyethyl)-1-piperazinepropanesulfonic acid] (EPPS), pH 8.0, 3 mM DTT, 0.4 M lithium chloride (LiCl), 10 mM MgCl$_2$, 5 mM ATP, 2% Trition X-100, 50 μM Holo-ACP, 50 μM octanioc acid, acyl-ACP synthetase from *Escherichia coli*, and H$_2$0 were combined to a final volume of 100.3 ml. The reaction was allowed to proceed for approximately 5.0 hours at 37ºC. The reaction was diluted to one liter in cold 25 mM MES, pH 6.1 (buffer A), titrated to pH 6.1, and loaded onto a DE-52 column (2.5 x 10 cm) equilibrated in buffer A. The column was washed with 500 ml of buffer A, and the octanoyl-ACP was eluted from the column with 0.8 M NaCl in buffer A. Fractions containing protein (method of Bradford) were pooled and titrated to pH 3.9 with concentrated acetic acid. The octanoyl-ACP was precipitated overnight at 0ºC (ice water bath). The precipitate was pelleted at 20,000 x g for 30 minutes, and then dissolved in a minimal volume of 1 M 2-[N-cyclohexylamino] ethane sulfonic acid (CHES), pH 9.0, containing 1 M urea (buffer B). The dissolved protein was loaded onto an S-300 column (30 x 1200 mm) equilibrated in buffer B. The protein was eluted from the column in a fraction size of approximately 2-3 ml. Fractions displaying an absorbance maximum at A$_{280}$ and reacting with Bradford reagent were analyzed by native gel electrophoresis.

Purification of Acyl-Acyl Carrier Protein by gel electrophoresis

In preparation of the gel electrophoresis unit a solution with a final concentration of 13% acrylamide / 5% bis -crylamide, 0.37 M Tris, and 1M Urea was mixed and degassed to form the running gel. To this mixture 100 μL of ammonium persulfate (APS) and 10 μl of *N,N,N',N'*-tetramethylethylenediamine (TEMED) were added to facilitate polymerization which was allowed to occur overnight. A stacking gel containing 6M Urea, H$_2$O, Stacking Gel Buffer (0.5 M Tris-HCl, pH 6.8), acrylamide / bis-acrylamide, 50 μL APS, 10 μL TEMED was poured one to two hours prior to loading the gel. The gel electrophoresis unit was filled with electrophoresis buffer (0.025M Tris, 0.192 M glycine, 0.1% SDS), 500 ml in the upper electrophoresis buffer reservoir and 750 ml in the elution buffer reservoir. The lower buffer reservoir was filled with enough buffer to cover the gel. The apo-ACP sample was diluted 10-fold in deionized H$_2$O and

10 mL of it was mixed with 3.3 mL of sample loading buffer (0.1 M Tris pH 6.8, 30% glycerol, 0.03% bromophenol blue). The protein-dye sample was loaded into the gel tube of the electrophoresis unit. The gel was run at a continuous 7W, with an elution rate of 1 ml min^{-1}. Fractions were collected and analyzed using UV-visible spectroscopy. See appendix.[3]

V. Results and Discussion

Assessment of the activity of *Escherichia coli* acyl-ACP synthetase was preformed through monitoring of the production of radioactive octanoyl-ACP, using [8-^{3}H] octanoic acid as one of the substrates. [8-^{3}H] octanoyl-ACP was easily separated from [8-^{3}H] octanoic acid by gel filtration using calibrated Nick^{TM} columns (Amersham Biosciences AB). Control experiments established that the time-dependent formation of [8-^{3}H] octanoyl-ACP is absolutely dependent upon the *E. coli* acyl-ACP Synthetase (Figure 3). The specific activity of the enzyme was determined to be 2.1 x 10^{-5} U mL^{-1}. The purified *E. coli* acyl-ACP synthetase was used to synthesize tritiated, deuterated, and protiated octanoyl-ACPs. The properties of octanoyl-ACP are not significantly different from those of holo-ACP, making the purification of the product from the substrate difficult. Acyl carrier proteins are protected from denaturation at high pH by long-chain acyl groups (>C-8). This denaturation results in an increase in the Stokes radius of the protein, causing it to

Assay of *Escherichia coli* Acyl-ACP synthethase

y = 357.5x + 234
R^2 = 0.9866
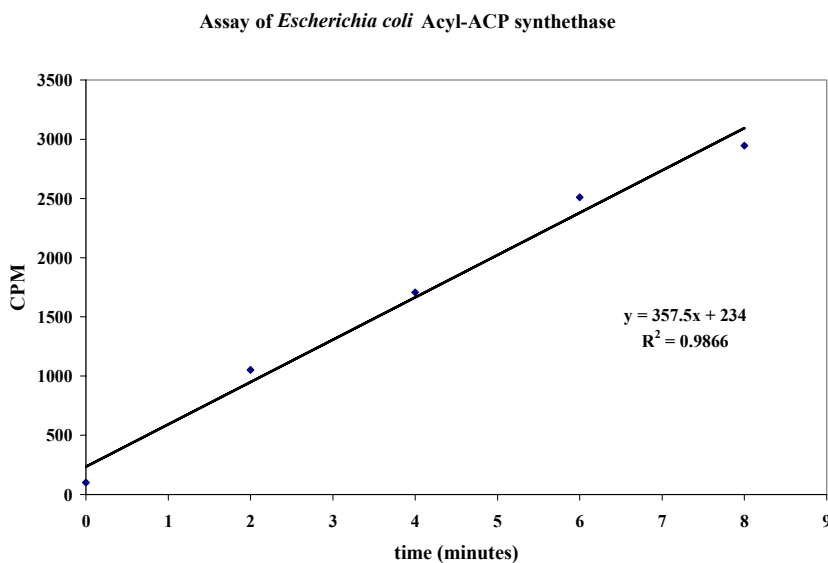
Figure 3: Production of octanoyl-ACP as a function of time.

migrate faster on gel-filtration columns, but slower by native gel electrohporesis. Thus, long chain acyl-ACPs migrate slower than holo or apo-ACP by gel-filtration.

This research was found to be a successful assessment of whether the *Escherichia coli* acyl-ACP synthetase could synthesize octanoyl-ACP, despite the fact that previous research has shown that this is not the enzyme's preferred substrate. It also has been

assessed that the presence of 1 M urea in the gel-filtration buffer may be a resource in the separation of ACPs.
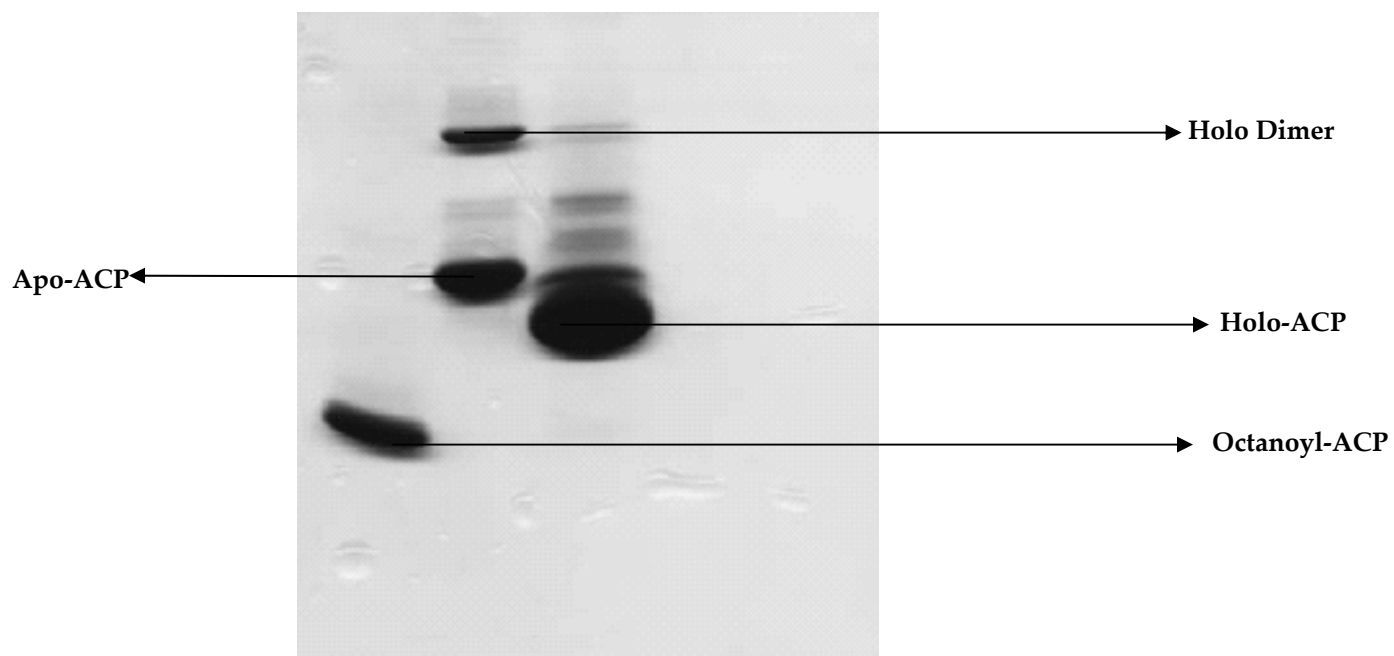
Figure 4: Native Gel Electrophoresis

REFERENCES

1. Garret, R.H., and C.M. Grisham. (1999) Biochemistry 2nd Ed. pp 101-601, Saunders College Publishing, Philadelphia.

2. Fuchs, J., L. Packer, and G. Zimmer. (1997) Lipoic Acid in Health and Disease, pp vii-10 Marcel Dekker Inc., New York.

3. Model 491 Prep Cell Instruction Manual. (2000) Biorad

4. Haas, J.A., and M.A. Frederick, and B.G. Fox. (2000) Chemical and Posttranslational Modification of *Escherichia coli* Acyl Carrier Protein for Preparation of Dansyl-Acyl Carrier Proteins. *Protein Expression and Purification* **20**:274-284.

5. Lambalot, R.H. and C.T. Walsh. (1997) Holo-[Acyl-Carrier-Protein] Synthase of *Escherichi coli. Meth. Enzym. Mol.* 27:254-262.

6. Shanklin, J. (2000) Overexpression and Purification of the *Escherichia coli* Inner Membrane Enzymed Acyl-Acyl Carrier Protein Synthase in an Active Form. *Protein Expression and Purification* **18**:355-360.

7. Juang, P and Cronan, J.E., Jr. (1994) Inhibition of Fatty Acid Synthesis in *Escherichia coli* in the Absence of Phospholid Synthesis and Release of Inhibition by Thioerterase Action **176**:2814-2821.

8. Indo, Y. and I. Matsuda, *Molecular Defects of the Branched–Chain α-Keto Acid Dehydrogenase Complex: Maple Syrup Urine Disease Due to Mutations of the E1α or E1β Subunit Gene*. In *Alpha-Keto Acid Dehydrogenase Complexes*, M. S. Patel, T. E. Roche, and R. A. Harris, Editors. 1996, Birkhäuser Verlag:Basel

Appendix

Native Gel

Running:

Reagent	Initial	Final	*Amount*
Acrylamide/Bis	30% / 1.12%	13% / 0.5%	8.7 mL
Tris	2 M	.37 M	3.7 mL
Urea	6 M	1 M	3.3 mL
H_2O			4.3 mL

* Note the above amounts are for a 20 mL gel; therefore, measurements should be multiplied by two.

APS	100 μL
TEMED	10 μL

Stacking:

Reagent	Amount
6 M Urea	3.3 mL
H_2O	9 mL
Stacking Gel Buffer	5 mL
Acrylamide/Bis	2.66 mL

APS	50 μL
TEMED	10 μL

Small Native Gel

Running:

Reagent	Amount
Acrylamide/Bis	6.7 mL
2 M Tris	2.8 mL
6 M Urea	2.5 mL
H_2O	3.3 mL

APS	38.5 μL
TEMED	3.8 μL (5.0)

Stacking:

Reagent	Amount
Acrylamide	1.3 mL
Stacking Gel Buffer	1.7 mL
6M Urea	1.25 mL
H_2O	1 mL

APS	12.5 μL
TEMED	2.5 μL

* Note the above amounts are for two small native gels.

30% Acrylamide Stock Solution (30%T/2.67%C)
Acrylamide   29.2g
Bis   0.8g

Dissolve into 70 ml H$_2$O

Stacking Gel Buffer Stock

0.5 M Tris – HCL, pH 6.8
Dissolve 6 grams Tris base in approximately 60 ml deionized water.
Adjust to 100 mL with deionized water and store at 4˚C

Loading the Sample:

Sample Loading Buffer:

pH 6.8
0.1 M Tris
30% glycerol
0.03% bromphenol blue

*Note samples are loaded 3 parts sample to one part sample loading buffer
 Ex: 10 µl sample; 3 µL sample loading buffer

Electrophoresis Buffer (10X Electrode[Running] Buffer):

1 Liter Stock Solution: Tris base 30.3g
                        Glycine 144.0g
                        SDS   10.0g
Dissolve and adjust to 1 L with deionized water. DO NOT adjust pH with acid or base.

*Note to run the native gel you use **1X Electrode (Running) Buffer.** To make 1 Liter of **1X Electrode (Running) Buffer** dilute 100 ml of **10X** stock with 900 ml of deionized water.

Native Gel Stain

Stain:

50% Methanol
10% Acetic Acid
0.1% Coomassie Blue

<u>1 Liter Stock</u>
500 ml Methanol
100 ml Acetic Acid
1 g Coomassie Blue
400 ml H_2O

Destain:

10% Methanol
10% Acetic Acid

<u>1 Liter Stock</u>

100 ml Methanol
100 ml Acetic Acid
800 ml H_2O

Dynamics of Nanotube Synthesis

Christina Guice, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Henry C. Foley, Head and Robb Chair
Department of Chemical Engineering
Pennsylvania State University

Bo Yi, PhD Candidate Graduate Researcher
Department of Chemical Engineering
Pennsylvania State University

Introduction

Carbon nanotubes are potential candidates for use in many functional devices such as field emitters and for use in nanoelectronics.[1] The synthesis of aligned nanotube bundles is of great importance in this area,[1] that is the goal of this experimentation. Carbon nanotubes are the world's smallest solid-state light emitters, and the first electrically controlled, single molecule emitter.[2] Nanotube transistors have been "successfully fabricated and tested using individual multi-wall or single-wall as the channel of a field-effect transistor."[2]
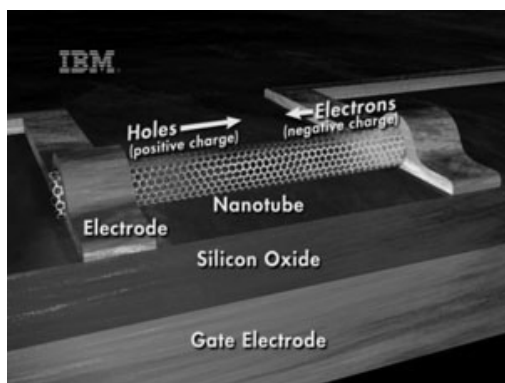
Figure 1. IBM scientists simultaneously inject positive and negative charges into a carbon nanotube through the source and drain electrodes at its two ends.[2]
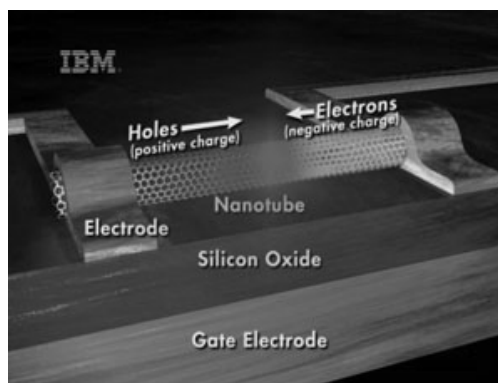
Figure 2. When the electrons and holes meet in the nanotube, they neutralize each other and generate light.[2]

Carbon nanotubes may be defined as greatly elongated fullerenes with exactly twelve pentagons and millions of hexagons, while a fullerene is a closed shell all carbon molecules with an even number of carbon atoms.[3] An alternate definition is that carbon nanotubes are graphite sheets wrapped into a cylinder and capped by fullerene-like structures[4] (See Figure 3).

Figure 3. A fullerene, graphite plates, and carbon nanotube to better aid in the description

A fullerene.[5]   Graphite Plates.[5]   Carbon nanotube.[5]
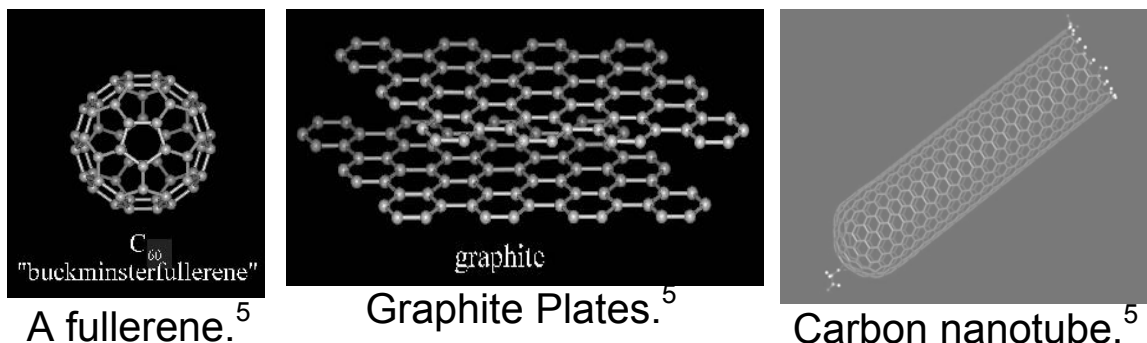
There are two forms of carbon nanotubes: multi- and single-walled. Multi-walled carbon nanotubes (MWNT) are comprised of concentric cylinders formed around a common central hollow with spacing between the layers close to that of the interlayer spacing of graphite (0.34 nm).[3] Single-walled carbon nanotubes (SWNT) differ from multi-walled nanotubes in the fact that they are composed of a single graphite sheet.[3]

MWNTs were initially discovered by Iijima in 1991 via arc-evaporation of graphite rods in an inert atmosphere inside the cathodic deposit.[6,7] This led to an explosion of nanotube research being conducted. SWNTs were first synthesized in 1993 in an arc discharge apparatus.[8]   In this experiment; the method of ferrocene pyrolysis was used to obtain bundles of aligned MWNTs.

Experiment

Via the method of injection chemical vapor deposition (CVD), "aligned nanotube films were obtained from temperatures of 550 °C and above"[7]. Using that same method, the length of the nanotubes increased until reaching a temperature of 680 °C, and then proceeded to decrease in length. However, in this experiment, using ferrocene pyrolysis, a mixture of 80% argon (Ar) and 20% hydrogen (H_2) is used as the carrier gas, and the carbon source is acetylene. All of the nanotubes in these experiments are synthesized at 700 °C on a quartz plate, both orthogonal and parallel to the gas flow, while placed inside a quartz boat.

Prior to heating, the Ar:H_2 mixture is flowed into the furnace. The ferrocene is heated at a constant rate in the first stage of the furnace, until it has sublimed, using a temperature controller. This heating takes place before the acetylene begins to flow through the reaction tube. There were various temperature settings for the heating of the ferrocene in the first stage. The temperature settings for the heating of ferrocene varied per sample. However, the starting temperature was always room temperature (approximately 23 °C), and the final temperature ranged between 160 and 170 °C.

For Sample 1, the heating rate of the ferrocene was: from starting temperature to 90 °C in the first 20 minutes, 90 °C to 130 °C in 5 minutes, 130 °C to 150 °C in five minutes, 150 °C to 160 °C in 5 minutes, and held constant at 160 °C for 15 minutes. The

flow rates of Ar, H_2, and acetylene, for this sample, are 150 sccm (standard cubic centimeters), 37.5 sccm, and 8 sccm respectively. The mass of ferrocene used was 200 mg.

For Sample 2, the heating rate of the ferrocene was: from starting temperature to 60 °C in the first 15 minutes, 60 °C to 90 °C in 5 minutes, 90 °C to 170 °C in 15 minutes, and held constant at 170 °C for 15 minutes. The flow rates of Ar, H_2, and acetylene, for this sample, are the same as those for Sample 1. The mass of ferrocene used was 300 mg.

For Sample 3, the heating rate of the ferrocene was: from starting temperature to 60 °C in the first 15 minutes, 60 °C to 120 °C in 5 minutes, 120 °C to 170 °C in 10 minutes, and held constant at 170 °C for 50 minutes. The flow rates of Ar, H_2, and acetylene, for this sample, are 210 sccm, 52.5 sccm, and ~6 sccm respectively. The mass of ferrocene used was 400 mg.

At the same time, the second stage of the furnace heats at a constant rate until it reaches 700 °C (during the first 20 minutes of the experiment), which is the temperature chosen to grow the aligned MWNT bundles. Once the temperature of 700 °C is attained, after 20 minutes, the acetylene flow begins, and the temperature is held constant at approximately 700 °C for 30 to 60 minutes. Once the process is completed, the furnace and nanotubes are allowed to cool until they are both at room temperature. The products were then characterized with a Hitachi S3000-H scanning electron microscope.

Results/Discussion

The temperature controller used for the experimentation was not very accurate. Therefore, for each sample, the actual temperatures of ferrocene after each time interval were observed. The actual heating rate for Sample 1 was: 18.7 °C to 87.5 °C in the first 20 minutes, 87.5 °C to 125.5 °C in 5 minutes, 125.5 °C to 147.3 °C in five minutes, 147.3 °C to 158.3 °C in 5 minutes, and held constant at approximately 160 °C for 15 minutes.

The actual heating rate for Sample 2 was: 20 °C to 58.8 °C in the first 15 minutes, 58.8 °C to 86.9 °C in 5 minutes, 86.9 °C to 166.5 °C in 15 minutes, and held constant at approximately 170 °C for 15 minutes.

The actual heating rate for Sample 3 was: 22.1 °C to 59.8 °C in the first 15 minutes, 59.8 °C to 115.8 °C in 5 minutes, 115.8 °C to 166.5 °C in 10 minutes, and held constant at approximately 170 °C for 50 minutes.

Effects of Ferrocene Mass

For Sample 1, 200 mg of ferrocene was used. This experiment yielded short, aligned nanotubes that were about 30 microns (μm) in length. These nanotubes proved to be the most aligned nanotubes that were synthesized in the series of experiments, and they can be seen in Figure 4. For Sample 2, 300 mg of ferrocene was used, and the results were aligned nanotubes, but they were shorter than and not quite as aligned as those found in Sample 1. The nanotubes from this sample had a length of 12 μm. However, the far edge of the plate, furthest away from the ferrocene source, showed nanotubes with the length of 22 μm. For Sample 3, 400 mg of ferrocene was used. The only material remaining on the quartz support was a very thin layer of carbon nanotubes.

Effects of plate positioning

The concentration of ferrocene was assumed higher near the top of the reaction tube than it is at the bottom. To confirm or disprove this theory, there were two positions in which the quartz plate was placed: parallel and perpendicular to the flow of the gases. When the plate is placed parallel to the gas flow, there is a low rate of aligned nanotube synthesis. However, when the plate is placed perpendicular to the gas flow, the rate of successful aligned nanotube synthesis is high in comparison to the plate being placed parallel to the gas flow. Samples 1 and 3 were both synthesized with the plate perpendicular to the flow of gases. However, in Sample 2, the plate was placed parallel to the gas flow.

Figure 4. SEM images of the MWNTs from Sample 1
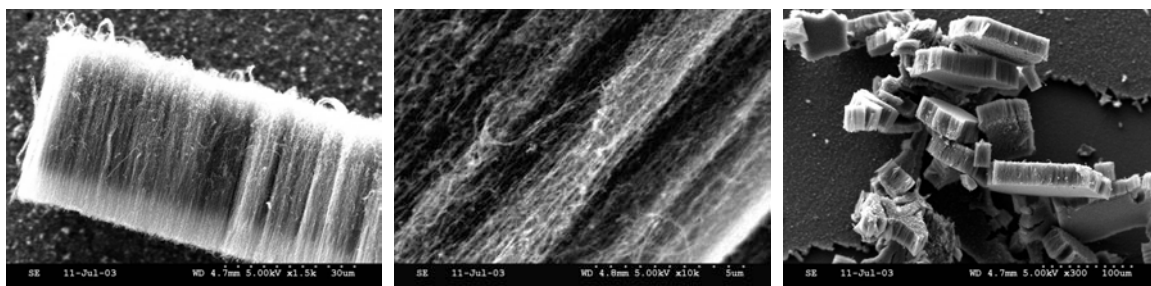
Figure 5. SEM images of the MWNTs from Sample 4
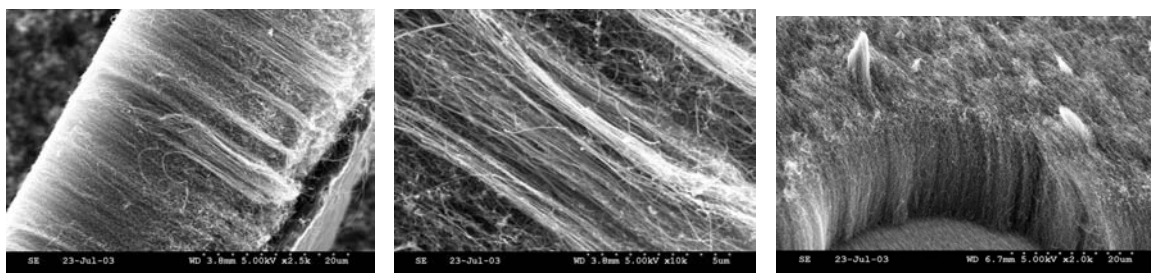
Conclusions

The mass of ferrocene was varied between 200 and 400 milligrams (mg). The experimentation concluded that 200 to 300 mg was sufficient for the growth of nanotubes. However, 400 mg was a little too much. Therefore, aligned nanotubes will form almost consistently when using between 200 and 300 mg of ferrocene as the catalyst. However, when attempts were made to synthesize nanotubes with 400 mg of ferrocene, the results showed that there was a possibility that nanotubes formed, but they were too short to be characterized by the SEM.

The nanotubes also form rather consistently when the quartz support is perpendicular to the flow of gases as opposed to parallel. It is assumed that this has to do with the concentration of ferrocene at different levels of quartz tube. It is believed that the concentration of ferrocene will be higher near the top of the reaction tube, in

comparison to the concentration near the bottom, where the nanotubes are being grown. In addition, it appears as though when the support is parallel to the flow, generally any nanotubes formed blown off the quartz support into the tube.

Acknowledgements

I would like to thank NSF, the McNair Scholars Program staff who put me in the position to conduct this research, Dr. Henry Foley and his research group, and my family for their constant support.

References

1. Satishkumar, B.C.; Govindaraj, A.; Rao, C.N.R. *Chemical Physics Letters*. 307, **1999**, 158.
2. *IBM Research News*. www.research.ibm.com/resources/news/20030501_cntle.shtml
3. Ajayan, P.M. *Chem. Rev.* 99, **1999**, 1787.
4. Huczko, A. *Applied Physics A*. 74, **2002**, 617.
5. *The Smalley Group at Rice University*. http://www.ruf.rice.edu/~smalleyg/image_gallery.htm
6. Rahul Sen, A. Govindaraj, C.N.R. Rao. *Chemical Physics Letters*. 267, **1997**, 276.
7. Charanjeet Singh, Milo S.P. Shaffer, Alan H. Windle. *Carbon*. 41, **2003**, 359.
8. Hongjie Dai, Andrew G. Rinzler, Pasha Nikolaev, Andreas Thess, Daniel T. Colbert, Richard E. Smalley. *Chemical Physics Letters*. 260, **1996**, 471.

Egypt's Bounty via the Humble Potstand

Maria A. Gutierrez, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Elizabeth Walters, Associate Professor of Art History
College of Arts and Architecture
Pennsylvania State University

"I have brought you all good offerings and nourishment, every good thing in Upper Egypt, in Lower Egypt, all life, stability and power…"

<div align="right">Sesostris I pavilion at Karnak, Dyn XII
(Lacau-Chevrier, 211)</div>

Offerings were of paramount importance to the ancient Egyptians. The passage above is spoken by the fecundity figure often known as Hapi and permits us to recognize how significant and encompassing the blessings desired. On the walls of this pavilion Hapi and others offer to their king Sesostris (fig. 1) the bounty of Egypt as also connoting the best in life.

In looking at various offering depictions throughout the Egyptian civilization, common patterns can be identified in the way offerings are presented to the gods, kings and the deceased. Some scenes like "the offering table scene" in mastabas repeat themselves over and over. They reveal religious connotation as well as cultural contexts. Just as some scenes became standardized, the objects used for the presentation of gifts also came into conformity. Offerings were usually elevated from the ground and placed on stands, table stands or potstands. The elevation of offerings goes back to Predynastic times and continues through Roman times. Potstands as a medium for offerings are the focus of my research.

Background

The purpose of this study is to investigate the types, functions and contexts of stands and potstands from the earliest examples (middle Predynastic c. 3600 B.C.) into the Roman era. I am interested in their significance and connections to religious and funerary rites. Central to my research is a deposit of potstands found in the 2003 field season of the Temple-Town Hierakonplis Project directed by Drs. Elizabeth Walters (Art History), Shelton Alexander, David Gold and Richard Parizek (Geosciences) of the Pennsylvania StateUniversity, and in which I had the privilege to participate.

My study involves a survey of ancient Egyptian potstands, published as early as the seminal reports of Flinders Petrie and for Hierakonplis the report of the first excavators Quibell and Green (early 19th c AD). My survey involves material from excavations also at Buto, El Kab, Elephantine, Naqada, Abydos, Amarna, Giza, Saqqara, etc. with material following Petrie and as recent as 2000. I have confirmed the date of the 2003 potstands (Dynasty I as recognized by Dr. Walters) and found information on other potstands from graves and temples, helpful to understanding the importance and use of offerings on stands, and helpful to assessing this deposit, suggestive of possible cultural context.

As ubiquitous as offerings on stands were in ancient Egypt, potstands have not been the subject of any serious study. A simple overview informs us that potstands had a significant place in Egyptian culture, judging from their longevity and the fact that the shape of the potstand was even used in hieroglyphs as a symbol for offerings (fig.2). In addition, tiny model pots on stands (fig. 3) are known from the excavations at Hierakonpolis found in the temple by Quibell and Green 1897-1899 and a model pot on stand found with a few votive figurines by Dr. Fairservis in 1981 (Fairservis, Hierakonpolis Project vol. III 1984). These little models possibly as early as Dynasty I seem to us curious because the vast numbers of votives from the 'main deposit' in the temple at Hierakonpolis were animals; that model potstands were the one item that represents man-made gifts, a model of gift, suggests that like the hieroglyph (pot) they are the essence of giving.

Surprising continuity exists for the tall, trumpet-like stand, serving as a table for bread in offering scenes. Evidence indicates this appeared as early as Dynasty II and certainly became a tradition by Dynasty III (which will be discussed later with Funerary rites), and continued past the beautiful examples of the last empire age, the New Kingdom, Dynasty XIX and into Roman times. Examples of the New Kingdom include the banquet of Vizier Ramose (Garis Davies 1941), funerary or ceremonial, and the tomb of Panehsy (Baud & Drioton 1932), where offerings for gods and specific deified rulers were placed on high stands. Burning incense we have already seen on such a stand in the earlier pavilion of Sesostris I at Karnak, Dynasty XII (see fig. 1); hence the tall stand could serve bread, food, water, burning incense, and as we will see even more offerings in the New Kingdom.

This study is the first to focus on potstands, helpful to scholars and archaeologists with my survey of these vases from Predynastic through Roman, 4th c A.D. The core material, the 2003 potstands, provides a preliminary report useful to our Temple-Town Hierakonpolis Project and helpful to future publications. Discussed here are the best examples to show cultural importance of potstands this is only a small part of my investigation. More examples are to be found in my catalogue, essential documentation for and serving as an appendix to my senior thesis (Schreyer Honors thesis for Art History and my mentor Dr. Elizabeth Walters). This catalogue includes descriptions and classifications given by the original excavators, and this organized material can be used as reference in the future. To my knowledge, this study is a pioneer in focusing simply on one type of pottery (stands) and exploring its art, history and significance.

The most exciting part of my study is the cultural context that will be explored. Culture and religion in ancient Egypt were intricate, and as found with many other civilizations, a large portion of traditions and beliefs were not written down or evident in excavated finds. Future and further work at Hierakonpolis may offer more evidence of use and meaning, hopefully even a glimpse at an ancient town life.

Potstands in Egyptian Temples

The use of potstands can be best learned from texts and illustrations from late Egyptian temples (Ptolemaic, 3rd to 1st c BC temples). Fortunately these late temples have remarkably preserved much of their original structure, wall inscriptions and reliefs. The inscriptions and illustrations are an invaluable source of information, since they include

descriptions of the temple room by room, giving their name, dimensions, purpose and decoration. Inscriptions also explain activities and ceremonies that took place everyday and for festivals throughout the year. They do so with such detail that, according to H. W. Fairman, it is possible to "reconstruct the furnishing and equipment of certain rooms, to tell when, how and where the offerings were prepared, to indicate the precise doors through which they were introduced into the temple, to trace the order of the ritual and the route of the great processions, and even know what happened to the offerings after the services and festivals were ended". Fairman also suggested that the texts are based on ancient traditions going back to very early times, judging from the vocabulary and content used (Fairman, 165-166).

At Edfu, fortunately for me, near Hierakonpolis, there is one of the best-preserved temples: the Temple of Horus, the falcon god. It was begun in 237 BC and completed in 57 BC (fig. 4) and is the only Ptolemaic temple that fully finished and standing with its original roof intact. Only the obelisk at the entrance and some chapels on the roof are no longer in place. The sacred lake, temple storehouses and other structures, however, are still buried under the modern town adjacent to the structure (Fairman, 166).

From the wall inscriptions translated by M. Alliot (1949) and Fairman (1954) on this temple at Edfu, we know that there was a Daily Service, composed of three main ceremonies, at dawn, at midday and at sunset. The morning service, which was the most important one, began with extensive preparations before sunrise. Two priests had the duty of filling the libation vessel from the sacred well (fig.5) for the replenishment of the water in the temple. Having done so, they both walked around the ambulatory (in a counter-clockwise direction), one carrying the vase and the other one walking in front him, censing the vessel. Then they entered the temple through the door on the west that leads to the Chamber of the Nile, and then towards the Inner Hypostyle Hall. The water was blessed and dedicated and then two other priests would have to refill all the libation vessels from the temple. Meanwhile, the offerings were brought through the east door of the Inner Hypostyle Hall. The priests had been busy before, slaughtering an ox and preparing the offerings that would be given to the gods. At the appointed time, the offerings were escorted and censed by priests and then taken into the Hall of Offerings.

We know that high stands were used when offerings were presented to the gods. Inscriptions from the Hall of Offerings indicate that high metal stands served as altars on top of which food was arranged. Sylvie Cauville states in *Edfou* "it was in this room that the food offerings were placed, either on wood dressers or   (≈)   on metal altars. The   (𝕀) purification and offering scenes that decorate the walls reveal the action of the priests and contribute to the magical nourishment of the god. Offerings did not leave the hall; the god was fed from the aroma of the food" as the doors to his innermost shrine were open (Cauville, 29).

Relief decorations in Sanctuary walls at earlier temples like the Temple of Seti I at Abydos Dynasty XIX (fig. 6) confirm that offerings, many on high stands, were part of temple service. The innermost shrine has in one of its walls a relief showing the god Amun's gifts framing the god's innermost shrine, a boat. This traveling boat (Barque to Egyptologists) was an actual container for the cult image. The statue of the god would be placed in the central receptacle of the boat, doors shut and hidden, as priests would carry the boat on poles. Beautiful flower offerings and libations are placed on a single metal

stand (fig.7) or on twin golden stands (fig.8). Lotus flowers seem to be prevalent, as well as the traditional Nemset vases for water.

Libation offerings were brought to 3 places: Hall of Offerings; the *Place where the Gods Reside* (Hall of the Ennead) and finally the innermost shrine, where the "portable traveling shrines" of the gods (deities that dwelled with Horus in the temple) were kept (Fairman, 178). Then the lead priest would enter with much formality by the main door of the Outer Hypostyle Hall. Upon doing so, he recited a Declaration of Innocence and then he was taken to the House of the Morning (place of purification) to be ceremonially cleansed, dressed and endowed with authority. When this was completed, hymns would start to be chanted as the officiating priest marched in procession towards the Sanctuary, whose doors were closed and seals unbroken. It is interesting to note at this point that the temple architecture conformed to the progression towards holiness and the structure was designed to maintain a deep sense of mystery and power. The only are beyond the Forecourt or Court of the Pylon (except for the Food-altar) that had access to sunlight was the southern wall of the Outer Hypostyle Hall, of which upper half was a screen. The rest of the temple was without external illumination and the inner most parts were in complete darkness. The light of torches "used during services [played] on brilliantly coloured reliefs, on the gilded surfaces of doors and shrines, and on the cult vessels [which] must have increased the sense of awe and majesty and grandeur" (Fairman, 172). Evident in the cross section of the temple of Karnak, the feeling was accentuated by the fact that as one progressed unto the holiest places, the floor level was raised and the height of the roof lowered (fig. 9)

As the priest reached the Sanctuary, the service consisted of seven stages. First, he went up the stairs of the naos, broke the seals of the doors and in doing this revealed the statue of Horus. Then came the ceremonies of uncovering the face of the god and, seeing the god, where the priest said: '*I have seen the god, the Power sees me. The god rejoices at seeing me. I have gazed upon the statue of the Divine Winged Beetle, the sacred image of the Falcon of Gold'* (Fairman, 180). This was the most important part of the service, since the Horus had come to dwell in his house again and entered his statue. Following this, the god was adored and myrrh was presented to him. It symbolized the presentation of a meal, and the formula indicates that 'The scent of myrrh is for thy nose, it fills thy nostrils, thy heart receives the meat-proportions on its scent' (Fairman, 191). The three last phases had to do with the cleaning, grooming and dressing of the god. First, the statue was touched with unguent and four colored clothes were presented to it. Second, the statue was purified with water from the customary green and red vases. Finally, the god was censed and fumed extensively, the priest withdrew and the shrine and Holy of Holies were closed again.

While these rituals were taking place in the Sanctuary, other priests went around the corridor around it and into the chapels, and performed short versions of the morning service. Hence all the gods and the entire temple were "awakened, washed, dressed, fed, and made ready for another day" (Fairman, 179-180). It is highly likely that after this, the rites of the Reversion of the Divine Offerings took place. Once Horus was satisfied with his offerings, the gifts would be reverted to the priests and distributed among them according to rank. The midday and evening services were also abbreviations of the morning ceremony and less important. They were repetitions of the morning one only

less elaborate, although the evening service took place in the Throne of Re and not in the Sanctuary (Fairman, 180-181).

Edfu had, in addition, two calendars of festivities throughout the year, which were beyond forty. These involved elaborate ceremonies, worship and offerings and thus the use of stands and potstands. The most important were the New Year, the Coronation of the Sacred Falcon, the Festival of Victory and the Sacred Marriage. (Fairman, 183).

Potstands and Funerary Rites

Special vases were not exclusive to temples, but were also used in the funeral such as the "Opening of the Mouth Ceremony" performed at least as early as the 4th Dynasty, evident from the tomb of Metjen at Giza (Kanawati, 31). The developed form of the Opening of the Mouth ceremony known from Dynasty XIX (Budge 1909) took place after mummification; the body was "reanimated" through more 30 rituals and the recitation of texts took place over the statue of the person. The mummy's bodily functions were restored; his eyes, ears, nose and mouth were opened and his limbs re-united. A very beautiful representation of a small part of this ceremony is found in the tomb of Tutankhamun (fig. 10). According to W. Budge, the first 7 rituals consisted of censing and sprinkling. The Sem-priest would walk around the statue sprinkling water, burning incense and reciting texts. During the eleventh ceremony, a bull was slain and his leg and heart were presented to the statue. The bull was a symbol of Osiris, and just as Osiris had been slain and the restored, so the deceased would be restored through the acquisition of the life and power of Osiris. The bull's leg symbolized the Eye of Horus and the heart served the function of transferring the seat of life and power of the animal to the statue. The Sem-priest also presented geese as a sacrifice because they symbolized the enemies of Osiris and killing them ensured protection. After this, the same priest would touch the statue with the "two divine axes" and thus opened his eyes and mouth. Also, he would rub them with a red substance, milk, oil, stibium, etc. and the statue was dressed with collars, given a scepter and a mace, among other things (Budge 1909, 82-111). The rituals guaranteed that the *ba* and the *ka* (soul and life force of the person) would return to the mummified body and the deceased could have an existence in the after-life.

Traditional vases apparently were used as part of the Opening the Mouth Ceremony. Simple flared cups have a long existence (Walters observation: Predynastic to Roman bronze cups). They are part of model vases from Dynasty V and VI (MMA NY), "Opening of the Mouth" kits(?), made in contrasting materials, a black set with a white stone set. Among the vases found in the royal tombs at Saqqara, Dynasty I (fig. 11), several are special, footed and made of different materials such as crystal or black basalt topping limestone. Although this royal grave was rich in vases, often in store-areas of the mastaba tomb, offerings next to the sarcophagi were lost as the grave was robbed and deprive us of the possible use of vases for the tomb occupant. However, cups of burnished black pottery and basalt found in great numbers in these early tombs have later counterparts in the low cups with incense seen in Tutankhamun's opening of the mouth ceremony.

Offerings for the graves beyond the grave goods were numerous and known since early times (Predynastic through Roman). Stands, table like stands and footed vases were part of the offerings shown on reliefs from Dynasty II through VI (pyramid age, Old

Kingdom) and later in painted coffins Middle Kingdom and the rich scenes on the tomb walls in the New Kingdom.

The stands come in a variety of shapes, heights and materials, and began to be part of standard representations as early as Dynasty II. An example of this is the relief of a Dynasty II princess (fig.12), found and documented by W. B Emery in Saqqara, a site near Cairo. The relief shows her seated before an offering table consisting of a big bowl with bread on a high stand. We know she is a princess because part of the inscriptions on the right contains hieroglyphs commonly used to describe a "royal daughter" (e.g. a Gardiner's G38 and M23).

From Dynasty III, we learn of a hierarchy in the offerings with Hesy-Ra. One of the most famous reliefs (also in Saqqara, fig.13) shows him proudly displaying his office, royal scribe (scribe's tools draped over his shoulder) as he is seated before his offerings. This panel is one of 11 different panels of precious wood that were once placed the in the niches of his tomb's chapel *serekh* or offering wall. On the right Hesy-Ra has immediately in front on him 2 potstands, first among the offerings. The upper one has a water vase and the lower one a bowl for food. And very prominent in this relief is the Canonic offering table. The stand on the Dynasty II relief is stumpy, but during the Old Kingdom it would develop into a slender high stand like Hesy-Ra's that would be used until the end of the Egyptian Civilization. Moreover, when observing these two reliefs, it is extremely interesting that both stands have a triangular piercing at the bottom, a very characteristic feature of the tall stands (table or flute-like stands for vases) during the Old Kingdom.

Potstands and table-like stands of various materials were prestigious. I have found no more than two small potstands and three platter stands in the published and excavated Prehistoric graves at Naqada and el Ballas (c. 3300-3200 BC) for example. J. de Morgan made a beautiful record of one of them (fig.14) which shows the remains of the deceased, in a fetal position, surrounded by goods deposited in his tomb. It is significant that the table stand of alabaster (like one found in the Cairo Museum, fig. 15) as placed by the man's head, next to the identity of the owner. The table was intentionally positioned, showing a hierarchy in which the stand was the most important. It is possible that stone or metal vases and stands were prototypes for the subsequently common clay examples. Pierced stands have a long tradition. Petrie documented various Prehistoric potstands from the same site (Naqada, fig. 16) with a vide variety of piercings: huge circular holes (no. 88), small circular openings (84b) and irregular sloppy triangles (84a, 85 and 86). Unfortunately, we do not see what the stands held. The ring stands would have supported a vase, but we can only make educated guesses about the shape, material, purpose, etc.

Potstands at Hierakonpolis and the Temple-Town context

The Temple of Hierakonpolis provides exciting evidence for religious and possible royal use of standed offerings. Previously, their significance had been overlooked because scholars and excavators at the site have focused only on the royal procession, the headdress, the costume and the entire concept of the first king of Egypt. Hierakonpolis is the site where the famous palette of Narmer was found (fig. 17), the visual document that

gives us evidence of that Narmer was the first king of Egypt, the Egypt we know. Important to my study is another object owned by Narmer, also found in the "main deposit" of the temple: his royal mace (fig.18). This object would have been used in ceremonial occasions by the king as a symbol his authority. It has various carved scenes, the most important one depicting a ruler seated in an elevated shrine. Narmer is seated within the a shrine and wears the crown of Lower Egypt while the goddess Nekhbet as a vulture flies overhead as protection and as a projecting of his rulership also in Upper Egypt. Below the throne are two fan bearers and immediately behind it, the high priest, bodyguards and a figure with the title *servant of the king* (Quibell, 9). Relevant to my study, however, is another detail of the mace (fig. 19). The scene shows, according to Quibell, "a vase upon a stand, and an ibis, possibly connected with offerings before Tahuti" (Quibell, 9). I agree that the ibis is an actual god itself (Thot, god of writing, or Tahuti according to Gardiner's list of hieroglyphs). The offering in its simplest form (the vase on a stand) is within the god's property, framed by a fence judging from the enclosed area beneath the Ibis and the standed vase. This scene is the first preserved depiction of a potstand in Egyptian history, dating to Dynasty I.

The Temple-Town Hierakonpolis Project extends the use of potstands to the new context of town, community, and possibly religious or royal use. Dr. Fairservis, the founder of the project, excavated from 1967 to 1994 and was the first archaeologist to re-map the site after the early excavations and to conduct professional work in the site. Also, his work was the first one to connect the property of the god (temple) with its town. From the map he developed (fig. 20) one can clearly distinguish the temple and its enclosure, and to the west of it the ancient town that sustained it. The darker structure in the town is believed to be a palace because of its elaborate entrance and niche wall decoration. At this period of time, the only royal buildings and tombs at Saqqara, Dynasty I had niche-decoration walls (Emery, 1949). During the 1981 excavation that he led, two noteworthy objects were found. One of them is a faience model vase on a stand (fig. 21) that looks very much like one found by Quibell (fig.3). The potstand part is obviously broken off. The second object is a fragment of an actual potstand (fig. 22). This fine piece is made of clay and it has both circular and triangular piercings. Curiously, both were found in the ancient town part, near the palace. The area is neither part of the temple or part of a cementery (the are no burials). It is exciting to learn that potstands were being used in a different way than discussed before.

During the 2003 fieldwork with Dr. Walters, an exploratory trench of 2 x 2m. was opened in the 10N11W quadrant, an area previously unstudied and distant from both the temple and the town (fig. 23). To our surprise, it was amazingly rich in potsherds, counting up to 7,446 in number and containing a wide variety of pieces that were as late as Roman and Islamic. The very special find at Hierakonpolis in the 2003 trench, however, was the first preserved deposit of potstands. They provide new evidence for activity in the western part of the ancient town and could be dated up to c.3200 B.C. A close up photograph (fig. 24) reveals the different types of potstands, all together near a limestone block that is thought to have been a working area (e.g. a table). Dr. Walters is currently working in the drawing reconstruction of the material, and as of now, her careful analysis reveals that there were at least three different types of stands: platter stands with a broad shallow mouth, bowl stands with deep dishes and ring stands. The deposit rests in a depth of 0.50-0.58 m., but it is intriguing that the potstands are not all

oriented in the same angle and few of them are complete. A much later industrial layer overlay and cut into the potstands, but it is clear that they rest on the original floor (0.50-0.58 m). Dr. Walters has also created a drawing of the trench along with side drawings and strata differentiation (fig.25). Evident from her documentation: soil barely covers a dense salt layer; immediately below is the thickest stratum: the industrial layer. It caps and cuts into the deposit of potstands, of which several retain their upright position as placed on a floor. The potstands lay next to a mud brick wall that is aligned with walls in the already known structures in the town, specifically the Palace (excavated by Fairservis in 1969-1981) but further exploration is needed to confirm this and to assess the town's occupation and growth near the Palace. The location of this unusual deposit of potstands could be explained as what remained of a workshop that in ancient times supplied pottery to the town, or palace or the temple (or to all?). These suggestions must await clarification from further excavation at Hierakonpolis.

The drawing and recording of the 2003 potstands are essential. These complete potstands and platter stands permit us to date them by close comparison with vessels other sites. Potstands of several shapes were found in the 1897-1899 excavations at Hierakonpolis. As seen from a drawing by Quibell (fig. 26), two stands resemble the 2003 potstands (5a and 5b). They come from the temple of Horus, but there is no discussion and no particular context given. Fortunately, there are many other sites that have potstands. Important is the work of Petrie. His *Prehistoric Egypt; Corpus of prehistoric pottery and palettes* (Petrie 1974) includes vases that he excavated from Naqada and Abydos, and date to Dynasty I. Most recently published example from Buto (fig. 27) has been dated stratigraphically to Proto-to Early Dynastic time; this potstand is bowl with stand with wide rims on the top and bottom; it has a relatively narrow "waist" in the middle and triangular pierced opening in the stand. It is similar to the 2003 Hierakonpolis potstands in the low bowl and squat proportions and confirms the sequence dating of Petrie (1890-1920's), a tool that excavators continue to use today.

A wonderful resource to learn how potstands were used are tomb walls. The tomb of Nefer and Ka-Hay (Dyn. V, Saqqara, c. 2500 BC) has an elaborate banquet with members of the family facing the tomb owner (fig.28); although Nefer and Kahay each have the high offering table as is customary, and the tomb owner " [is] leaning against a staff, looking down on a series of offerings piled on tables, plater-like stands and chests, [and this] dominates this scene of the south wall […] The top row shows a quantity placed on low offering plates and various beverages in jugs resting in ring-stands. The latter are always placed at the right side of the plates" (Moussa & Altenmüller, 31). These scenes enrich the customary west wall where offerings were given as well as represented to serve the dead into eternity. The deceased is given a wonderful variety of food and fruit on low platters-tables and broad bowls with stands. Here, the ring-stands held beer, the low-platters bread (triangular shaped loaves), fruit, vegetables and even other vessels. Platter stands from Hierakonpolis 2003 may have also served meat, bread and fruit, but it is important that they did so not for a tomb, but for use in the ancient town. Further study and future excavations will hopefully clarify that use in the town. By their location to the northwest of the archaic building, identified as a possible Palace (Dynasty I or earlier), were these potstands stored for use in the Palace or for other structures, the temple of Horus to south or thus far unknown shrines? Definitions of the architectural and cultural contexts await future excavations.

This study on potstands is ongoing and will be expanded with incoming new data. Several questions to be addressed include the following: Were potstands of different sizes and heights designed for specific types of vases or foods? How do potstands and platter-table stands vary through time and by region? The tall trumpet type stand has a very long life yet was it interchangeable with tall incense burners? Were the openings, triangular and pierced holes, useful or decorative? Based on metal prototypes? Or stone? Was the material, type of stone, metal, or quality of the clay and finish important to their use and possible meaning?

My conclusion thus far, recognizes that the high potstand first seen as a gift to the Ibis, Dynasty I mace of Narmer, may represent a separation in offerings. Potstands for tombs at the same time are not only rare but also low to the ground. It is interesting that the high stands (fluted ones) exist from Dynasty II on tomb reliefs for the prestigious table with bread offerings. Was the latter, this table stand, borrowed from the gods? Our data is very limited for Dynasty I and from my survey of potstands, archaeological context is often not given in the publications. More recent excavations (post 1950's) do provide specific information from the stratigraphy, associated finds, and tomb or structure context. Funerary evidence is dominant, hence material excavated from temple or town sites are very important. It is hoped that further work at Hierakonpolis Temple-Town will offer evidence of town life and clarify the potstand deposit of 2003.

References

Alliot, Maurice. <u>Le Culte D'Horus A Edfou Au Temps Des Ptolemees</u>. Institut Français
      D'Archaeologie Orientale Bibliotheque d'Etude, Tome XX, Premier Facsicule.
      Beyrouth, Liban: Libraire du Liban, 1949.

<u>Ancient Egypt. Discovering its Splendors</u>. Washington: The National Geographic
Society,1978.

Baud, Marcelle and Etienne Drioton. "Le Tombeau de Panehsy." <u>Tombes Thebaines</u>.
      Memoires Publies par les Membres de l'Institute Francais d'Archeologie Orientale
      du Caire 57. Cairo: Institut Français d'Archeologie Orientale, 1932.

Budge, E.A. Wallis, trans. <u>The Book of Opening the Mouth. The Egyptian Texts with
      English Translations</u>. 2 vols. London: Kegan Paul, Trench, Trübner & Co. Ltd.,
      1909.

Cauville, Sylvie. <u>Edfou</u>. Cairo: Les Guides archéologiques de l'Institut français du Caire
      vol.6. Cairo: Institut Français d'Archéologie Orientale, 1984.

Emery, Walter B. <u>Great Tombs of the First Dynasty. Excavations at Saqqara</u>. Cairo:
Govt.
      Press, 1949-1954.

Emery, W.B. A <u>Funerary Repast in an Egyptian Tomb of the Archaic Egypt</u>. Leiden:
      Nederlands Instituut Voor Het Nabije Oosten, 1962.

Fairman, H.W. "Worship and Festival in an Egyptian Temple." <u>Bulletin of the John
      Rylands Library</u> 37 (1954): 165-203.

Faltings, Dina. et. al. <u>MDAIK 2000</u>. Cairo: MDAIK, 2000.

Gardiner, Alan. <u>Egyptian Grammar. Being an Introduction to the Study of Hieroglyphs</u>.
3rd
      ed. London: Oxford University Press, 1964.

Garis Davies, N. <u>The Tomb of Vizier Ramose</u>. Mond Excavations at Thebes I. London:
      The Egypt Exploration Society, 1941.

Grimal, Nicolas. <u>A History of Ancient Egypt</u>. Oxford : Blackwell, 1994.

Kanawati, Naguib. <u>The Tomb and its Significance in Ancient Egypt</u>. Prism
Archaeological
      Series 3. Guizeh, Egypt: Foreign Cultural Information Department, 1987.

Lacau, P. and H. Chevrier. <u>Une Chapelle d'Sesotris Ier a Karnak</u>. Cairo: Institut Français
      d'Archeologie Orientale, 1956-1969.

Morgan J. de, et. al. <u>Recherches sur les origines de l'Egypte</u>. Paris: E. Leroux, 1896-
1897.

Moussa, Ahmed M. and Hartwig Altenmüller. <u>The Tomb of Nefer and Ka-Hay</u>. Verlag
      Phillip von Zabern, 1971.

Quibell, J.E. <u>Hierakonpolis I</u>. Egyptian Research Account, fourth memoir. London:
      Histories and Mysteries of Man, Ltd., 1989.

Petrie, W.M. Flinders. <u>Prehistoric Egypt ; Corpus of prehistoric pottery and</u>

palettes .Warminister, England: Aris & Phillips, 1974.

Petrie, W.M. Flinders. <u>Naqada and Ballas</u>. Warminister, Great Britain: Aris & Phillips Ltd.-Joel L. Malter, 1974

Chronological Table
(Adapted from Nicolas Grimal's *A History of Ancient Egypt*, 1992)

4500-3150 BC **Predynastic Period**

4500-4000 BC Badarian
4000-3500 BC Naqada I
3500-3300 BC Naqada II
3300-3150 BC Naqada III

3150-2700 BC **Thinite Period**

3150-2925 BC Dynasty I
2925-2700 BC Dynasty II

2700-2190 BC **Old Kingdom**

2700-2625 BC Dynasty III
2625-2510 BC Dynasty IV
2510-2460 BC Dynasty V
2460-2200 BC Dynasty VI

2200-2040 BC **First Intermediate Period**

2200-2160 BC Dynasties VII and VIII
2160-2040 BC Dynasties IX and X (Herakleopolis)
2160-2040 BC Dynasty XI (Thebes)

2040-1674 BC **Middle Kingdom**

2040-1991 BC Dynasty XI (all Egypt)
1991-1785 BC Dynasty XII
1785-1633 BC Dynasties XIII and XIV

1674-1553 BC **Second Intermediate Period**

1674-1633 BC Dynasty XIV
1674-1533 BC Dynasties XV and XVI (Hyksos)

1674-1533 BC Dynasties XVII (Thebes)

1552-1069 BC **New Kingdom**

1552-1314 or 1295 BC Dynasty XVIII
1295-1188 BC Dynasty XIX
1188-1069 BC Dynasty XX

1069-702 BC **Third Intermediate Period**

1069-945 BC Dynasty XXI
945-715 BC Dynasty XXII
825-715 BC Dynasty XXIII

747-525 BC **Late Period**

747-656 BC Dynasty XXV
747-672 BC Dynasty XXIV
672-525 BC Dynasty XXVI

525-404 BC **Dynasty XXVII** (First Persian Period)

404-343 BC **Dynasties XXVIII-XXX**

404-399 BC Dynasty XXVIII
399-380 BC Dynasty XXIX
380-343 BC Dynasty XXX

343-332 BC **Second Persian Period**

332 BC-AD 395 **Greco-Roman Period**

332-304 BC Macedonian Dynasty
304-30 BC Ptolemaic Period
30 BC-AD 395 Roman Period

Figures

Figure 1. King Sesostris I(18th c BC); his statue and his offerings to major gods on tables and stands at Karnak (photo: E.J. Walters).

R1, R2   W11, W12

Figure 2. Hieroglyphs representing temple furniture and vessels of stone and earthenware. R1: Table with loaves and jug; R2: table with slices of bread; W11: Ring-stand for jars, red earthenware pot (Dyn. XVIII form, round at bottom); W12: Ring-stand (O.K. form, straight at bottom) (Gardiner, 501, 529).
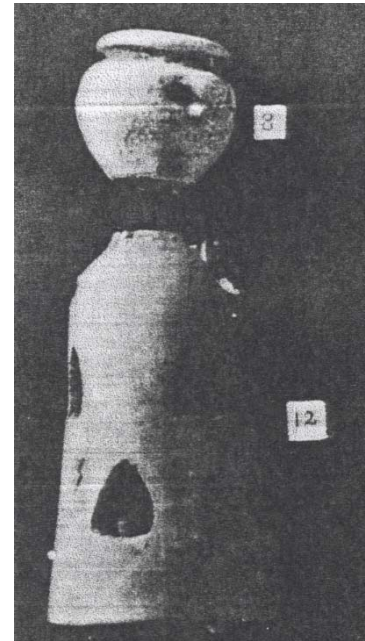
Figure 3. Dyn 0. Model vase (no.8) and model stand (no.12).Hierakonpolis,"Main Deposit" (Quibell, 7, pl. XVIII)
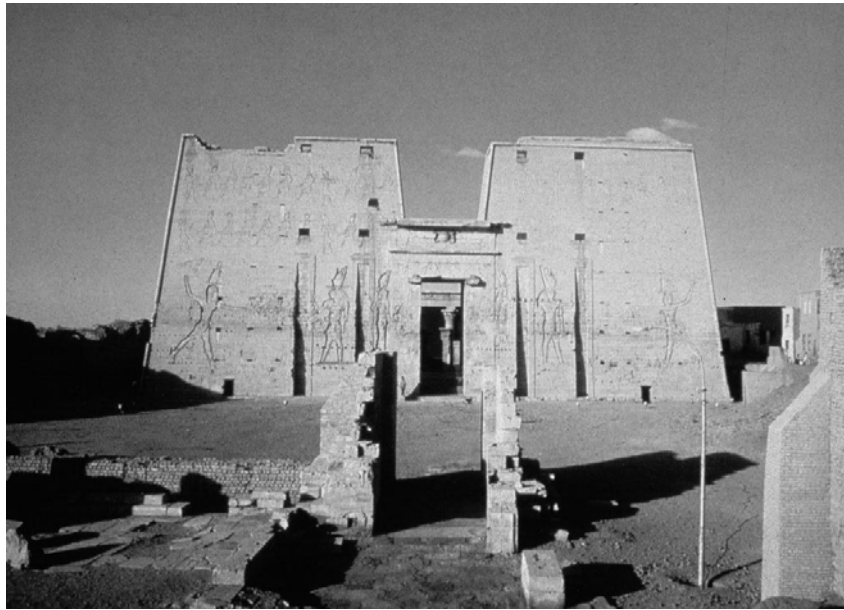
Figure 4. Temple of Edfu. Front view: pylon and main entrance (photo: E.J. Walters).
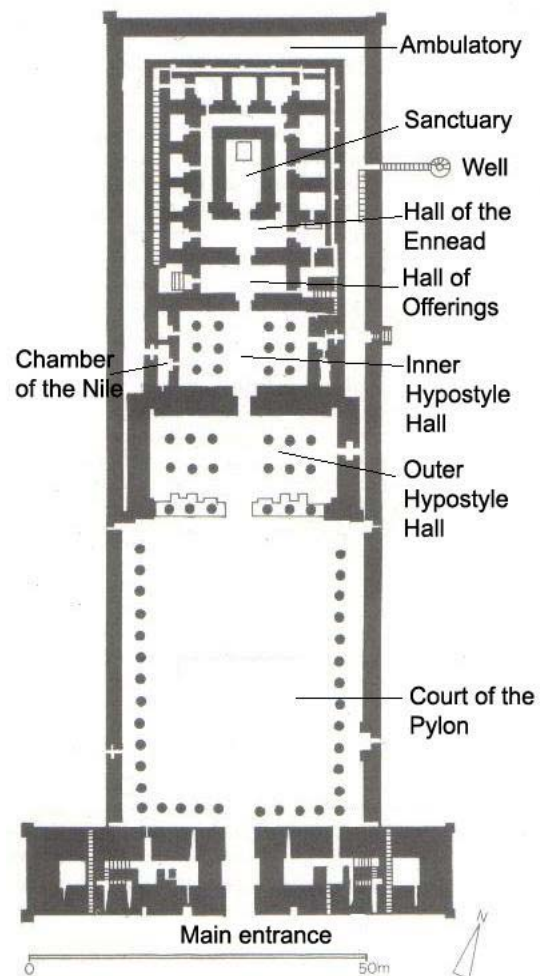
Figure 5. Plan of Temple of Edfu
(adapted from Fairman 1954)
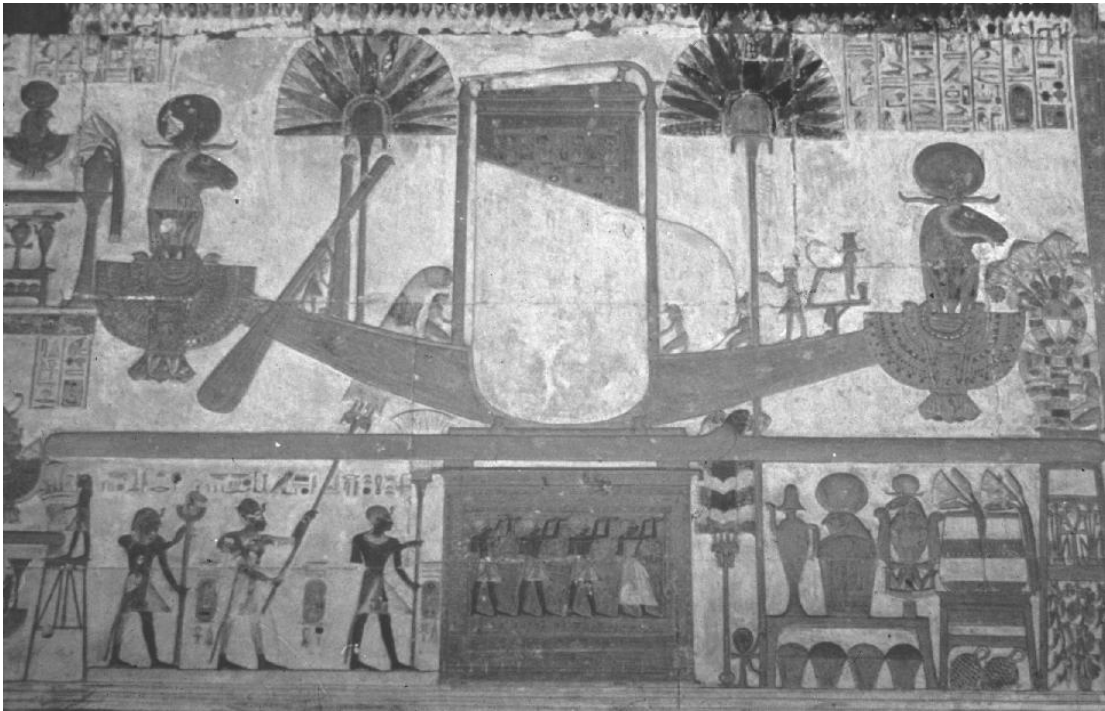
Figure 6. Relief painting in Sanctuary, Abydos, Temple of Seti I, 13th c BC (photo: E.J. Walters).

Figure 7. Holy of Holies of god Amon, Abydos, Temple of Seti I (detail, photo: E.J. Walters).
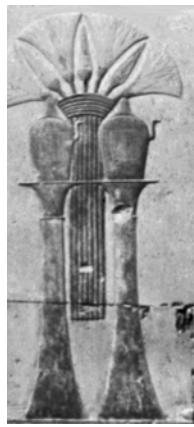
Figure 8. Holy of Holies of god Amon, Abydos, Temple of Seti I (detail, photo: E.J. Walters).
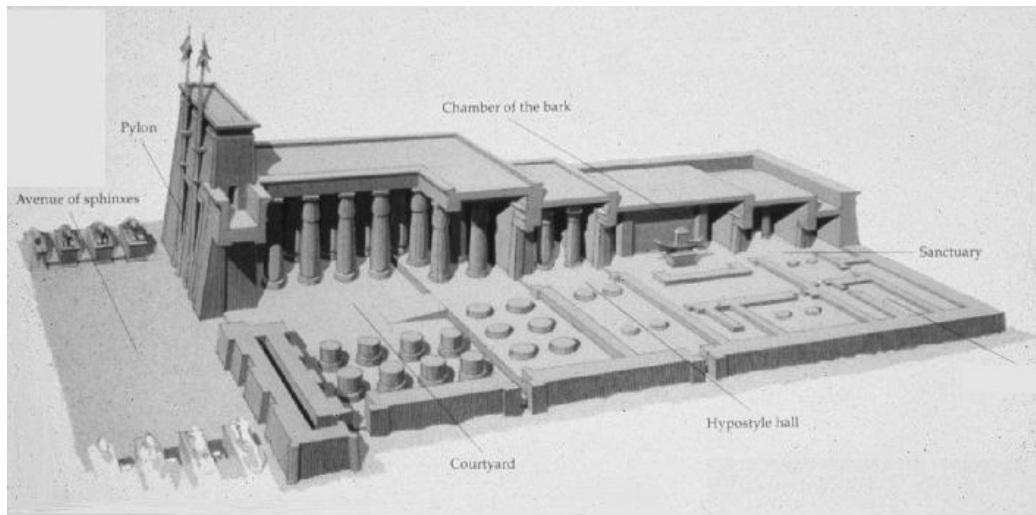
Figure 9. Cross section of the Temple at Karnak (Ancient Egypt, 167).

Figure 10. Wall painting, tomb of Tutankhamun, Valley of the Kings (Dyn. XVIII). King Ay as high priest on the right performing the "Opening of the Mouth" ceremony on the mummy of Tutankhamun (Ancient Egypt, 200-201).

Figure 11. Special vases from tomb no. 3503. Saqqara, Royal Tombs from Dyn. I (Emery, pl. LIII).

Figure12. Relief from Dynasty II tomb at Saqqara (Emery 1962, pl. 3A).
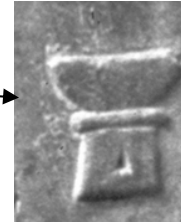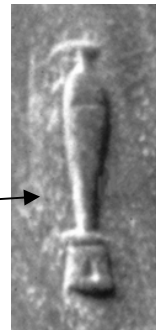
Figure 13. Hesy –Ra seated before his offering table
(photo: E.J. Walters).

Figure 15. Alabaster plate,
Cairo Museum (photo:
E.J. Walters).
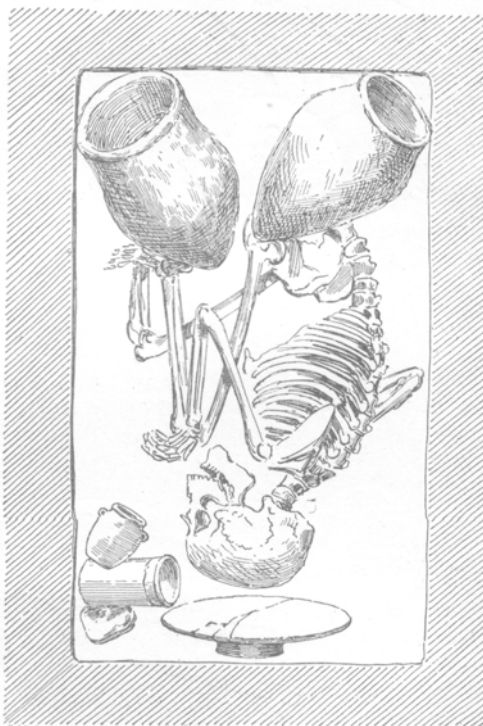
Sepulture de la necropole de Kawamil. J. de Morgan.
Ethnographie des Populations indigenes de l'Egypte. Fig, 466

Figure 14. Prehistoric grave from Naqada
(Morgan J. de, fig. 466).
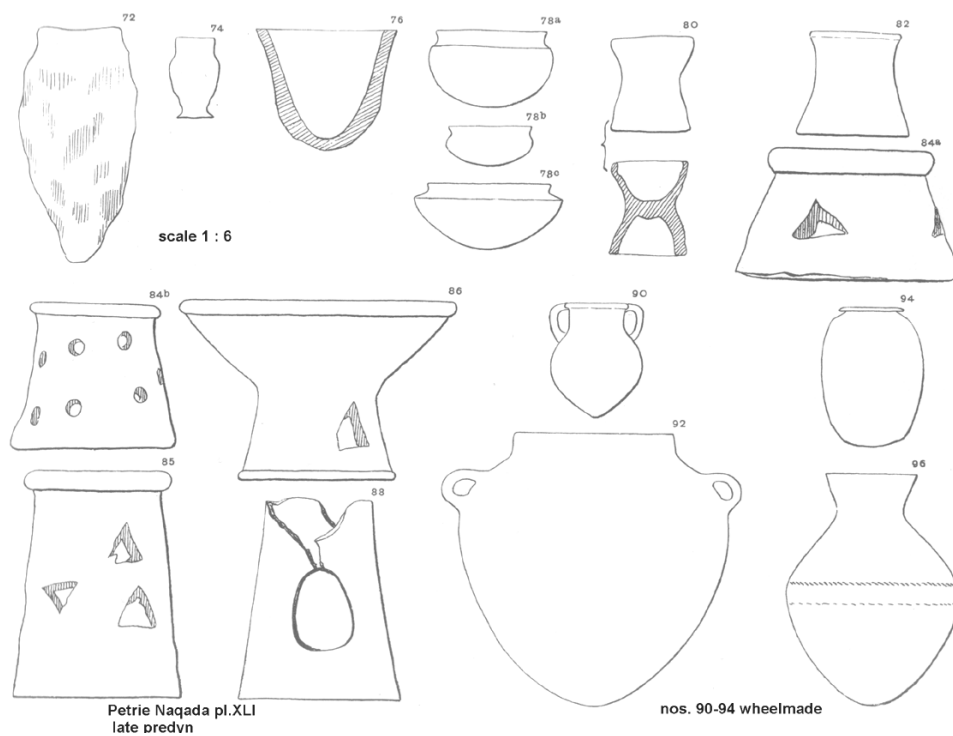
Figure 16. Predynastic potstands from Naqada (Petrie, pl. XLI).

Figure 17. Palette of Narmer, Hierakonpolis,
Dyn. 1 (photo: E.J.Walters).
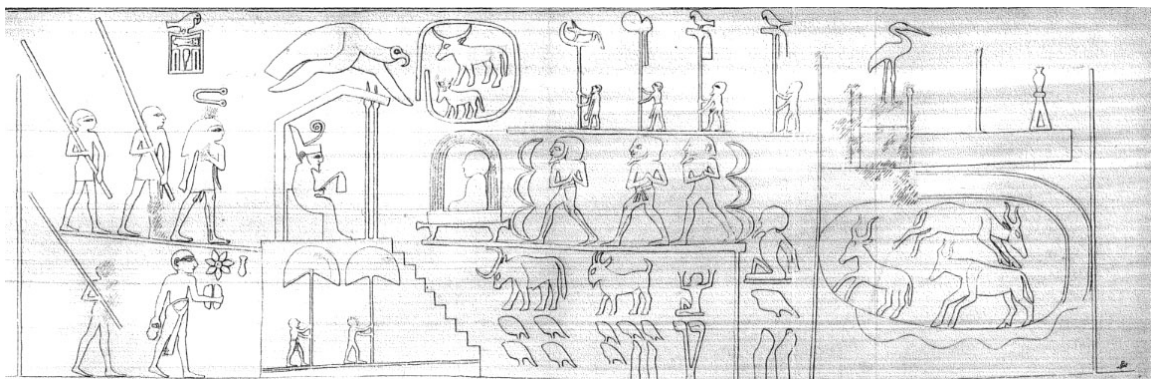
Figure 18. Drawing of carved scenes of Narmer's royal mace. Hierakonpolis (Quibell, pl. XXVI.B)
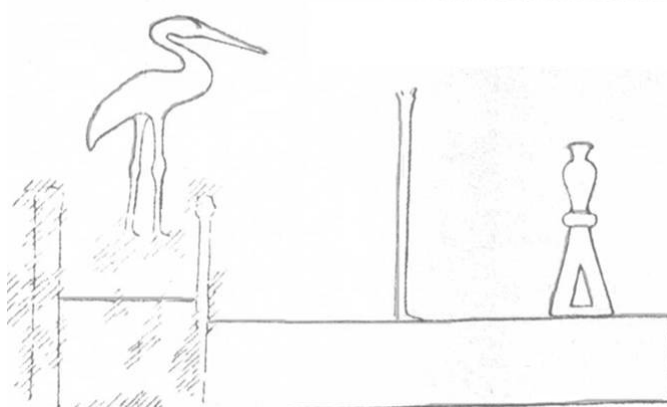
Figure 19. Offering before the Ibis god Thot; detail from royal mace of Narmer (Quibell, pl.XXVI)
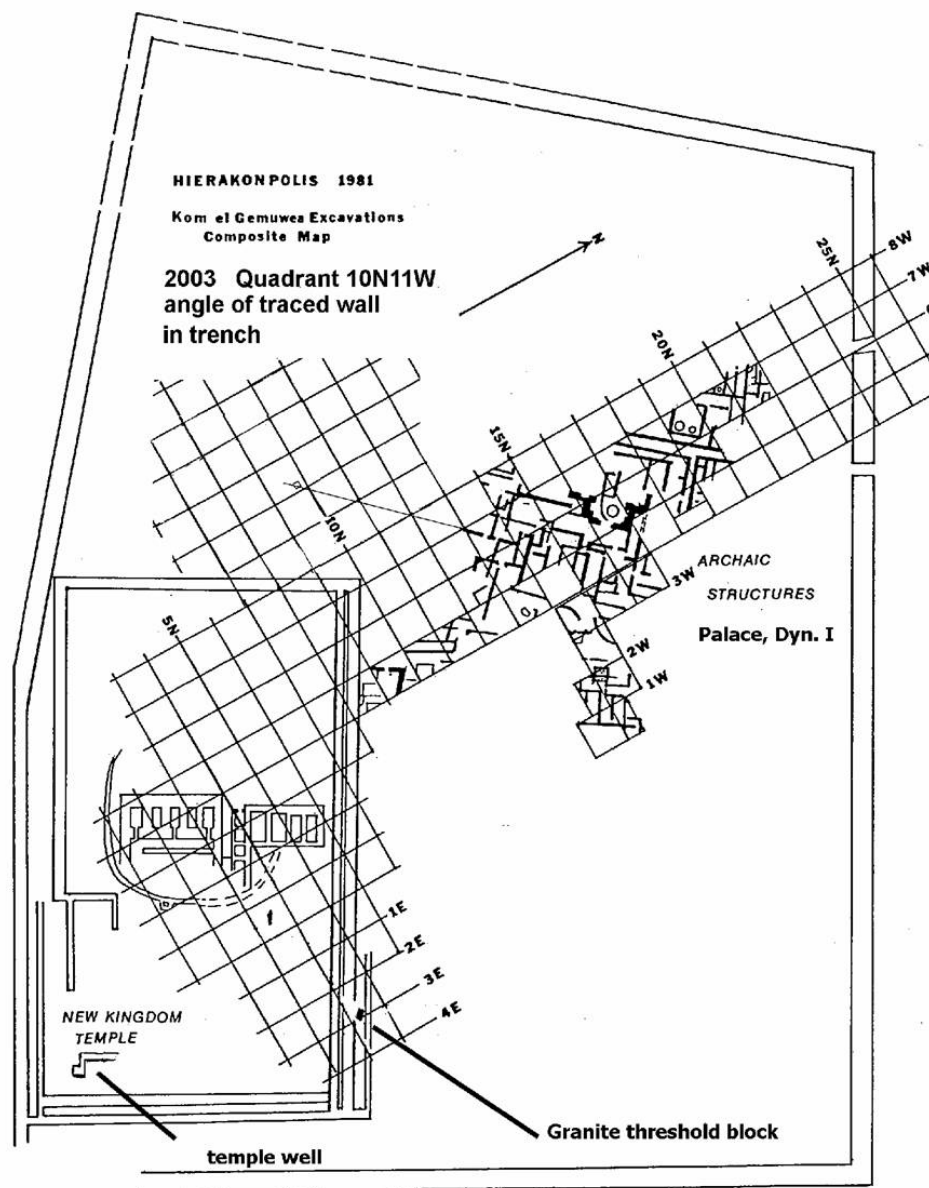
Figure 20. Map of Hierakonpolis from 1981 excavation.
Adapted by E.J.Walters for 2003 field season.
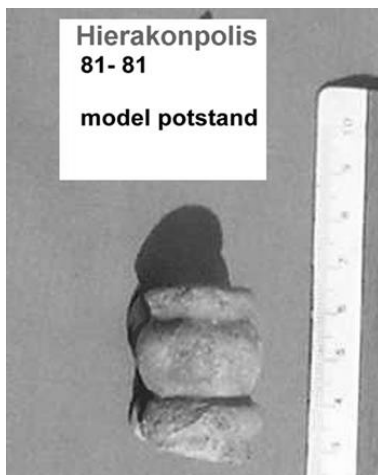
Figure 21. Model potstand from the 1981 field season (photo: E.J.Walters).
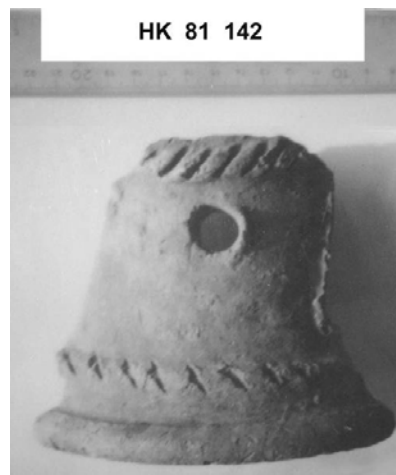
Figure 22. Fancy potstand fragment with 2 triangular and 1 circular piercings (photo: E.J. Walters).

Figure 23. Hierakonpolis Temple-Town project staff, Egyptian colleagues and village workers opening trench at 10N11W.

potstands from 10N 11W
(E.J. Walters)
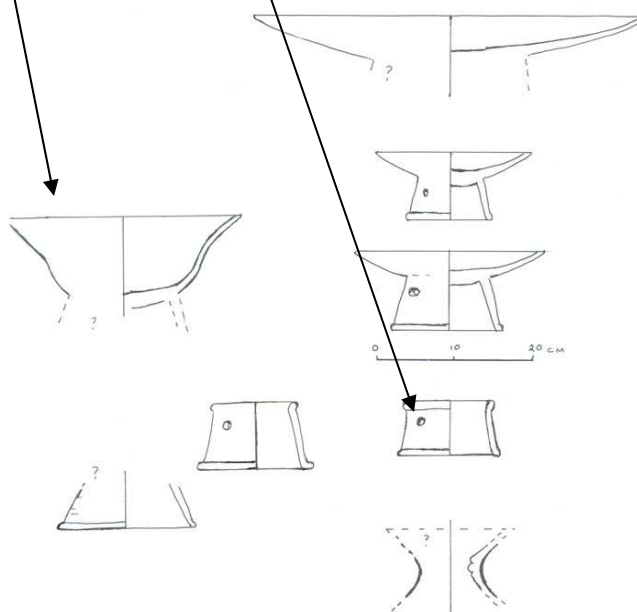
Figure 24. Close up of trench and drawings of potstands found (photos E.J. Walters).
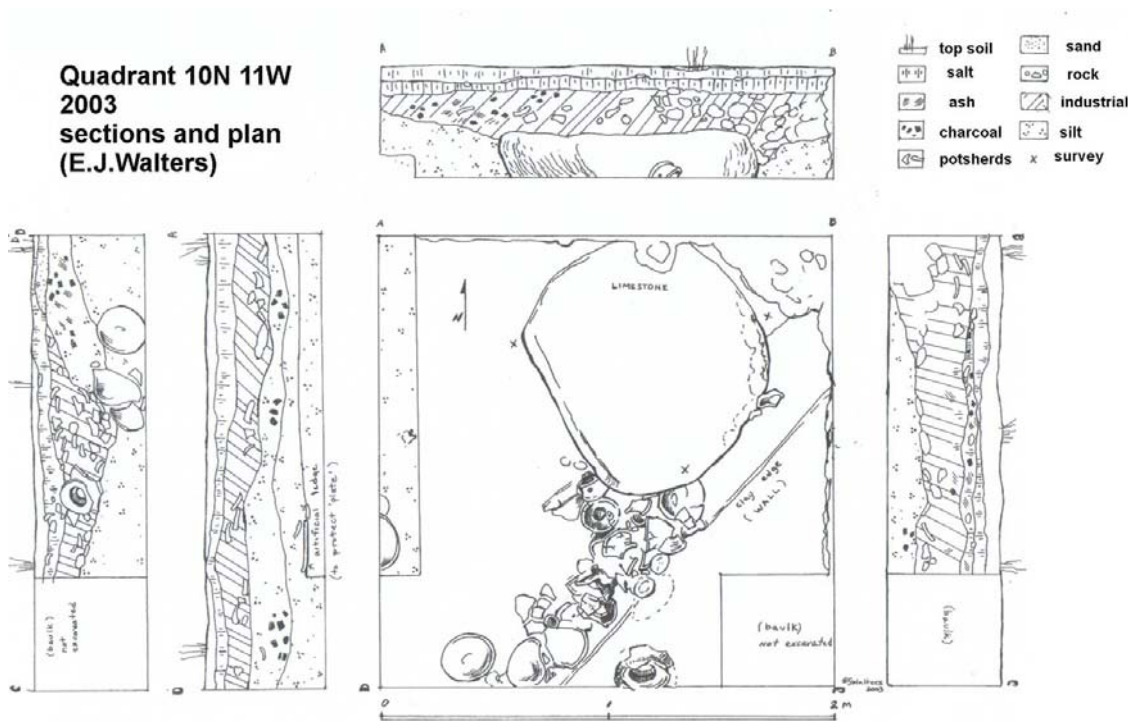
Figure 25. Drawing of trench with strata differentiation (photo E.J. Walters).
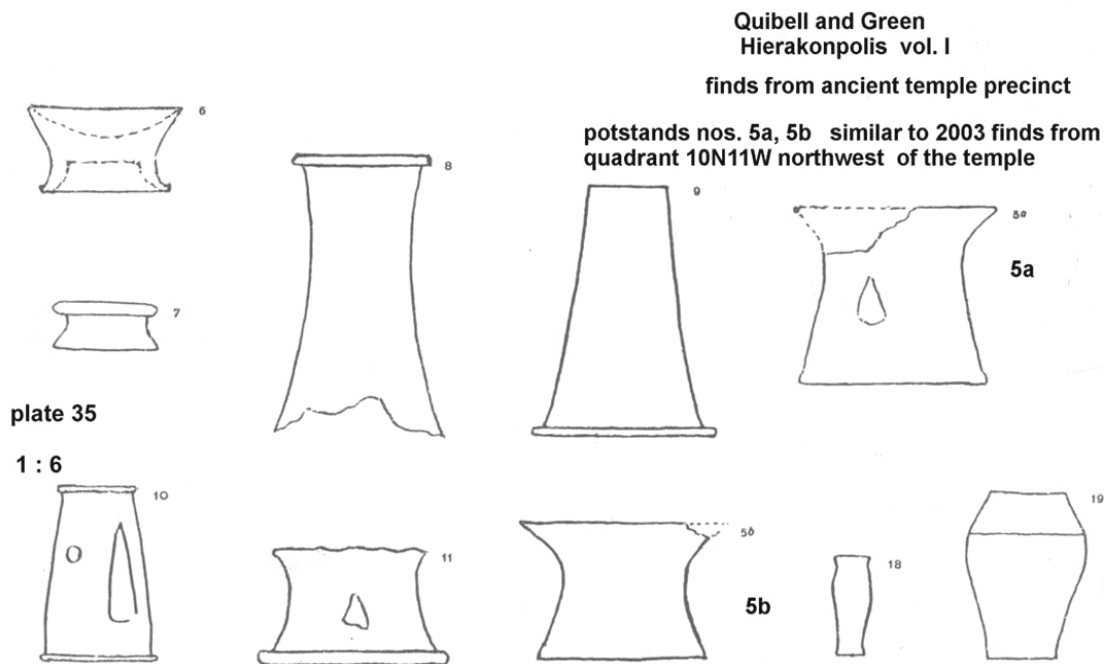
Figure 26. Potstands recorded by Quibell and Green from the Hierakonpolis Temple (Quibell, pl. XXXV).

Dina Faltings et al.
MDAIK 2000 v.56 p.153
Proto-Early Dyn.

10

13   14
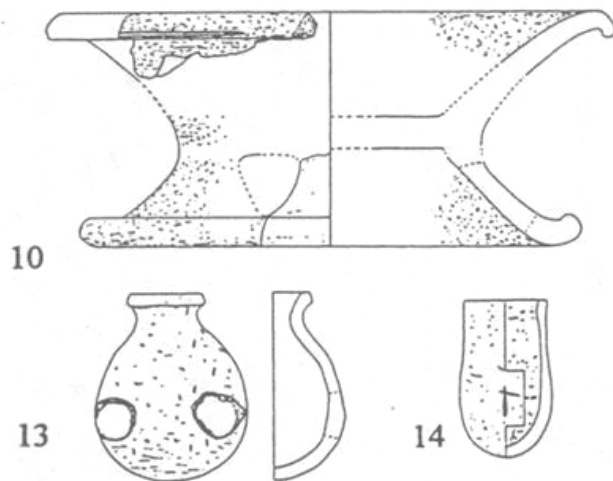
Abb. 8: Keramikformen (Auswahl) (10 und 16–19 M 1:8, sonst 1:4)

Figure 27. Recent potstand from the site of Buto
(Faltings, 153).
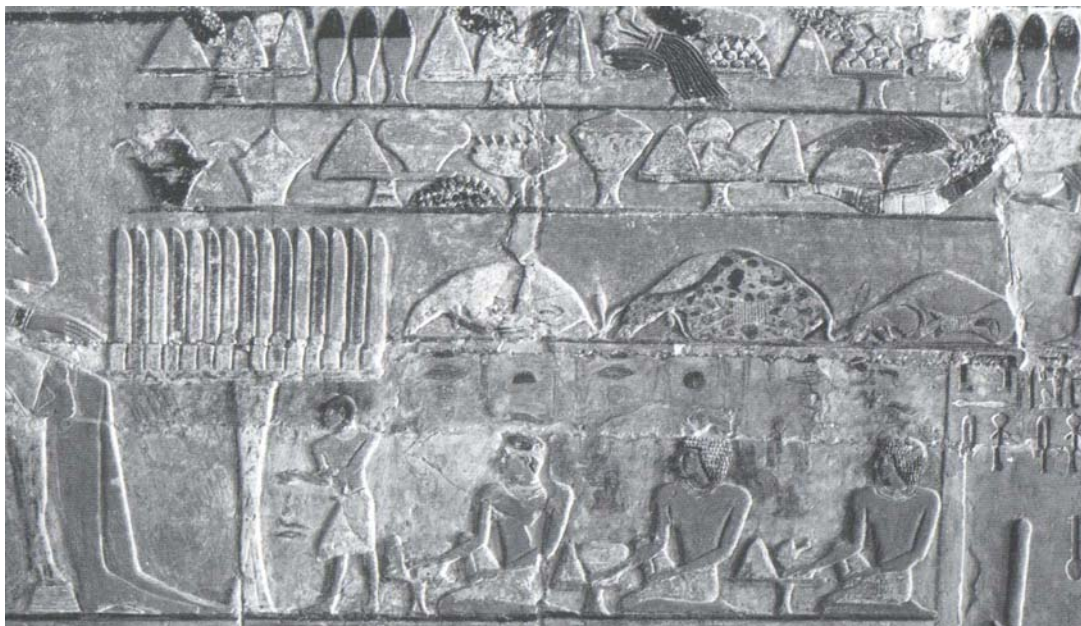
Figure 28. S.B. scene of tomb of Nefer and Ka-Hay (Moussa & Altenmüller, pl. 25).

Race Matters: Disparities in African-American Children with Attention Deficit Hyperactivity Disorder

Charmayne Maddox, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Keith B. Wilson, Associate Professor of Education
College of Education
Pennsylvania State University

Abstract

Attention Deficit Hyperactivity Disorder, ADHD, has become one of the most commonly diagnosed childhood behavioral disorders. African American children, specifically, have been found to be less likely to be treated for ADHD even after receiving a diagnosis when compared to their Caucasian counterparts. The purpose of this analysis of literature is to examine the relationship between race and healthcare disparities, as many researchers acknowledge an association between the two variables. This review of literature examines race and ethnicity, individual and institutional racism, as well as negative stereotypes in the healthcare system as possible explanations for disparities in ADHD.

Introduction

Attention Deficit Hyperactivity Disorder, or ADHD, has become one of the most commonly diagnosed childhood behavioral disorders. As defined by the American Psychiatric Association's Diagnostic and Statistical Manual of Mental Disorders-IV (DSM-IV-TR), ADHD is a disruptive behavior disorder characterized by persistent inattention and/or hyperactivity-impulsivity occurring in several settings repeatedly and more severely than is typical for individuals in the same age group. In the third addition of the DSM, the diagnostic criteria for what is now ADHD was presented in two variations: ADD + H (Attention Deficit Disorder with Hyperactivity) and ADD-H (Attention Deficit Disorder without Hyperactivity). Although given three categories of symptoms: (1) inattention, (2) impulsivity, and (3) hyperactivity, the American Psychiatric Association (APA) believed that the DSM-III criteria were inadequate because the criteria did not distinguish between attention deficit disorders and conduct disorders (Jordan, 1998). In 1994, the APA presented new guidelines for labeling an individual as ADHD. ADD (+ or – H) and ADHD are not to be used interchangeably, as the DSM-IV provides a new definition and a new set of diagnostic criteria for the disorder. The fundamental characteristics of the ADHD diagnosis remain to be inattention, hyperactivity, and impulsivity. The following is a list of symptoms that characterize inattention, hyperactivity, and impulsivity (APA, 2000):

<u>DSV-IV Criteria for ADHD</u>
Inattention
- Often fails to pay close attention to details or makes imprudent mistakes in schoolwork, work, or other activities.
- Often has difficulty sustaining attention to tasks or play activities.
- Often does not seem to listen when spoken to directly.
- Often does not follow instructions and fails to complete schoolwork, chores, or duties in workplace (not due to oppositional behavior or failure to comprehend instructions).
- Often has difficulty organizing activities.
- Often avoids, dislikes, or is reluctant to engage in activities that require sustained mental effort (e.g. schoolwork or homework).
- Often misplaces items necessary for tasks and activities (e.g. toys, school assignments, pencils, books, or tools).
- Is often easily distracted.
- Is often forgetful in daily activities.

Hyperactivity/Impulsivity (1) Hyperactivity
- Often fidgets with hands or feet or squirms in seat.
- Gets out of seat in situations in which remaining in seat is expected.
- Often runs about or climbs when and where it is not appropriate (adolescents or adults may be limited to feelings of restlessness).
- Is often "on the go" and often acts as if "driven by a motor".
- Often talks excessively.

(2) Impulsivity
- Often blurts out answers before questions have been completed.
- Often has difficulty waiting ones' turn.
- Often interrupts or intrudes on others (e.g. butts into conversations or games).

In most mainstream school classrooms, children may display some of these symptoms, as children can sometimes be inattentive, impulsive or overactive. While under pressure, bored, or tired, most people will not be able to concentrate and can become easily distracted. Characteristics of ADHD that are temporary, episodic, and directly associated with situational factors should not be mistaken as the behavior disorder (Cooper & O'Regan, 2001). Based on the criteria previously listed, three types of ADHD are identified: ADHD Combined Type, ADHD Predominately Inattentive Type, and ADHD Predominately Hyperactive-Impulsive Type. An ADHD diagnosis is usually determined by the pervasiveness of the symptoms. For an ADHD diagnosis to be made, the following conditions must be met as listed in the DSM-IV-TR:

- The child must display six or more of the nine symptoms listed in the DSM-IV-TR for both inattention and/or hyperactivity/impulsivity to meet the criteria for one of the three types of ADHD.
- Some evidence of ADHD symptoms must have been observable prior to age seven.
- The child's symptoms must have persisted for at least six months.

- ADHD symptoms must be present in two or more settings (i.e. school, work, and/or at home).
- There must be clear evidence of significant impairment within the social, school, or work setting.
- ADHD symptoms must not occur only during the course of a Pervasive Developmental Disorder, Schizophrenia, or other Psychotic Disorder, and are not better accounted for by another mental disorder (i.e. Mood Disorder, Anxiety Disorder, Dissociative Disorder, or a Personality Disorder)

ADHD currently affects three to seven percent of children or an estimated two million children nationwide. On average, at least one child in every classroom in the United States needs clinical assistance with ADHD. Boys are diagnosed with ADHD two to three more times than girls (Matthews, 2002). One author estimated the prevalence of ADHD as high as ten percent or a total of three million children between five and 12 years of age (Millichap, 1998). ADHD often continues into adolescence and adulthood, which can lead to medication dependency and a lifetime of treatment.

Treating children with ADHD requires medical, educational, behavioral, and psychological treatment. This comprehensive approach is called multimodal and includes parent training, behavioral intervention strategies, an appropriate educational program, education regarding ADHD, individual and family counseling, and medication when required (Matthews, 2002). Treatment plans are tailored to the specific need of the child and his/her family. For most children, medication is a vital part of treatment. Medication is used to improve ADHD symptoms, allowing the child to function more effectively. Psychostimulants are the most commonly used medication for the treatment of ADHD. Common psychostimulants include methylphenidate (Ritatlin), Adderall, and dexotroamphetamine (Dexedrine, Dextrosat). The most common side effects of psychostimulants are reduction in appetite and difficulty sleeping. Some children experience stimulant rebound—a negative mood or an increase in activity when the medication loses it effect (Matthews, 2002).

The cause of ADHD is still unknown, although researchers have examined several theories. In previous years, health professionals adopted the notion that ADHD stemmed from the home and environment. Researchers hypothesized that ADHD was the result of watching too much television, food allergies, excess sugar intake, and poor home or school life. Conversely, scientists at the National Institute of Mental Health (NIMH) currently suggest that ADHD has a biological basis. Scientists at the NIMH found a link between a person's ability to pay continued attention and his/her level of brain activity. Researchers measured the level of glucose used by the areas of the brain that inhibit impulses and control attention, as glucose is the brain's main source of energy giving a good indication of the brain's level of activity. In persons diagnosed with ADHD, the brain areas that control attention used less glucose, indicating less activity. Researchers suggest that lower levels of brain activity may cause inattention (NIMH, 1996). There is also consistent evidence that ADHD is heritable.

There is a higher rate of concurrent and past ADHD symptoms in immediate family members of children with ADHD relative to their non-ADHD counterparts (Faraone et al., 1993). The most practical explanation regarding the etiology of ADHD is that various neurobiological factors may predispose children to exhibiting higher rates of impulsivity along with shorter than average attention spans compared to other children (DuPaul & Stoner, 2003). While several researchers have focused on the causes of ADHD, many others have examined the treatment of the disorder, the relationship between ADHD and learning disabilities, special education, and the use of healthcare services. Unfortunately, most of the studies conducted on ADHD are not generalizable to an entire population, as most of the research participants have been Caucasian males. African American children have accounted for a small number of participants in several research studies on ADHD.

Data reported from a national health survey conducted by researchers with the Center for Disease Control and Prevention (CDC) estimated the prevalence of diagnosed Attention Deficit Disorder (ADD) and/or Learning Disorder (LD) in U.S. children. In 1997-98, over 2.6 million children aged six to 11 were reported to have an ADD or LD diagnosis. Three percent of children aged six to 11 had been diagnosed with only ADD, four percent with only LD, and four percent with both conditions. White non-Hispanic children were diagnosed more often with ADD than black non-Hispanic children. Caucasian non-Hispanic children represented over 65% of the participants surveyed.

Addressing various limitations to the current research done on ADHD, the review of literature here is set to examine several disparities in the diagnosis of ADHD. As noted earlier, most of the studies conducted on ADHD focus on Caucasian males. The analysis here will examine the prevalence of ADHD among African American children, as many African American children are currently faced with this disorder. Secondly, several researchers suggest that although African American children are diagnosed with ADHD, many fail to receive adequate treatment for the disorder, as their needs for services are left unmet (e.g. Samuel et al., 1998; Bussing et al., 2003). Others have noted racial disparities in the use of prescription medication and primary healthcare (e.g. Hahn, 1995; Zito et al., 1998; Zima et al., 1999). While several of these authors have discussed financial barriers (e.g. socioeconomic status, poverty) and inadequate health insurance as explanations for disparities in ADHD among African Americans, this review of literature is set out to examine race, racism, and discriminatory practices as the underlying reasons behind the inadequate treatment of ADHD for African American children. The purpose of this paper is to answer the following research question: Is there a relationship between race and healthcare disparities in Attention Deficit Hyperactivity Disorder in children? I hypothesize that race, racism, and discriminatory practices will dictate which group is given adequate treatment for ADHD. It is my inference that due to discriminatory practices and racist ideology within the healthcare system, African American children will be less likely to receive treatment for ADHD when compared to their Caucasian counterparts.

Disparities in ADHD: Review of Literature

In contrast to the plethora of research on ADHD in Caucasian children, there is limited information about ADHD among African American children. To further explore this inadequacy, Samuel et al. (1998) interviewed 19 African American children with DSM-III-R ADHD and 24 African American children without ADHD. Interviewers conducted a psychiatric assessment of the participants using the Schedule for Affective Disorders and Schizophrenia for School-Age Children—Epidemiologic Version (5), as well as material based on the DSM-III-R. Compared with children who did not have ADHD, African American children with ADHD had higher levels of psychiatric disorders other than ADHD (e.g. disruptive disorders, mood disorders, anxiety disorders, substance disorders). The findings of this study were compared with an earlier study of Caucasian children with ADHD. The comorbidity of ADHD with other disruptive behavior disorders has been associated with poor prognosis, delinquency, and substance abuse in Caucasian children. These preliminary findings suggest that the currently accepted definition of ADHD identifies a disorder with similar—but not identical—psychiatric correlates to those previously identified in Caucasians (Samuel et al., 1998). The results of this study should be interpreted with caution, as the number of participants was relatively small and it was difficult to detect group differences. We can infer that as compared to Caucasian children, the comorbidity of ADHD in African American children may have an association with poor prognosis and insufficient treatment. Several other epidemiological studies have shown that African American children with ADHD and/or psychiatric disorders remain untreated.

Cuffe et al. (1995) examined race and gender differences in the treatment of adolescent psychiatric disorders. Data was collected on 478 adolescents during a two-stage, school-based, epidemiological study of depression using the Schedule for Affective Disorders and Schizophrenia for School-Age Children and the Children's Global Assessment Scale (K-SADS). Items covered in the K-SADS included the diagnostic criteria for affective, schizophrenic, anxiety, phobic, and conduct, and eating disorders during the past 12 months as defined by the DSM-III. The diagnostic evaluation consisted of a semi-structured interview of the adolescent and his or her mother. Symptom information was obtained independently from the adolescent and the parent to arrive at summary diagnoses. Diagnoses were grouped into three categories: affective disorders, non-affective disorders, and affective disorders comorbid with nonaffective disorders. Affective disorders included major depressive disorder, bipolar disorder, mania, hypomania, and schizoaffective disorder. Non-affective disorders included schizophrenia, generalized anxiety, separation anxiety, phobia, panic disorder, obsessive-compulsive disorder, conduct disorder, anorexia, and bulimia. African Americans and Caucasians had similar prevalence rates of an affective disorder and an affective disorder comorbid with a nonaffective disorder. Caucasians had a higher prevalence of having a nonaffective disorder than did African Americans. Caucasian males were more likely to receive outpatient treatment than any other race-gender group. Caucasians were also more likely to receive outpatient treatment than African Americans in all disorder categories. In addition, the frequency of receiving more than two weeks of treatment was 13 and six percent for Caucasians and African Americans, respectively. Increased risk

for under treatment was found for African Americans. African American subjects reported higher levels of symptomatology on the Center for Epidemiologic Studies Depression Scale (CES-D). After controlling for CES-D scores, African American males received treatment at half the rate of Caucasian males. After controlling for demographic variables (e.g. race, gender, socioeconomic status), Cuffe et al. (1995) found that low socioeconomic status was not a significant factor for the underutilization of treatment services. This finding suggests that income is not a barrier to receiving care in this sample. There may be referral bias against African Americans (Cuffe et al. 1995).

To further explore disparities in the treatment of ADHD, Bussing, Zima, Gary, and Garvan (2003) conducted a study to identify barriers to detection, help seeking, and treatment. For help-seeking analysis, 389 children were chosen who were considered high risk for ADHD according to scores on the Swanson-Nolan-and Pelman-IV (SNAP-IV), or had a previous diagnosis for ADHD, and/or were currently under treated. For the barriers to care analysis, 91 children were selected who met the DSM-IV criteria for ADHD identified as having unmet service needs for ADHD care in the past year. Caucasian children were more than twice as likely to receive an evaluation, be diagnosed, or be under current treatment for ADHD than African American children. Fifty percent of the unserved children with ADHD were African American. Seventy percent of the unserved children received subsidized school lunch. ADHD treatment was higher for non-poor children than for their impoverished peers. ADHD treatment was also higher for students receiving regular or gifted education services compared to children receiving special education services. After controlling for enabling and need variables (e.g. health insurance status, SNAP-IV scores), Caucasian children were more likely to receive an ADHD evaluation than were African American children. Barriers to care included not being sure as to where to go for help, system barriers (e.g. could not get an appointment), and financial barriers. African American parents had higher rates of negative treatment expectations than did Caucasians. This may be a reflection of the racial disparities in the quality of care, including mental health treatment (Bussing et al., 2003). Several other studies indicate that minority children are significantly less likely to receive ADHD treatment, including psychotropic medications, than Caucasians.

A retrospective investigation was conducted to analyze racial disparities in psychotropic medication prescribed to African American and Caucasian children with Medicaid insurance in Maryland. Zito, Safer, Dosreis, and Riddle (1998) analyzed (1) quantitative estimates of the medication prevalence (defined as the proportion of eligible recipients with any prescription claims) of the leading types of psychotropic and non-psychotropic medication classes in relation to race (Caucasian opposed to African American); the relationship between race and enrollment status (continuous versus non-continuous) for the study year; and (3) the relationship between race and region (county of residence) for psychotropic and non-psychotropic medication classes. Psychotropic classes included psychostimulants, antidepressants, antipsychotics, benzodiazepines, and lithium. Non-psychotropic classes included antibiotics, skin preparations, antitussives, antihistamines, and ear, eye, nose, and throat preparations. The sample consisted of 99,217 African American and Caucasian children aged five to 12 years of age that were enrolled and eligible to receive health services during the study year. Caucasian children

with Medicaid insurance were more likely to have prescription claims for the five classes of psychotropic medications than African American children. Caucasian children were twice as likely to receive a psychotropic prescription compared with African American children. After adjusting for geographic region, Caucasians remained twice as likely to receive psychotropic prescriptions compared to African Americans. Within the stimulant pharmacological class, methylphenidate was the most frequently prescribed medication, accounting for 89.6% of the stimulant use. For methylphenidate, there was a 2.5-fold lower use among African Americans than Caucasians. Moreover, these findings suggest that there are racial differences in the use of psychotropic medications despite having comprehensive, cost-free access to medical services and medications through the Medicaid insurance program. Explanations for unequal access to healthcare and medication include questions of quality and discriminatory practices among healthcare providers (Zito et al., 1998).

Zima, Bussing, Crecelius, Kaufman, and Berlin (1999) conducted a study to also investigate the use of psychotropic medication use among school-aged children. Two phone interviews were conducted among foster parents to determine the levels of psychotropic medication use among their school-aged foster children and how these levels relate to severe psychiatric disorders for which medication use is a fundamental component of treatment. In addition, the study also examined potential predictors (e.g. socio-demographic characteristics, placement history) of receiving medication treatment. Psychotropic medication use was assessed from the foster parent reports. Of the 472 children randomly selected for the study, 82% of the children were from minority backgrounds. Sixteen percent of the subjects were reported to have ever taken psychotropic medication, and most of these children (89%) had received treatment in the previous year. Stimulants were the most common class of medication taken in the previous year. Fifty-two percent of the children whose clinical status merited medication evaluation had not received any psychotropic medication in the previous year. Children with ADHD were more likely than those without the disorder to have taken each of the classes of psychotropic medication in the previous year. However, almost 49% of the children in this study with ADHD had not received any psychotropic medication in the previous year. These findings are consistent with Zito et al. (1998) in that disparities in the treatment of ADHD persist among minority populations, as the majority of participants in this study were from minority backgrounds. Researchers have consistently found racial differences in the use of services and prescription medication.

Using data from the 1987 National Medical Expenditure, Hahn (1995) examined the relationship between race, ethnicity, physician visits and the use of prescription medication for two samples of children: 1,347 children aged one to five and 2,155 children aged six to 17. Of the children with at least one physician visit, black and Hispanic children in both age cohorts averaged one fewer physician visit compared with white children. The percent of children whom their mothers reported as in fair or poor health for both groups of minority children was over 2.5 times higher than that of white children. Hispanic and black children of any age had lower incomes compared with whites. Nearly half of minority children were in families with an income at or near the poverty level compared with 17% of white children. Black and Hispanic children were

significantly more likely to be uninsured compared with white children. Young minority children were also less likely to have a regular source of care compared with white children and were more likely to use hospital based treatment, such as emergency rooms, as their usual source of medical care. After controlling for predisposing and enabling variables, black children were approximately half as likely to receive a prescription medication compared with white children. Predisposing variables included the child's age and education of the child's mother. Enabling variables consisted of the type of health insurance coverage, poverty status, usual source of medical care, and geographic location. Despite the age group, whether or not predisposing and enabling variables or number of physician visits were included, white children had the highest use of prescription medication and black children had the lowest. These findings suggest that the relationship between racial and ethnic status and the use of prescription affirms that minorities receive fewer services than whites. Differences may be due to forms of discrimination in treatment of minority patients compared with white patients (Hahn, 1995). Other researchers have demonstrated that disparities in access to basic health care services have been a result of race, income, and insurance status.

Newacheck, Hughes, and Stoddard (1996) examined the use and access patterns to primary health care for four groups of children: (1) children in families with incomes below the federal poverty line; (2) children representing minority ethnic or racial groups; (3) children without health insurance; and (4) children exhibiting none of the above characteristics. Using the National Medical Expenditure Survey (NMES), a national household survey conducted for the Agency for Health Care Policy and Research, Newacheck et al. (1996) sampled 7,578 children aged one to 17 to determine access to healthcare. Children without any health insurance were least likely to have usual sources of health insurance. These children were more than twice as likely than children from white, non-poor, and insured families (the reference group) to report not having usual sources of care. Children from poor families and minority children were also less likely than white, non-poor, and insured children to have usual sources of care. Minority children and those from poor families reported having usual sources of care approximately 11% less often than children in the reference group. Although not an exhaustive list, usual sources of care included physician's office or group practice physician's clinic, company industrial clinic, school clinic, family health center, hospital outpatient clinic, hospital emergency department, walk-in center and/or patient's home. The children in the reference group were identified as using the physician's offices as their location for usual care approximately one third more times than minority children and children from poor families. Children from poor families were nearly nine times as likely to identify neighborhood or family health centers as their usual source of care compared to the reference group. Children from poor families, minority children, and uninsured children were all at increased risk of having no access to after-hours emergency care. Children from white, non-poor, insured families had significantly higher rates of annual physician visits per 100 bed days than children from the three at-risk groups. These results indicate that minority children, children who live in poverty, or are uninsured are at a much greater risk than their white, non-poor, insured counterparts to experience barriers in access to primary healthcare (Newacheck et al.,

1996). These results also indicate that minority and poor children have difficulties obtaining primary care even when insurance status was statistically held constant.

To further explore disparities in the diagnosis and treatment of ADHD, several researchers have examined behavior rating scales. Behavior rating scales are one of the most commonly used methods in the assessment of ADHD (Reid, Casat, Norton, Anastopoulos, & Temple, 2001). Concerns over the possibility of disproportionate diagnosis of ADHD among African American children, researchers have examined the validity and reliability of behavior rating scales.

Assessment of ADHD

In a study of 3,998 (2,124 African American and 1,874 European American) elementary school children aged five to 11, Reid et al. (2001) examined the normative and construct equivalence of the teacher IOWA Conners Rating Scale (IOWA). IO referred to Inattention/Overactivity and WA referred to Aggression. African American children screened positive for IO and WA at a much higher rate than the European American group. There were significant main effects for ethnicity, as the scores on the IO and WA subscales were higher for African American children compared with European American children. Differences were found in the distributions of IOWA scores and means across African American and European American children, which lead to an increased likelihood of African American children screening positive for IO and WA. On average, African American boys were approximately 2.5 times more likely to screen positive for IO and WA. African American girls were 3.5 times more likely to screen positive for IO and WA. All of the participants were selected from schools with a high proportion of at risk, low socioeconomic status children. There were no pronounced differences in the socioeconomic statuses of the participants, therefore socioeconomic status alone could not account for the observed differences (Reid et al., 2001). For both the IO and WA subscales, teachers rated the African American students higher than the European American students. More specifically, African American girls were much more likely to screen positive for IO and WA when rated by European American teachers. Disparities in Reid et al. (2001) study adds the IOWA to the list of behavior rating scales that have documented significantly higher scores for African American children when compared to their European American counterparts.

Like the IOWA behavior-rating scale, The Conners Teacher Rating Scale has also documented racial differences on ratings of hyperactivity and impulsivity. The CTRS is a commonly employed rating scale used to assess classroom behavior problems related to ADHD (Epstein, March, Conners, & Jackson, 1998). In a study of 1,027 children aged 10 to 16 years old, Epstein et al. (1998) examined mean differences on the CTRS between African American and Caucasian children. For both male and females, teachers tended to rate African American children higher than Caucasian children on factors relating to externalizing behavior (e.g. conduct problems, hyperactivity). Of the 39 items on the CTRS, the most profound differences between the races were found on the Conduct Problems and Hyperactivity factors, with African American males scoring significantly higher than Caucasian males. Effect sizes for these differences were

moderate suggesting that there is a tendency for teachers to rate African American children higher on CTRS items that reflect externalizing behaviors (Epstein et al., 1998). African American females were rated significantly higher on the Conduct Problems factor than Caucasian females. African American females were also rated higher on Anxious/Passive factors than their Caucasian counterparts. These findings suggest that there is a possible teacher bias on ratings of hyperactivity for African American children.

To further explore the assessment of culturally different students for ADHD, Reid et al. (1998) examined the cross-cultural equivalence of the ADHD-IV Rating Scale School Version. Three hundred African American and 1,359 Caucasian public school students aged five to 18 years old were selected for the study. Teachers were asked to rate the behavior of two randomly selected students from their class roster. The mean scores for the African American group were significantly higher than the Caucasian group for both Hyperactivity-Impulsivity (HI) and Inattention (IA). Moreover, there were significant differences in group variances and distinctly different distributions across both racial groups for HI and IA factors. These findings suggest that if Caucasian norms were used for African American students, approximately twice the number of African American students would screen positive on the HI and/or IA factors. The norms for the Caucasian group may not be appropriate for the African American group. The ADHD-IV scale may not satisfy the conceptual equivalence requirement across groups. This suggested that at least some of the observed group differences are due to variations in the performance of the scale across groups as opposed to differences in the actual behavior exhibited by the participants (Reid et al., 1998).

As research on behavior-rating scales became more popular in ADHD literature, growing evidence suggested that African American children consistently have higher scores on inattention and hyperactivity measurements compared with Caucasian children. Based on a review of research studies using ADHD behavior rating scales with culturally different groups, Reid (1995) made the following conclusions: (1) insufficient data existed to determine the extent to which psychometric properties of rating scales were consistent across different groups; (2) evidence suggested that culturally different individuals may be over identified; (3) culturally different individuals were not adequately represented in the norm groups of many of the available scales; and (4) the possibility of rater bias could exist when individuals from one cultural group rate children from a different cultural group. Racial biases could be inherent in behavior-rating instruments. Behavior-rating scales and symptom checklists may not be equivalent across ethnicities (Cuffe, Waller, Cuccaro, Pumariega & Garrison, 1995). Both ADHD and behavior-rating scales were derived from the perspective of Western professionals, using Western concepts of disorder and measurement, and without regard to cultural differences (Reid et al., 1998).

Although several studies have investigated behavior-rating scales and the assessment of ADHD, few studies have examined the perceptions of educators and other professional staff who work with ADHD children. Davison and Ford (2002) interviewed 25 participants consisting of African American and White educators, medical personnel, and social workers/counselors who work with parents of children attending four inner

city schools within a large African American population. The interview questions focused on the perceptions of individuals working with children with ADHD within the school, home, and medical setting. African American and those interacting with African American parents expressed a socially constructed view of ADHD and were less likely to accept a biological determinist point of view as opposed to their white counterparts. Five themes emerged from the participants: (1) distrust of the educational system; (2) perceived lack of cultural awareness of White educators; (3) perceived social stigma of the ADHD label; (4) concern about drug addiction; and (5) pressure from political forces. To provide an example of one of the above themes, a white counselor offered an explanation during her interview as to why African American parents distrust the system:

> The rating scales we use to determine ADHD are ethnocentric. They are made to the white woman system, which is what elementary school teachers basically are. There is also a problem with a minority student going to schools with a white majority…they don't fit into the norm there and are seen as having ADHD because they don't fit into how those teachers would define the norm. (p. 269)

This idea is concurrent with several studies (e.g. Reid et al., 1998; Reid et al., 2001; Epstein et al., 1998) suggesting that many behavior and teacher-rating scales are culturally biased resulting in more African American children being misdiagnosed with ADHD. Majority culture norms may not be practical in the assessment of children and adolescents of color (Cuffe et al., 1995). The current reality of African American children being misdiagnosed with ADHD has led several authors to investigate the relationship between misdiagnosis and the overrepresentation of African American children in special education classes. High prevalence rates of ADHD have been found among children in special education classes (Bussing, Zima, Belin, & Widawski, 1998). Almost 50% of children with ADHD will be placed in special education programs for learning disabilities and behavioral disorders (Reid, Maag, Vasa, & Wright, 1994). Teacher biases in the referral process combined with biases in the assessment of ADHD contribute greatly to the overrepresentation of African American children in special education classes.

Overrepresentation of African American Children in Special Education

The overrepresentation of African American children and youth in special education programs for students with learning, severe emotional or behavioral, and mental disabilities has remained a persistent reality throughout our history. The proportion of African Americans identified as mentally disabled has not changed much within the past few decades. In 1975, African Americans identified as mentally disabled constituted for 15% of the nation's school population and 38% of the special education population. In 1991, African American children constituted for 16% of the nation's school population and 35% of the special education population. Conversely, it has been well documented that African American males are particularly overrepresented in certain special education programs, as well as other disciplinary practices (i.e. recipients of corporal punishment and suspension). African American males have also been found to receive their special education in segregated classrooms or buildings (Patton, 1998). The underlying assumption is that the proportion of different ethnic groups in any program

should be equal to the proportion of that ethnic group in the general school population if there is no discrimination (MacMillan & Reschly, 1998). To no coincidence, the problems of overrepresentation are only evident in categories that we characterize as "judgmental"—that is, "those is which subjective judgments may influence decisions because the disabilities involved do not have a clear biological basis and in which contextual factors are important and in which cases are filtered through the referral process of general education teachers" (MacMillian & Reschly, 1998, p. 16). ADHD is a leading example of a disability in which the etiology is unclear and the definition is often socially constructed.

The overrepresentation of African American students in special education classes is quite problematic in part because students are not receiving adequate, effective services. The reliance on standardized examinations as measures of intelligence and success burden poorly taught children with worries of diploma denial and grade level retention. Secondly, the relationship between special education and larger sociopolitical issues tends to be overlooked. As one researcher stated, "special education, grounded in structured power relationships, is designed to serve the interests of the dominant, social, political, and economic classes and to place [and keep] African Americans in a disvalued position" (Patton, 1998, p. 27). The current reality of the disproportionate representation of African American children in special education perpetuates the sociopolitical history of the United States. Issues of inequality and oppression predate the field of education and continue to manifest today leaving African Americans in a disvalued position.

African American children in special education are especially adversely affected because they are not only discriminated against on the grounds of race, but also on the grounds of disability. The overrepresentation and misclassification of African American children in special education have resulted in a denial of equal opportunity. African American children are not only misdiagnosed for ADHD, but also, when given a correct diagnosis African American children remain under treated. Health disparities in ADHD have greatly impacted the number of children represented in special education, as African American children with ADHD are oftentimes labeled with a learning disorder. The larger sociopolitical issue behind these disparities is often sugarcoated with sprinkles of explanations and further justifications that all overlook, and furthermore, discount the impact of race. Explanations often include discussion of socioeconomic status and other financial barriers. It is imperative that we look beneath the surface and consider race, racism, and discrimination within the healthcare system as forces behind disparities in ADHD. It is also imperative that we understand the meaning of race, as it is a sociopolitical construct that is oftentimes misconstrued and misunderstood.

Why Race?

Defining Race and Ethnicity

In the recent decades, there has been debates regarding which concepts should be used as the most appropriate research and/or census variables depicting differences in population groups: race, ethnicity, or both race/ethnicity (Kendall & Hatton, 2002). Race and ethnicity, often interpreted as the same concept and also used interchangeably, have very different definitions. Race is defined as "a local geographic or global human population distinguished as more or less a distinct group by genetically transmitted physical characteristics; perceptions of genetics and physical characteristics; biologically based" (The American Heritage, 2001). Race is a socially constructed classification system that was created to define individuals by some physical characteristic (e.g. skin color, facial features). Ethnicity is defined as common ancestry through which individuals share behavioral attitudes, beliefs, lifestyles, food, spirituality, and language.

The Great Debate: Abandoning 'Race' as a Research Variable

Several researchers argue (e.g. Oppenheimer, Cooper) that we should abandon the concept or race because, ideally, we are all members of one human race. Many believe that we should no longer place surveillance on one particular group, and instead use ethnicity as an appropriate classification for public health and research practice. Shifting away from biological differences among racial groups should broaden our appreciation of various cultures and lifestyles that may in fact affect health. Others argue (e.g. Thomas, 2001; Kendall & Hatton, 2002) that the shift away from "race" to "ethnicity" will in many respects minimize the health impact of racism, especially for people of color subjected to prejudice and discrimination based on darker skin or facial features (Thomas, 2001). "Simply knowing the ethnicity of an individual or group of individuals does little to explain specific social, emotional, and mental health outcomes" (Phinney, 1996, p. 918). While we all belong to one human race, our experiences as members of different racial groups have varied.

To emphasize health differences between the races, Geronimus (2000) reported that by 1990, African American youths in some urban areas faced lower probabilities of surviving to age 45 than Caucasian youths nationwide faced to surviving to age 65. Media emphasizes the role of homicide among African American youth, although chronic diseases in early and middle adulthood are key contributors to health inequalities (Geronimus, 2000). African Americans have higher all-cause morbidity and mortality rates. African Americans also have an overall death rate that is higher than that of Caucasians. African Americans suffer from high-blood pressure, diabetes, heart disease, certain cancers (e.g. breast, prostate), lupus, and HIV/AIDS at much higher rates than Caucasians (Kendall & Hatton, 2002). The association between health and race is profound. "These disparities involve health experiences, health outcomes, and access to healthcare services and are driven by the sociopolitical realities of discrimination,

disparate educational attainment and income levels, poor working conditions, residential segregation, and material deprivation" (Kendall & Hatton, 2002, p. 22).

An important question to ask ourselves is, "Who really benefits when race is removed as a research variable"? We would lose the ability to link health status and race. We would be blinded to the subtle ways in which racism continues to shape the attitudes and behaviors of healthcare providers toward people of color (Thomas, 2001). Moreover, the dominant culture could easily overlook discriminatory practices and become "blind" to the fundamental role racist ideology plays in healthcare (Kendall & Hatton, 2002). It is important to preserve the term "race" in order to fully understand the impact of racism in the healthcare system, as racism has served as a root cause for inequalities in practices, services, and treatments.

<div align="center">Health Disparities and Racism</div>

To further comprehend the impact of racism on the healthcare system, the definition of racism needs to be understood. Racism refers to "institutional and individual practices that create and reinforce oppressive systems of race relations whereby people and institutions engaging in discrimination adversely restrict, by judgment and action the lives of those whom they discriminate" (Kreiger, 2003, p. 195). Racism is any action, attitude (conscious or unconscious) that subordinates an individual group based on skin color or race. Racism can adversely affect health in that its perpetuation in societal institutions can lead to truncated socioeconomic mobility, differential access to material resources, access to healthcare, residential segregation, and poor living conditions (Williams & Williams-Morris, 2000). Racial prejudice and discrimination measured at both the individual and institutional levels are two very important indicators of the presence of racism and its' adverse affects on health.

Individual Racism

Much of the individual racism includes assumptions and stereotypes about a person or group of people. Stereotypes are defined as unreliable generalizations and like prejudice, stereotypes "pre-judge" an individual based on assumptions. Prejudices are negative attitudes towards an entire group of people. Both prejudices and stereotypes are learned and support a larger system of social relationships. A key characteristic of racial prejudice has been an overt desire to maintain social distance from stigmatized groups. "Overwhelming support of egalitarian attitudes coexist with a desire to maintain at least some social distance from blacks and a less resounding commitment to policies to eradicate entrenched inequalities" (Williams & Williams Morris, 2000, p. 245).

Research on stereotypes revealed that many Caucasians view African American and other minorities negatively (e.g. Williams & Williams-Morris, 2000). Researchers found that 29% of Caucasians viewed most blacks as unintelligent, 44% believed that most blacks are lazy, and 56% that most blacks prefer to live on welfare. Moreover, only relatively small percentages of Caucasians portrayed positive stereotypes of blacks. Twenty percent of Caucasians believed that most blacks are intelligent, 17% that most

blacks are hard working, 13% that most blacks prefer to be self-supporting, and 15% that most blacks are not prone to violence (David & Smith, 1990 as cited in Williams & Williams-Morris, 2000).

Davison and Ford (2002) research study on the perceptions of ADHD in one African American community provides examples of assumptions, generalizations, and stereotypes that were made about African American parents and their children. Educators, medical personnel, social workers, and counselors who work with African American parents and their children were interviewed to explore their perceptions of ADHD. The following are statements from two of the interviewees.

A white medical practitioner who has worked extensively with African American families stated:

> I always had a gut sense that they [African Americans] accepted much more activity and they expect more activity and voice response and less of the compulsively well-behaved kids. In white families, you're expected to sit and listen and you attend. You don't get into things that aren't yours. It's not necessarily the expectations of African Americans, but to be very open and busy and boisterous and robust is accepted. I have also noticed, over the years, that when we do intervene with medical systems, they [African American parents] aren't necessarily pleased with the results because it really changes their child a great deal, things they valued in that child were gone. They see it as a loss of spirit.

Similarly, a white nurse who has worked with the African American community stated:

> I think there's a negative perception in the African American community. Its' viewed as a control aspect. Physical expressiveness is more accepted in the African American culture and exuberance is a desired characteristic and not something to squelch. Moms tell me all the time—that it's [the ADHD diagnosis and stimulant treatment] taking the soul out [of African American children].

The authors even noted that, "African American culture allows its members considerably greater freedom to assert and express themselves, whereas the White culture values the ability of individuals to rein in their impulses" (Davison & Ford, 2002, p. 269). Although the intentions of the authors were not to perpetuate stereotypes about African Americans, in many respects their findings contributed to several generalizations and assumptions. The perspectives above are the perceptions of educators and other personnel who work with African American parents and not that of the parents themselves. This study illustrates the importance of communicating with the primary source, which in this case would be the parents. It is dangerous to assume that African American parents want less for their children than other parents (Davison & Ford, 2002). Negative stereotypes and generalizations arise from our own assumptions about particular groups. Davison and Ford's (2002) research study is a prime example of how stereotypes can affect healthcare. Acknowledging that there are cultural differences and practices amongst the races, it is important for educators and physicians to increase their awareness of these differences, as they will work with children from races other than their own.

The perpetuation of negative stereotypes of African Americans, as they are not without consequence, suggests that there may be considerable cultural support for racist societal institutions and policies. Historically, the beliefs about the inferiority of African Americans have translated into policies that restricted the access of African Americans to educational, employment, and residential opportunities. Residential segregation has been driven by beliefs of black inferiority and an overt desire to avoid social contact with African Americans (Williams & Williams-Morris, 2000). Residential segregation in the United States illustrates institutional racism at its core.

Institutional Racism

Residential segregation has been the central mechanism by which racial inequality has been created and reinforced in the United States. Segregation has determined access to education and employment opportunities that have led to truncated socioeconomic mobility for African Americans. Segregation affects the quality of life for African Americans. African Americans reside in areas where the quality of schools is poor. Urban schools receive less funding, as it is controlled by the local government. Community wherewithal often determine the quality of the school and many urban communities do not have adequate resources. In the last several decades there has been a mass movement of low-skilled high-pay jobs from the urban areas where African Americans are mostly concentrated (Williams & Williams-Morris, 2000). The lack of educational opportunities for African Americans and the big shift in the job market have resulted in a 'spatial mismatch' and a 'skills mismatch'. Spatial mismatch refers to residing in an area where the residents lack proximity to entry-level jobs. Skills mismatch refers to the availability of jobs in an area where the residents do not have the level of skill and training required (Williams & Williams-Morris, 2000). Lack of job access leads to high rates of unemployment, underemployment, and poverty-stricken conditions.

The physical separation of the races continues only because of the cooperative efforts of major institutions, including real estate, banking institutions, and housing policies. The institutional policies combined with the efforts of individual discrimination ensured that African Americans were limited in housing choices to the least desirable residential areas (Williams & Williams-Morris, 2000). The individual discrimination includes 'white-flight', or the effort of Caucasians to move out of communities when the African American population increases. Isolating African Americans in segregated communities contribute to inadequate employment and educational opportunities, as well as access to healthcare.

> Race continues to determine health status and the social allocation of resources and opportunities. Inequality in socioeconomic status, educational and work opportunities, and residential segregation remain the strongest indicators of the viability of racism within organized institutional structures. The vast majority of discriminatory practices lie within the macro societal structures of our society, not in individuals (Kendall & Hatton, 2002, p. 24).

It is important to understand the centrality of racism to further understand racial health disparities. At an institutional level, it is clear that discrimination influences the economic opportunities that people have, as well as the quality of health services they receive (Nazroo, 2003).

<center>Conclusion</center>

A basis indicator of health disparities among racial groups lie in the experiences of overt discrimination, fundamental inequalities that exist in this society, and the effects of living in a society that still binds its prejudices to the color of one's skin. The discussion of race and why race matters is fundamental because the experiences of people of color in the United States are often dictated by race and skin color. "We need to do a better job at understanding how to measure race, racism, and social inequality in medical care and public health practices" (Thomas, 2001, p. 1046).

To decrease the racial influence on health, further research needs to be conducted identifying the various health disparities among the races. Researchers should also continue to uncover the root causes of these health disparities acknowledging that socioeconomic status and financial barriers are not the only explanations. After controlling for possible mediating and/or moderating variables (e.g. socioeconomic status, maternal education, health insurance status), researchers, as noted earlier, have found health disparities in ADHD for African American children, as African American children have not received adequate treatment. It has become quite clear that an improved explanation is needed to understand why these health disparities exist between the races. The explanation behind racial health disparities, as presented in this paper, is racism. A comprehensive discussion of racism and the effects of discriminatory practices are long overdue. Eliminating racial disparities in healthcare will require the efforts of policymakers, educators, physicians and other medical personnel who will all need to make a conscious attempt to discuss the impact of racism on health.

A greater emphasis should be placed on multiculturalism. Multiculturalism values the perspectives and viewpoints of various racial groups, discussing issues related to race and race relations such as discrimination and individual and institutional racism. Multiculturalism is described as a perspective of cultural pluralism, which acknowledges the cultural context of all health processes, as they may differ from racial group to racial group (Kendall & Hatton, 2002). Diversity and multiculturalism training should be required for educators and healthcare providers. The impact of racism on the healthcare system needs to be greatly considered and understood before we can expect an elimination of health disparities in ADHD.

References

American Psychiatric Association (4th ed.). (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.

Bussing, R., Zima, B.T., Gary, F.A., & Garvan, C.W. (2003). Barriers to detection, help-seeking, and services for children with ADHD symptoms. *The Journal of Behavioral Health Services & Research, 30*, 176-189.

Bussing, R., Zima, B.T., Perwien, A.R., Belin, T.R., & Widawski, M. (1998). Children in special education programs: Attention deficit hyperactivity disorder, use of services, and unmet needs. *American Journal of Public Health, 88*, 880-886.

Cooper, P., & O'Regan, F.J. (2001). *Educating children with ADHD: A teacher's manual*. London: RoutledgeFalmer.

Cuffe, S.P., Waller, J.L., Cuccaro, M.L., Pumariega, A.J., & Garrison, C.Z. (1995). Race and gender differences in the treatment of psychiatric disorders in young adolescents. *Journal of the American Academy of Child and Adolescent Psychiatry, 34,* 1536-1543.

Davison, J.C. (2001). Attention deficit/hyperactivity disorder: Perspectives of participants in the identification and treatment process. *Journal of Emotional Thought, 35(3),* 227-247.

Davison, J.C., & Ford, D.Y. (2002). Perceptions of attention deficit hyperactivity disorder in one African American community. *Journal of Negro Education, 70,* 264-274.

DuPaul, G.J., & Stoner, G. (2003). *ADHD in the schools: Assessment and intervention strategies.* New York: The Guilford Press.

Entwistle, P.C. (2000). Assessment of attention in children with attention deficit hyperactivity problems in primary care settings. *Journal of Clinical Psychology in Medical Settings, 7(3),* 159-166.

Epstein, J.N, March J.S., Conners, C.K., & Jackson, D.L. (1998). Racial differences on the Connors Teacher Rating Scale. *Journal of Abnormal Psychology, 26*, 109-118.

Geronimus, A (2002). To mitigate, resist, or undo: Addressing structural influences on the health of urban populations. *American Journal of Public Health, 90(6)*, 867-872.

Jordan, D.R. (1998). *Attention Deficit Disorder: ADHD and ADD syndromes.* Texas: PRO-ED.

Hahn, B.A. (1995). Children's health: Racial and ethnic differences in the use of prescription medications. *Pediatrics, 95(5),* 727-732.

Karlsen, S., & Nazroo, J.Y. (2002). Relation between racial discrimination, social class, and health among ethnic minority groups. *American Journal of Public Health, 92*, 624-631.

Kendall, J., & Hatton, D. (2002). Racism as a source of health disparity in families with children with attention deficit hyperactivity disorder. *Advances in Nursing Science, 25*, 22-39.

Krieger, N. (2003). Does racism harm health? Did child abuse exist before 1962? On explicit questions, critical science, and current controversies: An ecosocial perspective. *American Journal of Public Health, 93(2),* 194-199.

MacMillan, D.L., Reschly, D.J. (1998). Overrepresentation of minority students: The case of greater specificity or reconsideration of the variables examined. *The Journal of Special Education, 32*, 15-24.

Matthews, D.D. (2002). *Attention Deficit Disorder sourcebook.* United States: Omnigraphics.

Millichap, J.G. (1998). *Attention Deficit Hyperactivity and Learning Disorders*. Chicago: PNB Publishers.

National Institute of Mental Health. (1996). *Attention Deficit Hyperactivity Disorder*. Retrieved July 14, 2003, from http://www.nimh.nih.gov/publicat/adhd/cfm.

Nazroo, J.Y. (2003). The structuring of ethnic inequalities in health: Economic position, racial discrimination, and racism. *American Journal of Public Health, 93(2)*, 277-284.

Newacheck, P.W., Hughes, D.C., & Stoddard, J.J. (1996). Children's access to primary care: Differences by race, income, and insurance status. *Pediatrics, 97*, 26-32.

Pastor, P.N., & Reuben, C.A. (2002). Attention Deficit Disorder and Learning Disability: United States, 1997-98. *Vital and Health Statistics*, 10, 1-12.

Patton, J.M. (1998). The disproportionate representation of African Americans in special education: Looking behind the curtain for understanding and solutions. The *Journal of Special Education, 32(1),* 25-31.

Phinney, J.S. (1996). When we talk about American ethnic groups, what do we mean? *American Psychologist, 51(9),* 918-927.

Reid, R. (1995). Assessment of ADHD with culturally different groups: The use of the behavioral rating scales. *School Psychology Review, 24(4),* 537-560.

Reid, R., DuPaul, G.J., Power, T.J., Anastopoulos, A.D., Rogers-Adkinson, D., Noll, M., et al. (1998). Assessing culturally different students for attention deficit hyperactivity disorder using behavior rating scales. *Journal of Abnormal Psychology, 26,* 187-198.

Reid, R., Casat, C.D., Norton, H.J., Anastopoulos, A.D., & Temple, E.P. (2001). Using behavior rating scales for ADHD across ethnic groups: The IOWA Conners. *Journal of Emotional and Behavioral Disorders, 9,* 210-218.

Reid, R., Magg, J.W., Vasa, S.F., & Wright, G. (1994). Who are the children with attention deficit-hyperactivity disorder? A school-based survey. *The Journal of Special Education, 28,* 117-137.

Safer, D.J., Zito, J.M., & Fine, E.M. (1996). Increased methylphenidate usage for attention deficit disorder in the 1990s. *Pediatrics, 98*, 1084-1088.

Safer, D.J., & Krager, J.M. (1994). The increased rate of stimulant treatment for hyperactive/inattentive students in secondary school. *Pediatrics, 94*, 462-464.

Samuel V.J., Biederman, J., Faraone, S.V., George, P., Mick, E., Thornell, A., et al. (1998). Clinical characteristics of attention hyperactivity disorder in African American children. American Journal of Psychiatry, 155, 696-698.

The American Heritage (4th ed.). (2001). New York, NY: The Dell Book.

Thomas, S.B. (2001). The color line: Race matters in the elimination of health disparities. *American Journal of Public Health, 91(7),* 1046-1048.

Williams, D.R., & Williams-Morris, R. Racism and Mental Health: The African American experience. *Ethnicity and Health, 5,* 243-268.

Wilson, K.B, & Senices, J. (In press). Exploring the vocational rehabilitation acceptance rates of Hispanic vs. Non-Hispanics in the United States. *Journal of Counseling and Development*.

Zima, B.T., Bussing, R., Crecelius, G.M., Kaufman, A., & Berlin, T.R. (1999). Psychotropic medication use among children in foster care: Relationship to severe psychiatric disorders. *American Journal of Public Health, 89*, 1732-1735.

Zito, J.M., Safer, D.J., Dosreis, S., & Riddle, M.A. (1998). Racial disparity in psychotropic medications prescribed for youths with Medicaid insurance in Maryland. *Journal of the American Academy of Child and Adolescent Psychiatry, 37,* 179-184.

The Beach Study: An Empirical Analysis of the Distribution of Coastal Property Values

Christopher Major, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Kenneth Lusht, Chairman
Insurance & Real Estate Department
Pennsylvania State University

Abstract: Much casual and less empirical evidence suggests that coastal properties, and particularly those proximate to a beach, have outperformed most real estate market segments over the past decade. Less well understood is the relative importance of the property and location characteristics that drive selling prices in these markets. This information is useful to investors, developers, and taxing jurisdictions, as well as those purchasing mainly for consumption.   In this study we examine the impact on property values of proximity to waterfront. We explain differences in the prices of residential properties in Stone Harbor and Avalon, New Jersey from January 2002 through June 2003. We find a steeper land value gradient than have prior studies.

Introduction

Proximity to negative and positive externalities is a key component of location. There is a rich literature that provides empirical estimates of the impact of proximity to amenities such as schools (Brasington, 1999), golf courses (Do and Grudnitski, 1997), and parks (Harner, et al, 1974), and to airport noise (Bell, 2001), overhead transmission lines (Wolverton and Bottenmiller, 2003), and toxic waste sites (Reichert, 1997) on the potentially negative side. Estimates of the impacts of these kinds of location-specific variables are necessary to make informed valuation decisions used by buyers, sellers, lenders, and tax assessors.

The location attribute of interest in this study is proximity to the ocean. Much casual and less empirical evidence suggests that coastal properties, and particularly those proximate to a beach, have outperformed most real estate market segments over the past decade. Less well understood is the relative importance of the property and location characteristics that drive selling prices in these markets. Prior research has focused mainly on the impact of view. Plattner and Campbell (1978) measured the impact of a water view on the prices of new condominiums in western Massachusetts. They found view premiums of 4%-12%, and that the percentage premium tended to be higher for lower-priced units. Gillard (1981) found a 9% view premium in Los Angeles.   Our study comes closest to the work of Benson et. al. (1997). They found that in the Point Roberts, Washington market, oceanfront views added 147% to value, ocean views added 32%, and partial ocean views added 10%. These results, spanning over two decades, show a consistent premium for water proximity, and that the premium has tended to increase over time.

We measure the impact of proximity to the ocean in the contiguous markets of Stone Harbor and Avalon, New Jersey. Our focus is distance and not views, as there is reason to believe that proximity impacts include, but are not limited to, differences in

view. Unlike some prior research, we use transactional data rather than assessed values. Our sample period of January 2002 through June 2003 provides a fresh measure of price distribution in what has been a historically volatile type of market.

Data and Methodology

Stone Harbor and Avalon, New Jersey are contiguous communities located on a seven-mile island about 90 miles southeast of Philadelphia and 45 miles south of Atlantic City. In the 2000 Census, the combined total housing units and permanent population of the study market were 8,709 and 3,271 respectively. The Chambers of Commerce estimate that in the summer months, tourists increase the population to 50,000. Figure 1 is a map of the study market. It shows the two communities side by side; in fact it is a continuous market as shown in the 7-Mile Island inset map. This is in some ways an ideal laboratory to measure proximity-to-water impacts. Because it is an island, there is the opportunity to measure both ocean- and bay- front impacts, and there are fewer intervening variables that must be controlled.

| **Figure 1.** | **Stone Harbor and Avalon** |
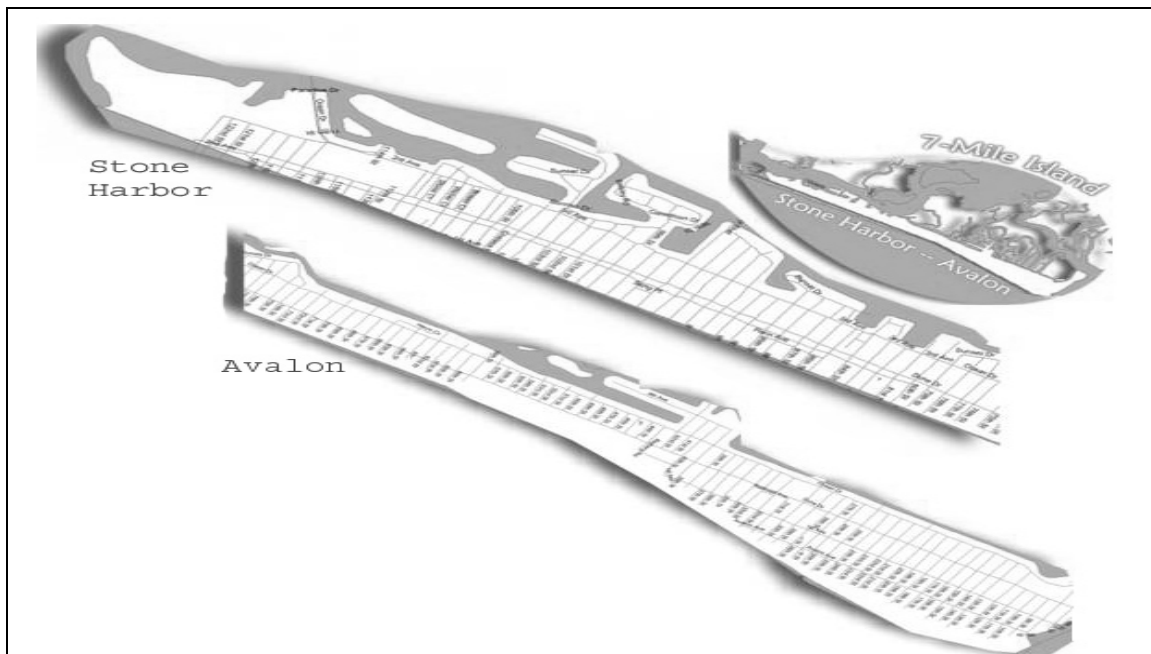
The data are from a sample of 249 residential sales that occurred from January 1, 2002 through June 26, 2003. This is 100% of MLS transactions during that period.

Equation 1 is the initial model specification.

166

Equation 1.	**Initial Model Specification**

PRICE = f(LOT SIZE, BATHROOMS, LAVATORIES, BEDROOMS, NEW, BEST
NEIGHBORHOOD, NORTH/SOUTH, EAST/WEST, 2ND BLOCK, BEACH
BLOCK, BAYFRONT, OCEANFRONT, AVALON, 1ST HALF 2002,
2ND HALF 2002)

where:

PRICE = 2002 – 2003 MLS Sale Price for Stone Harbor or Avalon, NJ Property,

LOT SIZE = square feet of lot size,
BATHROOMS = number of bathrooms,
LAVATORIES = number of lavatories,
BEDROOMS = number of bedrooms,
NEW = 1 or 0 if building built before or after 1975,
BEST NEIGHBORHOOD = 1 in the best neighborhood, 0 if not,
NORTH/SOUTH = 1 if property is on the north side of street, 0 if not,
EAST/WEST = 1 if property is on the east side of the street, 0 if not,
2ND BLOCK = 1 if property is located on the second block from the beach, 0 if not,
BEACHBLOCK = 1 if property is located on the block closest to the beach, 0 if not,
BAYFRONT = 1 if property is located on the bayfront, 0 if not,
OCEANFRONT = 1 if property is located on the oceanfront, 0 if not,
AVALON = 1 if property is in Avalon, 0 if in Stone Harbor,
1st HALF OF 2002 = 1 if property was sold between 1/1/02-6/30/02, 0 if not,
2nd HALF OF 2002 = 1 if property was sold between 7/1/02-12/31/02, 0 if not.

Descriptive statistics and expectations are shown in Table 1.

Table 1. Descriptive Statistics and Expectations

	N	Mean	Minimum	Maximum	Expectation
Dependent Variable					
PRICE		$1,115,199	$215,000	$5,700,000	
Independent Variables					
LOT SIZE	249	6,348	925	25,950	+
BATHROOMS	614	2.5	1	6	+
LAVATORIES	100	0.5	0	3	+
BEDROOMS	988	4	1	8	+
NEW	125				+
NEIGHBORHOOD					
Best	44				+
Other	205				-
SIDE OF STREET					
North	82				-
South					Default
East	67				+
West					Default
PROXIMITY TO WATERFRONT					
Oceanfront	8				+
Beachblock	55				+
2nd Block	82				+
Bayfront	25				+
Not Proximate	79				Default
STONE HARBOR/AVALON					
Avalon	168				?
Stone Harbor	81				?
DATE OF SALE					
2002 1st Half	62				?
2002 2nd Half	126				?
2003 1st Half	61				?

Many of our variables are standard in property valuation models, and their empirical expectations shown in Table 1 are well known. We expect lot size, bedrooms, bathrooms, and the best neighborhood to be positively associated with selling price. Other variables are somewhat unique to this kind of market and our empirical expectations are less certain. Being on the north or south side of streets (that run east and west) affects the amount of sunlight a property receives. Being on the east or west side of streets (that run north and south) affects proximity to the beach, not so much because the distance is significantly different, but because a western location requires crossing one extra street to get to the beach. The Stone Harbor/Avalon variable controls for any pricing difference between the two communities. In casual conversations with agents, most thought Stone Harbor would command a premium, but estimates of magnitude varied widely. The date-of-sale variables control for price trends over time. The default

is January 1- June 26, 2003. Although property values have increased significantly over the past ten years, that may not be the case during our sample period. Agents claim the recent collapse of the stock market and the weak economy has put downward pressure on prices.

The variables of primary interest are those that measure the impact of proximity to water. Proximity to the ocean is measured as either oceanfront or the number of blocks to the ocean. Proximity to the bay is measured as either bayfront or not. The default variable includes those properties not on the bay and more than two blocks from the ocean. We expect the coefficients on the distance variables to be positive. Finally, a property characteristic typically found to be strongly associated with selling price is square feet of living area. It is omitted in our model because data were not available. According to agents, bathrooms, lavatories, and bedrooms drive values in this predominantly rental market, with total size secondary.

The model was initially estimated using ordinary least squares. The results were acceptable with respect to the R^2 and the signs and magnitudes of most independent variables. Bedrooms, sides of street, and dates of sale were insignificant, and one of them, bedrooms, had the wrong sign. Inspection of the correlation coefficients showed high colinearity between bathrooms and bedrooms. A potentially more serious concern was the presence of mild heteroskedasticity (based on inspection of residual patterns) and the suspicion of autocorrelation based on a marginal Durbin-Watson statistic. These problems were corrected by reestimating Equation (1) using weighted variables in an AR1 model.

Results

The best regression, shown in Table 2, was one that omitted the three variables found insignificant in our initial estimation. Most variables are significant at the 1% or better level. The significance of bayfront is marginally higher than 10%, and one variable, Avalon, is insignificant. All signs are in the expected direction, and the magnitudes of the coefficients on most variables are reasonable. We attribute the seemingly high bathrooms and lavatories coefficients to the fact that they are serving as proxies for the bedroom and living area variables not in the model. The estimation produced an acceptable R^2 (.82), and the Durbin-Watson statistic and an inspection of the residuals indicated autocorrelation and heteroskedasticity were no longer an issue.

Table 2. | **Regresion Results with Ten Independent Variables**

Variable	Coefficient	Standard Error	Approx t Value	Pr > |t|
Intercept	-2,318	2,277	-1.02	0.3098
Lot Size	73.8980	16.4130	4.50	<.0001
Bathrooms	141,440	18,085	7.82	<.0001
Lavatories	135,183	23,813	5.68	<.0001
New	54,846	29,778	1.84	0.0668
Oceanfront	1,651,392	127,524	12.95	<.0001
Bayfront	162,416	99,800	1.63	0.1050
Beach Block	487,340	66,496	7.33	<.0001
2nd Block	111,810	36,325	3.08	0.0023
Best Neighborhood	64,959	34,264	1.90	0.0592
Avalon	45,613	90,502	0.50	0.6147

R-Square = 0.817   Durbin-Watson = 2.014

Focusing on proximity impacts, we use as a base price an "average" property; that is, a property with the sample average lot size, number of baths, and so on. We assume the property is in Avalon. The predicted price of this property not proximate to the waterfront (that is, not on the bayfront or within two blocks of the ocean) is $1,055,601. The same property on the 2nd Block would sell for $1,167,411 (+10.5%), on the Beach Block would sell for $1,542,941 (+46%), on the Bayfront would sell for $1,218,017 (+15%), and on the Oceanfront would sell for $2,706,933 (+156%). These results show exponential growth in the price premium as proximity increases. The lower premium for proximity to the bayfront compared to proximity to the oceanfront is as expected.

These findings are similar to the results of prior research. An exponentially increasing premium has typically been found, and in the past several decades these price gradients have become steeper. Our results show these trends continue, and there are reasons to believe that will not change in the foreseeable future. With an aging population and the loss of some confidence in financial markets, the demand for waterfront retirement and vacation homes should remain strong. At the same time, more stringent constraints on development mean the effective supply of shoreline locations is stagnant or effectively declining.

Summary and Conclusion

We estimate a model using a 100% sample of residential sales in a market with varying proximities to two types of water, ocean and bay. The results are consistent with prior findings, with respect to the direction of proximity impact, and suggest an acceleration of the historical trend toward steeper price gradients. One implication is that estimates of property values and distributions of values in these kinds of markets are likely to become stale relatively quickly.

References

Bell, Randall. "The Impact of Airport Noise on Residential Real Estate," The Appraisal Journal**.** Chicago: Jul 2001. Vol. 69, Iss. 3; p. 312.

Benson, E. D., J. L. Hansen, A. L. Schwartz, Jr. and G. T Smersh, "The Influence of Canadian Investment on U. S. Residential Property Values," journal of Real Estate Research (1997, 13: 3): 231-249.

Brasington, David M. "Which Measures of School Quality Does the Housing Market Value?" The Journal of Real Estate Research**.** Sacramento: Nov/Dec 1999. Vol. 18, Iss. 3; p. 395.

Gillard, Q, "The Effect of Environmental Amenities on House Values: The Example of a View Lot," Professional Geographer (1981, 33: 2): 216-220.

Grudnitski, G. and A. Q. Do, "Adjusting the Value of Houses Located on a Golf Course," The Appraisal journal (July, 1997): 261-266.

Plattner, R. H. and T. J. Campbell, "A Study of the Effect of Water View on Site Value," The Appraisal journal (January 1978): 20-25.

Reichert, Alan K. "Impact of a Toxic Waste Superfund Site on Property Values," The Appraisal Journal. Chicago: Oct 1997. Vol. 65, Iss. 4; p. 381

Wolverton, Marvin L. and Bottemiller, Steven C. "Further Analysis of Transmission Line Impact on Residential Property Values," The Appraisal Journal**.** Chicago: Jul 2003. Vol. 71, Iss. 3; p. 244

Elucidating the Neurobiological Basis for the Locomotor Effects of Ethanol

Elizabeth K. O'Malley, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Kyung-An Han
Assistant Professor, Department of Biobehavioral Health
College of Health and Human Development
Pennsylvania State University

Abstract

The effects of alcohol have been studied immensely in humans. It is still unknown what specifically causes a person to develop alcoholism. By looking at the locomotor effects of alcohol on two *Drosophila* mutants one possible cause for alcohol resistance could be established. Previous studies have revealed that flies deficient in the D_1 dopamine receptor DAMB (D_5-like) and the neighboring gene oxoglutarate carrier (OGC) display attenuated locomotor activities compared to the control parental line ry upon exposure to various concentrations of ethanol. The goal of my study is to identify which gene, *damb* or *ogc*, is responsible for the reduced response to ethanol. The results of this study show that the *damb* mutant was less responsive to alcohol than the *ogc* mutant and the DAMB receptor may play a role in mediating the locomotor activating effect of alcohol.

Introduction

Alcoholism is defined as a disease or illness which is characterized by the inability to limit or control alcohol use. More specifically, alcoholics are unable to limit the duration of the drinking episode, the quantity of alcohol consumed, and the behavioral consequences of drinking (www.ncadd.org). The statistics on alcoholism are staggering. The disease has a prevalence rate of approximately 8 to 14%, and affects twice as many men as women. Measuring the true prevalence of alcoholism is difficult, in part because sufferers do a very good job of hiding their illness. It is known that alcoholism is the fourth leading cause of death for people between the ages of 35 and 55, and 100,000 people die from this disease every year (Enoch and Goldman, 2002). The exact etiology of alcoholism is unknown. It is likely that both environmental and genetic factors are involved in the development of the disease. There are currently several different genetic studies being conducted to determine the specific genes involved in alcoholism (Dick and Foroud, 2003).

Drosophila melanogaster is a very popular animal model in studying developmental processes and nervous system function (Heberlein, 2000). The fly makes an excellent model organism for several reasons. It has a relatively small genome; 14,000 genes compared to a human's with about 70,000 genes. *Drosophila* has a short

life cycle of about two weeks. This makes it extremely easy and inexpensive to generate a large number of offspring for experiments in a short amount of time (http://www.ceolas.org/fly/intro.html). Most important to this study, intoxicated flies display many of the same behaviors as intoxicated humans. After exposure to alcohol the flies display hyperlocomotor activity, become disoriented, and then become sedated (Heberlein, 2000). This is how we can gain insights into human drug addiction by investigating fruit flies.

This study was focused on determining the neuromodulator receptors involved in the locomotor activating effect induced by a relatively mild ethanol exposure. Two mutant *Drosophila* genotypes were employed in this study. The first mutant, *damb,* lacks the genes for the DAMB receptor and the genes for the OGC receptor. DAMB has been identified as a dopamine D1 receptor which is highly concentrated in the mushroom bodies (Han et al., 1996). The mushroom bodies in the fly brain are a center for learning and memory (Han et al., 1996). The other mutant used in this study lacks only the genes for OGC. The exact function of OGC is yet to be determined, but it is found in mitochondrion and may be involved in energy metabolism (Fly Base, 2003).

Dopamine is one of the main neurotransmitters involved in the rewarding effects of ethanol. Dopamine is released when we engage in a dopaminergic neurotransmission activity that is found pleasurable. Alcohol increases dopamine release, therefore likely increasing the feeling of pleasure (Julien, 2001). Many studies have been conducted on the exact role and mechanism that dopamine plays in alcohol-induced responses. Bainton et al. tested flies with reduced dopamine levels and found that these flies had a decreased response to ethanol-induced hyperlocomotion (Bainton et al., 2000). This suggests that dopamine plays a role in the locomotor activating affect of ethanol. Cohen et al. conducted a study looking specifically at the dopamine D1 receptor (Cohen et al., 1999). The results of this study showed that the blockade of the dopamine D1 receptor causes decreases in ethanol self-administration in rats. The study also showed that the dopamine D1 receptor plays a significant role in ethanol induced locomotion (Cohen et al., 1999).

Research Questions
1) Does DAMB mediate the locomotor effects of ethanol?
2) Does OGC mediate the locomotor effects of ethanol?

Significance of the Study

If a person were more resistant to the effects of alcohol, due to malfunction or altered function of a certain receptor, they would drink more to achieve the desired effect of intoxication. If it is known which cellular mechanisms in the brain control the behavioral effects of alcohol, and also the rewarding effects, this may lead to a pharmaceutical treatment for alcoholism. Also, it will help identify people who are more at risk for developing an alcohol abuse problem (i.e. those with malfunctioning receptors).

Methods

The purpose of this study is to elucidate which molecules mediate the locomotor effect of ethanol in fruit flies. The flies were drawn from a random population of laboratory mutants. Only male flies were examined because eggs present in female flies might interfere with the results. This study examined 24 flies. The sample includes two different mutants: the *damb* mutant and the *ogc* mutant.

To elucidate the effects of ethanol on behavioral responses of flies, we employed a locomotor assay system. In this system, individual flies were placed in a glass tube with both ends plugged with cotton balls. The cotton ball in one side was soaked with various concentrations of ethanol (15% and 30%). The other cotton ball was soaked with distilled water. In order to account for side preference, the cotton ball soaked in ethanol was placed on different sides of different tubes. The number of times that flies displayed distinct locomotor behaviors during an hour was recorded using a video camera. The behaviors that were monitored include: single turn 1 (turning around in the middle of the tube), single turn 2 (turning around at the end of the tube), multiple turn 1 (walking in circles), and multiple turn 2 (turning side to side without walking). It is important to note that single turn 2 is normal fly walking behavior.

In this study, there are two different sets of independent variables. The first set includes the two different genetic mutants (*damb* and *ogc*). The next set of independent variables is two different concentrations of ethanol: 15% and 30%. The dependent variables in this study are: the number of single turn 1, single turn 2, multiple turn 1, and multiple turn 2.

The data recorded after all runs of the experiment was put into an Excel spreadsheet for analysis. T-tests were run to test for statistical significance.

Results

The ethanol induced in *ogc* mutants more single turning 1 than *damb* mutants with both 15% and 30% ethanol ($p < .05$) (shown in Figures 1 and 2). This indicates that the *damb* mutants are less responsive to ethanol than the *ogc* mutants. There was no difference in the number of single turn 2 (which is normal fly walking behavior) with water and ethanol exposures in both genotypes (both *damb* and *ogc*) ($p > .05$). This was observed in both trials using 15% ethanol and 30% ethanol (shown in Figures 3 and 4). This observation suggests that ethanol does not influence normal fly walking behavior. The ethanol-exposed *ogc* mutants displayed more multiple turning 1 than the ethanol-exposed *damb* mutants ($p < .05$) (shown in Figures 5 and 6). Ethanol at 15% and 30% induced more activity in *ogc* flies compared to *damb* mutants. The ethanol-exposed *ogc* mutants also displayed more multiple turning 2 than the *damb* mutants ($p < .05$) (shown in

174

Figures 7 and 8). As seen with single turn 1, the *damb* mutants are less responsive to ethanol than the *ogc* mutants when multiple turns were measured.
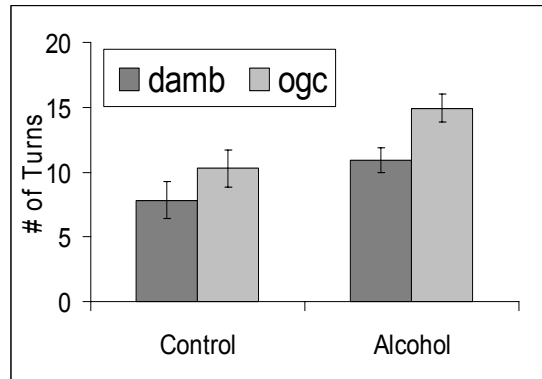
Figure 1: *ogc* and *damb* mutants' number of single turn 1 under 15% ethanol. Each bar represents a mean of 24 and standard error of the mean.
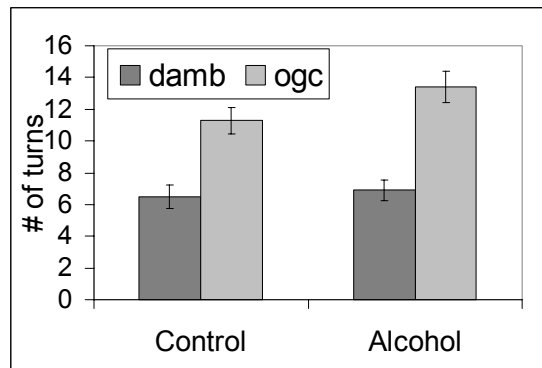
Figure 2: *ogc* and *damb* mutants' number of single turn 1 under 30% ethanol. Each bar represents a mean of 24 and standard error of the mean.
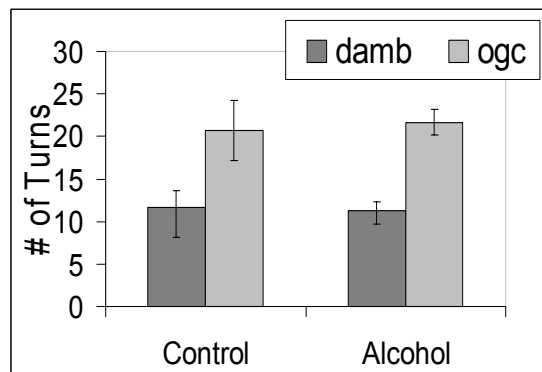
Figure 3: *ogc* and *damb* mutants' number of single turn 2 under 15% ethanol.

Each bar represents a mean of 24 and standard error of the mean.
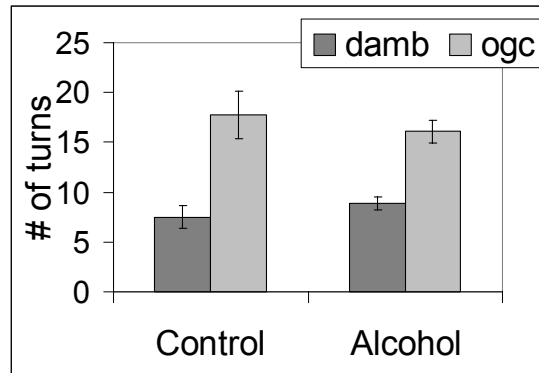
Figure 4: *ogc* and *damb* mutants number of single turn 2 under 30% ethanol
Each bar represents a mean of 24 and standard error of the mean.
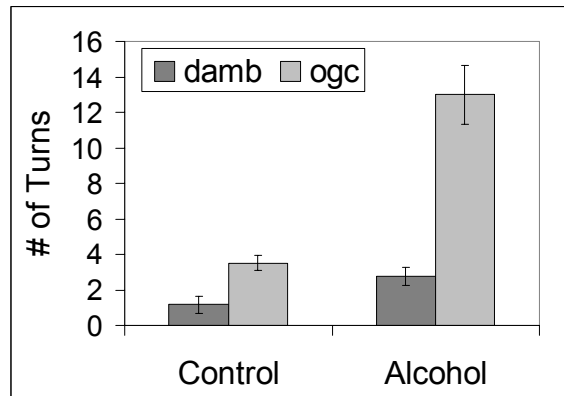
Figure 5: *ogc* and *damb* mutants number of multiple turn 1 under 15% ethanol
Each bar represents a mean of 24 and standard error of the mean.
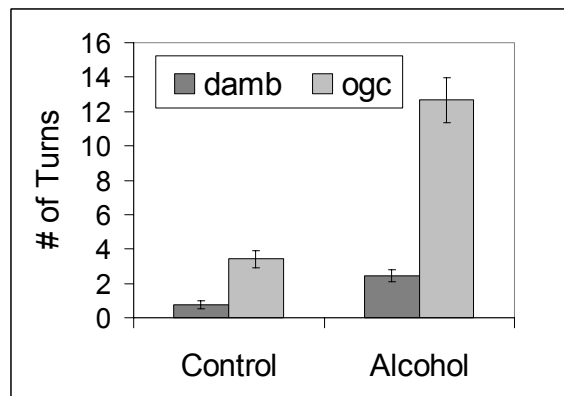
Figure 6: *ogc* and *damb* mutants number of multiple turn 1 under 30% ethanol
Each bar represents a mean of 24 and standard error of the mean
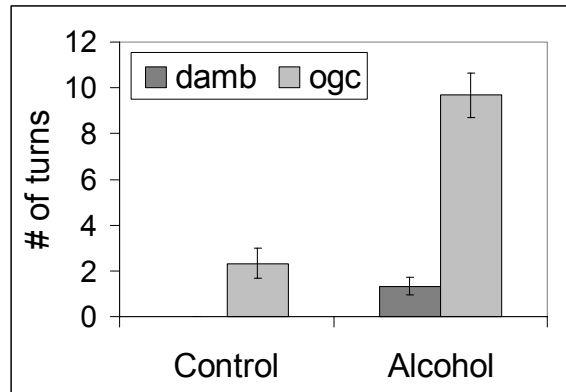
Figure 7: *ogc* and *damb* **mutants number of multiple turn 2 under 15% ethanol
Each bar represents a mean of 24 and standard error of the mean.**
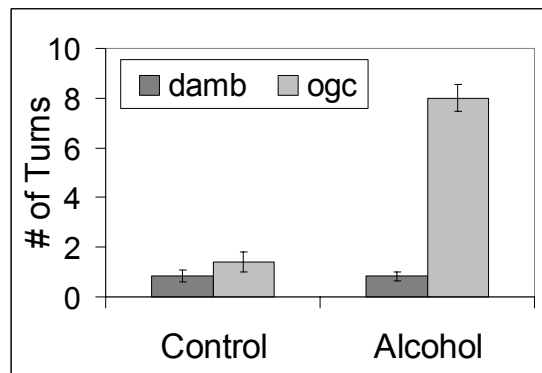
Figure 8: *ogc* and *damb* **mutants number of multiple turn 2 under 30% ethanol
Each bar represents a mean of 24 and standard error of the mean.**

Conclusions

A number of conclusions can be made from these results. Since the *damb* mutants were less responsive to ethanol than the *ogc* mutants, it can be concluded that DAMB plays a role in mediating the locomotor activating effect of alcohol (single turn 1 and multiple turn 1) and also disorientation (multiple turn 2). In broad terms this implies that people, who have fewer numbers of the DAMB receptor, or malfunctioning or less active DAMB receptors, may be more resistant to the effects of alcohol, which may cause them to drink more.

For future work, these experiments need to be repeated with the same variables to test reliability and validity, and with different concentrations of ethanol in order to establish a dose-response relationship.

References

Bainton R, Tsai, L, Singh C, Moore M, Neckameyer W, & Hebelein U (2000). Dopamine modulates acute responses to cocaine, nicotine, and ethanol in *Drosophila*. *Current Biology, 10*, 187-194.

Cohen C, Perrault G, & Sanger D (1999). Effects of D1 dopamine receptor agonists on oral ethanol self-administration in rats: comparison with their efficacy to produce grooming and hyperactivity. *Psychopharmacology, 142*, 102-110.

Dick D, and Foroud T (2003). Genetic Strategies to Detect Genes Involved in Alcoholism and Alcohol-Related Traits. *The National Institute on Alcohol Abuse and Alcoholism*. Retrieved on October 23, 2003, from http://www.niaaa.nih.gov/publications/arh26-3/172-180.htm.

Enoch M: Problem Drinking and Alcoholism: Diagnosis and Treatment. *American Family Physician* 2002 Feb 1;65(3): 441-448.

Han KA, Millar NS, Grotewiel MS, & Davis RL (1996). DAMB, a novel dopamine receptor expressed specifically in *Drosophila* mushroom bodies. *Neuron, 16*, 1127-1135.

Heberlein U (2000). Genetics of alcohol-induced behaviors in *Drosophila*. *Alcohol Research and Health, 24*, 185-188.

Julien R: *A Primer of Drug Action*. 2001. New York: Worth Publishers. 41-47.

The Flybase Consortium (2003). The Flybase database of the *Drosophila* genome projects and community literature. *Nucleic Acids Research, 31*, 172-175.

The National Council on Alcoholism and Drug Dependence. (n.d.) *Definition of Alcoholism*. Retrieved on March 23, 2003 at http://www.ncadd.org.

The WWW Virtual Library: *Drosophila.* (n.d.). *A quick and simple introduction to Drosophila melanogaster.* Retrieved October 23, 2003, from http://ceolas.org/VL/fly/intro.html.

Optimization of Mineral Activation for CO$_2$ sequestration

Hui X. Ou, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
M. Mercedes Maroto-Valer, Assistant Professor
Energy and Geo-Environmental Engineering
Virginia S. and Philip L. Walker, Jr., Faculty Fellow
Pennsylvania State University

Abstract

Mineral carbonation, the reaction of CO$_2$ with non-carbonate minerals to form stable mineral carbonates, has proved to be a promising concept for permanent CO$_2$ sequestration. However, there are some drawbacks of this technology: the reaction kinetics that require pulverization of the raw materials, long reaction times and high partial pressures. In the previous studies conducted at the Energy Institute of Penn State University, a novel active carbonation concept, which utilized surface activation to accelerate the reaction rates and efficiencies for forming carbonates from minerals, was developed. This research project took a step forward to further optimize the active carbonation process. A parametric study was conducted in order to increase the efficiency of mineral activation process for subsequent CO$_2$ sequestration.

Introduction

Fossil fuels are formed in the earth from the plant or animal remains; include natural gas, petroleum, and coal. They account for nearly 85% of the world's energy supply (DOE, 1999). In the past 60 years, the amount of anthropogenic carbon dioxide (CO$_2$) emitted to the atmosphere, primarily because of expanding use of fossil fuels for energy, has risen from pre-industrial levels of 280 parts per million (ppm) to present levels of over 365 ppm (Keeling and Whorf 1998). Prediction of global energy use in the next century suggests a continued increase in carbon emissions and rising concentrations of CO$_2$ in the atmosphere. Although the effects of CO$_2$ levels on global climate are uncertain, there is scientific consensus that a doubling of atmospheric CO$_2$ concentrations could have a variety of serious environmental consequences in the next century (DOE, 1999).

There are three main approaches to carbon management: (1) to increase the efficiency of primary energy conversion; (2) to substitute lower-carbon or carbon-free energy sources; and (3) to capture and sequester CO$_2$ emission. It is generally accepted that the first two alternatives will only provide incremental improvements, and therefore, carbon sequestration technologies must be developed to achieve zero emissions (DOE, 1999).

Carbon sequestration refers to the removal and long-term storage of carbon dioxide from the atmosphere or emission source. It is essential for the continued large-scale use of fossil fuels. Some other benefits for carbon sequestration includes: manufacture of commercial products; improved agricultural practice that could reduce

179

soil erosion, conserve water, and increase the sustainability of food production; the restoration of wetlands, which would help preserve wildlife and protect estuaries; increased biodiversity; enhanced recovery oil and methane; and the development of exportable technologies to help the U.S. economy (DOE, 1999).

Mineral carbonation, which involves the reaction of CO_2 with non-carbonate minerals to form stable mineral carbonates, has proved to be a promising CO_2 sequestration technology. There are several advantages, including a vast natural abundance of the raw minerals, the permanent and benign storage of CO_2 in solid form, and the overall reaction being exothermic, and therefore potentially economically viable (Goff et al., 2000; Herzog, 1997). The theorized reaction equation, where magnesium-rich or calcium-rich minerals, such as olivine (Mg_2SiO_4) and serpentine ($Mg_2Si_2O_5(OH)_4$) react with CO_2 to form geologically stable mineral carbonates, such as $MgCO_3$, is shown below:

$$(Mg, Ca)_x Si_y O_{x+2y} + xCO_2 \; x(Mg, Ca) CO_3 + ySiO_2$$
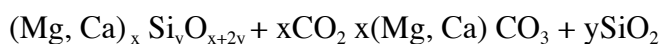
However, there is a primary drawback to mineral carbonation: reaction kinetics. Aging rocks is an example of direct mineral carbonation but the processes are extremely slow of the order of a hundred of thousand years (Herzog, 2002). The processes where the minerals are firstly dissolved before reacting with CO_2 are preferred for faster kinetics (Herzog, 2002). Previous studies have focused on crushing magnesium-rich minerals to very fine particle size $<37\mu$m to promote surface reactions that are known to control most mineral dissolution reactions (O'Connor et al., 2000). Therefore, higher surface areas will facilitate the magnesium dissolution. In addition to the extensive communion of the raw minerals ($<37\mu$m), these mineral carbonation studies also require high partial pressures (>1950psig), and long reactions times (>6 hours) (Maroto-Valer et al., 2003). However, all these operations are very energy intensive and consequently mineral carbonation will only become a viable cost-effective sequestration technology through innovative development of fast reaction routes under milder regimes in a continuous integrated process (Maroto-Valer et al., 2001).

The research group led by Dr. Maroto-Valer from the Energy Institute at Penn State University has developed a novel active carbonation concept (referred here as Phase I), which utilized surface activation to accelerate significantly the reaction rates and efficiencies for forming carbonates from minerals. This concept allows the integration of various synergistic features for the development of a cost-effective sequestration technology, including acceleration of carbonation efficiency without extensive mineral particle communion and without heat treatment. The concept increased the surface area of the serpentine and reacted it with CO_2 at low temperature and pressure (Maroto-Valer et al., 2001). However, the experimental conditions used in Phase I have to be optimized.

Accordingly, the objective of this research was to conduct a parametric study to optimize the operating conditions, including leaching temperature, acid concentration and activation time in order to increase the efficiency of mineral activation process for subsequent CO_2 sequestration. The overall goal of this study was to dissolve the maximum amount of magnesium from serpentine minerals for its subsequent carbonation, while producing high surface silica solid byproducts.

Experimental Procedures

The serpentine sample for this research was from Cedar Hills Quarry, SE Pennsylvania. A concentration of 1.5 M sulfuric acid was used as the best agent to dissolve the magnesium and made it readily available for the carbonation reaction. There were six experiments conducted at various operating conditions, as listed in Table 1. For each experiment, approximately 50 grams of serpentine and 53 grams of sulfuric acid were reacted in a continuous-stirred pitcher for a given period of time. The first five experiments were conducted at 51-52°C while the last experiment was conducted at room temperature (25°C). After the reaction, the acid treated product was filtered and both the liquid and the solid products were analyzed by gas sorption and ICP method.

Table 1: Serpentine activation operating conditions

Experiment	Temp (°C)	Time (hr)
1	51	1
2	51	3
3	52	6
4	51	8
5	51	16
6	25	24

Gas adsorption was used to measure the surface area of the sample. The operation procedure as follows: degassing, or surface cleaning from contaminants such as water and oils, is first carried out by placing the powder sample in a glass cell and heating it under a vacuum; then the clean sample is brought to a constant temperature by means of an external bath; later small amounts of a gas are admitted in steps into the evacuated sample chamber. As more gas molecules are introduced into the system, the adsorbate molecules tend to form a thin layer that covers the entire adsorbent surface. Based on the Brunauer, Emmett and Teller (BET) theory, the number of molecules required to cover the adsorbent surface with a monolayer of adsorbed molecules, Nm, can be estimated. Then multiplying Nm by the cross-sectional area of an adsorbate molecule yields the sample's surface area (http://www.quantachrome.com/SurfaceArea.htm).

Inductively Coupled Plasma (ICP) is a technique to determine the concentration of different elements of a sample. When an atom encounters a high-energy source, the electrons in the atom are excited to specified higher energy levels. When these electrons return to their ground state(s), they emit characteristic wavelengths of radiation. By determining what wavelengths are being emitted, the analyst can determine what

elements are present in the sample. Then by measuring the intensities of these wavelengths and comparing them to those generated by known standards, the concentrations of the different atoms can be determined (www.mri.psu.edu/mcl/techniques/icp.asp).

Results and Discussion

The ICP and BET data for the six experiments are presented in Table 2. As the reaction time increased (experiments 1-5), it is expected that the MgO concentration of the solid product in the first five experiments decreased, since more Mg^{2+} dissolves into the solution. The element Si does not dissolve, hence SiO% in the solid decreased as MgO% in the solid increased. Figure 1 also shows that MgO% in the mineral decreased after the acid treatment and leveled out between 3 hours to 16 hours reaction time. However, experiment four, which was conducted for 8 hours at 52°C showed relative low surface area ($90.55 m^2/g$) and Mg extraction yield percentage (61%) compared to the rest of the experiments. It is suggested that the experiment four needed further investigation and therefore would not be included for all the following analysis. Experiment six which was conducted at room temperature showed less competitive result ($88.5 m^2/g$ of surface area and 53% of Mg extraction yield percentage) and thus was not considered as optimum operating conditions.

Table 2: ICP and BET test results for MgO and SiO concentration, surface area of the solid product and Mg extraction yield%

Exp.	MgO%	SiO_2%	Fe_2O_3%	Surface Area (m^2/g)	Mg extraction yield %
1	30.30	51.30	2.33	91.oo	53%
2	25.60	56.00	1.86	172.40	63%
3	25.10	58.70	1.60	167.80	71%
4	27.00	57.10	1.97	90.55	61%
5	24.20	60.00	1.39	156.50	67%
6	32.00	50.20	2.85	88.50	53%

Figure 1: MgO concentration and SiO concentration in solid product
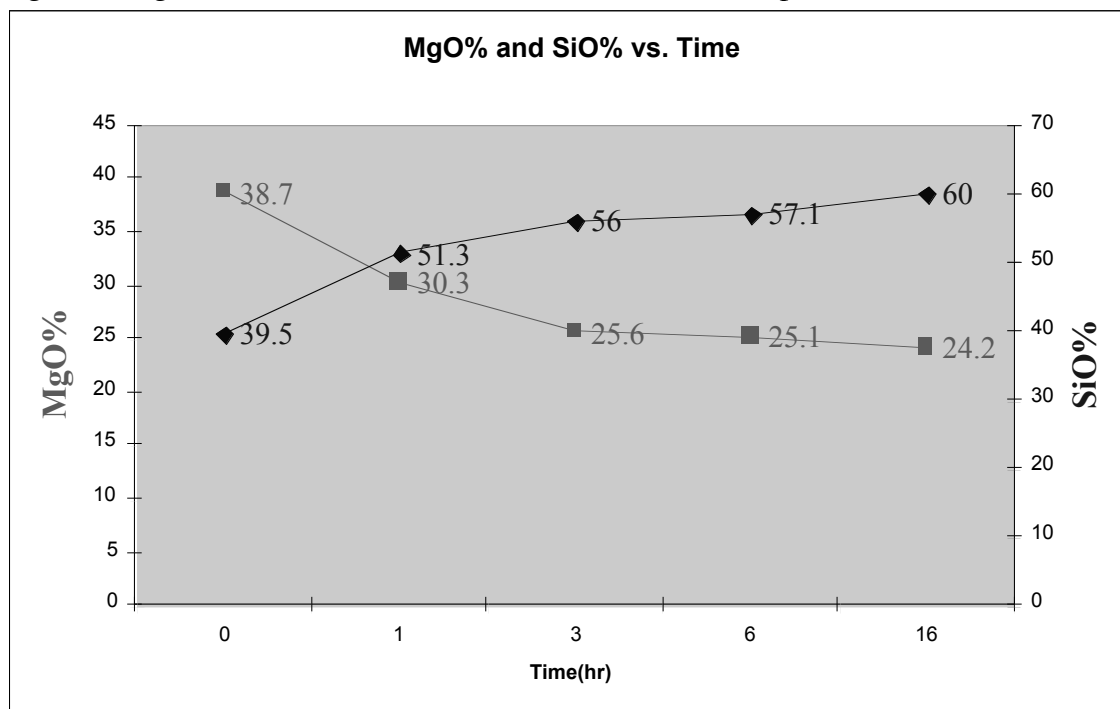
Figure 2 shows the variation of surface area with reaction time. As the reaction time increased, the surface area of sample also increased. This is because the acid treatment dissolved Mg and left more fine pores on the mineral, and thus increased its surface area. The surface area of serpentine sample could reach as high as 172 m^2/g compare to the original sample of 8m^2/g. As reported for the MgO and SiO concentration, the surface area did not change much from 3 hours to 16 hours reaction time. Therefore, 3 hours seem to be the optimum reaction time for the experimental conditions used here (51°C and 1.5 M acid).

Figure 3 shows the relationship between the MgO% in the solid and surface area of solid product. It can be seen that the surface area increased as more Mg dissolved into liquid and left less amount of Mg in the solid.

Figure 2: Surface areas of activated samples

Surface area vs. Time
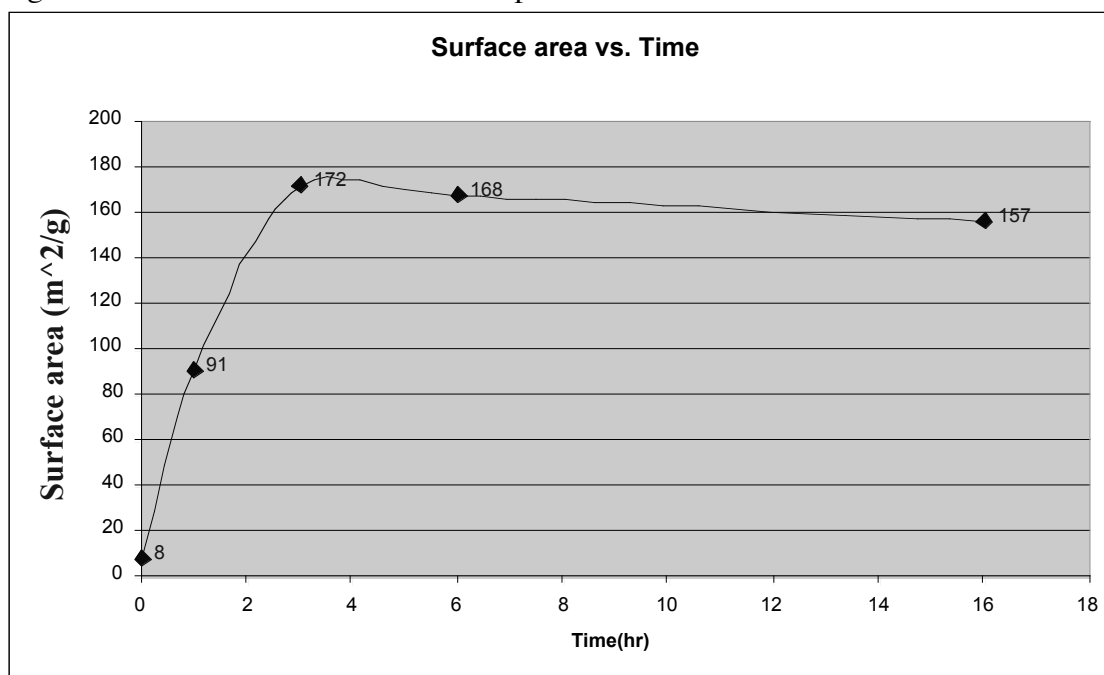
Figure 3: MgO concentration in solid and surface area of activated sample
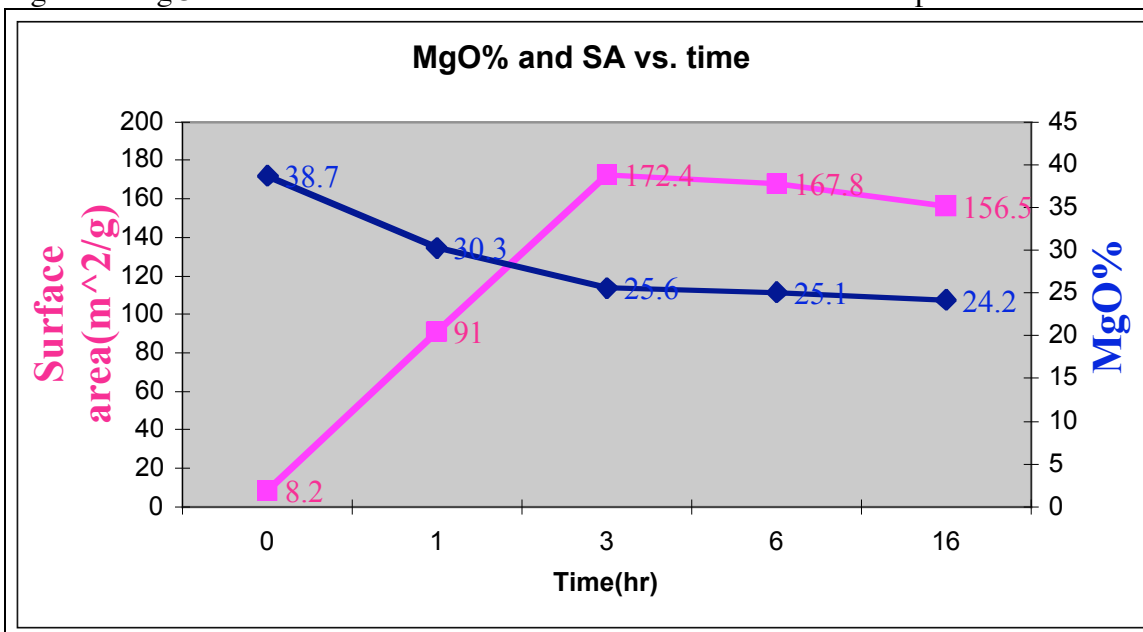
Figure 4 compares the surface area of the parent sample and the sample activated for 3-hours (experiment 3) with previous samples that were produced in Phase I. Although the sample showed much large surface area after treatment of 24 hours at room temperature, the sulfuric acid concentration was significantly higher (6.7M vs. 1.5M), and consequently, increased the cost of the process. The experiment conducted here for only 3 hours with 1.5M of sulfuric acid showed relatively close surface area, around 10% difference, to the experiment conducted for 12 hours using 2.3M of sulfuric acid.

Figure 4: Comparison of the surface areas of the parent sample and its counterparts activated under different conditions
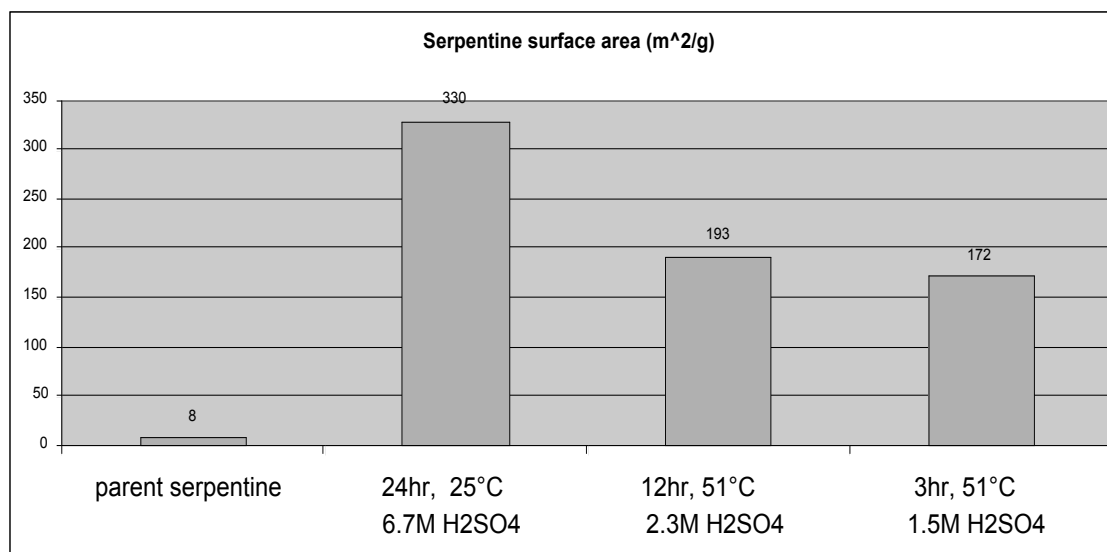
Figure 5 compares the Mg extraction yield percentage among these three experiments. As previously described, the purpose of this research was to dissolve the maximum amount of magnesium from serpentine mineral. The 24-hours experiment showed the highest percentage of Mg extracted, but it used relatively high concentration of 6.7M sulfuric acid and increased the cost. The 3-hours experiment with low acid concentration showed only 5% difference of Mg extraction yield percentage compare to the 12-hour experiment with 2.3 M sulfuric acid. This is a definite improvement considering the fact that short reaction time and low acid concentration were used.

Figure 5: Mg extraction yield%

Magnesium extraction yeild %
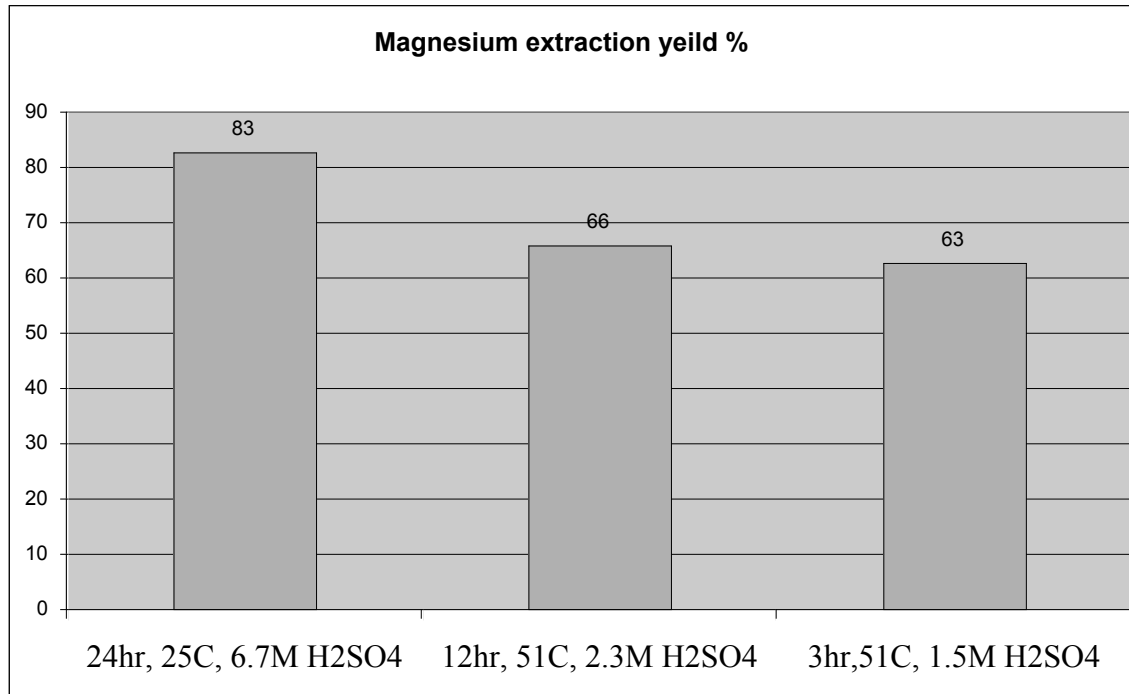
Conclusions

Findings from this summer research conclude that it is possible to increase the surface area of serpentine from $8m^2/g$ of untreated parent sample to $177m^2/g$ after 3 hours sulfuric acid treatment. This research has also optimized the operating conditions for mineral activation and, therefore, has led toward the development of CO_2 sequestration module. The best operating conditions for mineral activation were identified: 3 hours reaction time at 51 C with reduced concentration of 1.5M sulfuric acid.

The mineral carbonation project is still ongoing. After completing the optimization of mineral activation, the extracted magnesium in solution from serpentine will be reacted with CO_2 for aqueous carbonation. For the completion of the project, integration of the activation and carbonation units into a CO_2 sequestration module will be developed.

Acknowledgements

References

Department of Energy / Office of Science / Office of Fossil Energy, 1999, *US DOE Report: Carbon Sequestration: Research and Development.*

Fauth, D.J., Baltrus, J.P., Soong, Y., Knoer, J.P., Howard, B.H. Graham, W.J., Maroto-Valer, M.M. and Andrésen, J.M., Eds: Maroto-Valer, M.M., Soong, Y. and Song, C., 2002, 101-117Goff, F., Guthrie, G., Lipin, B., Fite, M., Chipera, S., Counce, D., Kluk, E., Ziock, H., Los Alamos National Laboratory, LA-13694-MS, 2000.

Goff, F., Guthrie, G., Lipin, B., Fite, M., Chipera, S., Counce, D., Kluk, E., Ziock, H., Los Alamos National Laboratory, LA-13694-MS, 2000

Herzog, H., Drake, E., Adams, E., Report to U.S. DOE, DE-AF22-96PC01257, 1997.

Herzog, H., *Carbon Sequestration via Mineral Carbonation: Overview and Assessment*, Available on-line at http://web.mit.edu/energylab/www/hjherzog/publications.html#co2, 2002

Keeling, C.D., and T.P. Whorf 1998. "Atmosphere CO2 Records from Sites in the SIO Air Sampling Network," in *Trends: A Compendium of Data on Global change*, Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory.

Maroto-Valer, M. M., Zhang, Y., Kuchta, M.E. and Andrésen, J.M., *Integrated Carbonation: A Novel Concept to Develop A CO$_2$ Sequestration Module For Vision 21 Power Plants*, ZEP1.pdf (CD-ROM publication).

Maroto-Valer, M. M., Fauth, D. J., Kuchta, M. E., Zhang, Y. Andrésen, J. M. and Soong, Y., *18th Annual International Pittsburgh Coal Conference*, 2001, Paper 23-01.pdf (CD-ROM).

O'Connor, W.K., Dahlin, D.C., Nilsen, D.N., Walters, R.P., and Turner, P. C., *Proc. 25th Int. Conf. on Coal Utilization and Fuel Systems*, Clearwater, Florida, 2000, CD-rom.

Human Mate Selection: An Exploration of Assortative Mating Preferences

Kristin Liv Rauch, McNair Scholar, Pennsylvania State University

Faculty Research Advisers
Dr. Jeffrey Cohen, Assistant Professor of Cultural Anthropology
Dr. Patricia L. Johnson, Associate Professor of Cultural Anthropology
Pennsylvania State University

This study seeks to identify mate choice preferences for dating, sexual encounters, and marriage. Specifically, preferences concerning demographic characteristics such as age, race, religion, income, and education were of interest. A survey was used to test whether people preferred mates whose characteristics matched their own, a phenomenon known as assortative mating. Past studies on assortative mating have focused on the behavior of individuals (i.e. who they actually choose as mates in their decisions about marriage). This study differs in two ways: one, it focuses on preferences, rather than behavior, and two, it includes questions about dating partners and sex partners in addition to questions about marriage partners. The results suggest that mate choice preferences reflect actual mate choice behavior. In light of this, one might very well ask, "What's love got to do with it?"

Background

Assortative mating addresses the question of who we choose as mates, particularly in terms of marriage. Studies over several decades have revealed a strong tendency for people to select mates who are similar to themselves with respect to a variety of demographic characteristics (see Atkinson & Glass 1985, Glenn 1982, Labov & Jacobs 1986, Rockwell 1976, Tucker & Mitchell-Kernan 1990, and Vera, Berardo, & Berardo 1985 for recent examples). These characteristics include, but are not limited to, age, race, religion, nationality, education, and income. Two major theories explain the tendency toward marital homogamy, or 'like' marrying 'like' (Warren 1968). The first pertains to the effect residential propinquity, or nearness has on our pool of potential mates, and the second involves cultural norms of endogamy, or the choice to marry within a group.

Propinquity

There are two aspects of propinquity: proximity and differential association. Proximity has an effect on our choice of a partner because we tend to live near potential mates. The people with whom we are perpetually in contact are the people who share our spheres of activity (e.g. work, school, church, stores, gyms, restaurants, etc.). It is this close contact with those who live near us that enables romantic interests to develop. Various studies have determined that propinquity does in fact play an important role in our mate choices. Brossard (1932) used the addresses found on 5,000 consecutive Philadelphia marriage licenses to determine that one-sixth of the couples lived within a block of one another, one-third lived within five blocks, and 51.9% lived within twenty blocks of each other (Kephart 1961:268). Similarly, a study in New Haven, Connecticut found that 51.3 percent of the applicants for marriage licenses had lived within a twenty-block radius (Davie & Reeves 1939). Other research, in Columbus, Ohio found that over 50% of married couples lived within one mile at the time of their first date together (Clarke 1952). At least thirteen subsequent published studies have supported Brossard's original findings concerning the relationship between residential propinquity and marital selection (see Katz & Hill 1958 for an overview).

Marital homogamy can be seen as a result of proximity due to the effects of differential association. Not only do we tend to live near our potential mates, but we also have a strong tendency to live near people who are like ourselves. As Richard Udry puts it, "Cities are found to be patterned with people who are alike on important social variables living together and those who are different from one another spatially separated" (1971:185). Our neighborhoods are demarcated along socioeconomic, racial, ethnic, religious, and even educational lines. Rarely will a High School dropout be found living next door to someone with a Ph.D. Unskilled laborers don't generally share their neighborhoods with white-collar professionals. So, if our pool of potential mates consists mainly of those who live near us, and those who live near us tend to be similar to ourselves, then more than likely our mates will reflect our own demographic characteristics. In other words, we will mate assortatively. This is consistent with the findings of Davie & Reeves, who state: "nearly three-quarters (73.6 per cent) of all persons marrying within the city [of New Haven] chose mates residing in the same type of neighborhood. Practically no intermarriage (3.9 per cent) occurred between areas far removed in social, economic, and cultural traits" (1939:517).

Most studies that focus on propinquity to explain assortative mating are out of date. People are quite mobile within their cities in the modern era. The number of young, single people who own or have access to an automobile has risen substantially over recent decades. These mate-seekers have the means to go beyond their demographically homogamous neighborhoods to shop, attend church, work, and play; and therefore they have opportunities to encounter a more diverse pool of potential mates. The number of people who leave their home area to attend college is also larger than it was fifty years ago. For these reasons, Udry argues propinquity may not exert as great an influence over current mate choices as in the past (1971:185).

A second point of interest is that propinquity studies are based on information derived from research on married couples. The data collected is based on the outcomes of decisions (who did I marry), not preferences (who might I marry). This means that homogamous mate choices may reflect spatial constraints on one's pool of potential

mates, but not necessarily reflect a specific *preference* for assortative mating. Perhaps if dissimilar people were included in one's sphere of activity, they would be as likely to be chosen. The results of my study suggest otherwise. When surveyed about their general mate preferences (which were not limited to any particular pool of choices), the subjects responded that, even when given a diverse set of options, they preferred mates who reflected their own characteristics. This lends much greater support to the second theory of why people mate assortatively: "Norms of Endogamy".

Norms of Endogamy

If a group prefers that its members select mates from within the group, it is said to be endogamous. If the social norm is to select mates from outside of the group, then it is exogamous. Whether or not a mate comes from inside or outside of the group depends, of course, on how the boundaries of the group are defined. As Levi-Strauss puts it, " . . . it is merely a question of knowing how far to extend the logical connotation of the idea of community, which is itself dependent upon the effective solidarity of the group" (1969:46). People can divide and categorize themselves along an endless number of lines—religion, race, ethnicity, social status, geographic location, political affiliation, sexual orientation, and so on. As social animals, a great deal of our personal identity is formed according to our membership in certain groups. A sense of security often accompanies group membership, and for this reason people are reluctant to behave in opposition to group norms. Too much rebellion could lead to expulsion from the group, and a consequent loss of security and identity. The "distress resulting from collective hostility" (Levi-Strauss1969: 42) is enough to make most people behave according to the norms of the group. This, of course, includes social rules concerning the choice of an appropriate mate—be they endogamous or exogamous.

Why would a given group hold norms of endogamy? One possible explanation for religious endogamy is offered by Albert Gordon (1964) who states: "Catholics, Protestants, and Jews respond in much the same way. The religious teachings emphasizing the unique, if not the superior and distinctive qualities of each of these religions, clearly urge young people to marry within their group to maintain these special qualities" (p. 69). In other words, members of groups share a belief in the superiority of their own group over other groups. Otherwise, why not belong to some other group? It is easy to see how this logic can extend beyond religion to other types of groups. There are elements in each group that attract and retain its members, in spite of competition from other groups. Selecting a mate from a different group implies settling for less. After all, if your group is indeed special, then members of different groups must be inferior to the members of your own group (there is an element of community-level narcissism in this). Thus, norms of endogamy maintain group integrity.

This same idea manifests itself as ethnocentrism when applied to racial or ethnic groups (Boas 1928). Racial endogamy results from a desire to preserve a group's existence. Although the concept of 'race' in any biological sense is now disputed, many people believe that one's racial identity is a fundamental genetic trait. Following this view, racially mixed marriages will result in the elimination of distinct racial groups, and therefore the loss of identity and security.

Evidence can be found which suggests that American culture does indeed have norms of endogamy, especially concerning certain types of social groups. One readily

available source of information is the Media, which at the very least reveals—and often endorses—cultural expectations for behavior. For example, consider the types of couples that are featured in movies, television shows, and advertisements (both print and TV). When was the last time you saw an interracial couple in an ad for floor wax? How about a sitcom where the husband drives a taxicab and the wife is a scientist? How often do Hollywood films feature couples that have different religious backgrounds? Rarely, and when the Media does feature some type of mixed couple, the story revolves around the scandals caused by their being together. <u>Jungle Fever</u>, <u>Guess Who's Coming to Dinner?</u>, <u>Brooklyn Babylon</u>, <u>Zebrahead</u>, <u>Mississippi Masala</u>, and <u>My Big Fat Greek Wedding</u> are all films in which the plot revolves around the very fact that the couple is not homogamous.[1] Never do we see a couple that "just happens" to be mixed, with the content of the story revolving around some other issue. What this pattern reveals is that mixed marriages (however the mix is defined) are considered abnormal, a novelty, or scandalous. In other words, they go against prescribed cultural norms.

Further evidence for strong norms of endogamy in American culture can be found in recent history. Antimiscegenation laws, forbidding the union of two people from different races, existed from the slavery era until merely 36 years ago. In June of 1967, the U.S. Supreme Court ruled the law unconstitutional, and the sixteen states that still had antimiscegenation laws were forced to abandon them (Wadlington 1966). The fact that norms of endogamy found their way into the law books illustrates the degree to which such norms are enforced.

Rules have also existed within certain churches concerning religious intermarriage. In sixteenth and seventeenth century Europe, there were strict laws forbidding marriage between Catholics and Protestants (Barbara 1989:40). There was also a "previous requirement of the Catholic church, annulled in 1966, that non-Catholic partners in Catholic marriages agree that children are to be raised Catholic" (Johnson 1980:11). This rule ensured that any progeny would be fully committed to the Catholic group, and would not have allegiances to any outside groups. The rule essentially said that a heterogeneous union was permissible, as long as it resulted in the production of more Catholics. Levi-Strauss interprets a 1914 German text on Mormonism ('Der Sexuelle Anteil an der Theologie Der Mormonen') as claiming: "if a girl cannot find a partner possessing the true faith, it is better for her to marry her father, for it is the possession of this faith which is the prime essential in [the Mormons'] definition of a human being" (1969:47). This suggests that religious norms of endogamy are so strict among Mormons (or were at that time) as to supersede incest taboos. Overall, norms of religious endogamy are so important that nearly two out of every three couples belong to the same religion at the time of their marriage (Warren 1970:144). In fact, seven out of every ten men who change their religious affiliation when they get married, change to the same religion as their wife (Warren 1970:144).

Our society is selective about which characteristics are most important for a couple to share. For instance, Hollywood movies commonly feature a leading man who is noticeably older than his love interest (sometimes by a few decades). Sean Connery,

---

[1] As an interesting side note, a Google search for 'films featuring mixed couples' yielded page after page of pornography. Comparing the short list that I was able to think of to this veritable flood of sexually explicit material leads one to believe that homogamy loses its importance in the case of strictly sexual relationships (a point I'll return to shortly).

Harrison Ford, Michael Douglass, Clint Eastwood, and Richard Gere have all played opposite twenty-something, up and coming actresses.[2] In many cases, these actors have married youthful rising stars in real life. Most Americans barely raise an eyebrow over this type of heterogeny. Therefore, our culture's norms of endogamy are differentially valued. Certain lines are more permissible to cross than others. Since all of us belong to a multitude of groups, only our membership in certain of those groups will have a bearing on public approval of a non-homogamous relationship.

One assumption is that people will always mate assortatively, regardless of whether they are choosing dating, sex, or marriage partners. However, due to the effects of endogamous norms, dating and marriage partner preferences should be more homogamous (i.e. endogamous) than casual sex partner preferences. I argue that this is because of the difference between biological and cultural drives behind mate selection.

The pursuit of a casual sex partner involves biological interests, and the act itself usually occurs in private. Marriage, on the other hand, is a cultural behavior and as such invokes social norms and public rules. It involves many people, such as family, friends, and coworkers, and is subject to public scrutiny. Unlike a sexual partnership, which can be secretive and clandestine, marriage is often announced to a community via the newspaper, and traditionally involves a third party's consent. Although many people might be interested in reading it, I have yet to see a newspaper with a section announcing sexual intimacies; and intercourse only requires the consent of the people directly involved in the act. The point is this: norms of endogamy exert a much greater force on our selection of dating or marriage partners than on our selection of casual sex partners. This is because sex involves internal and personal needs and drives, while marriage is a public declaration and involves social concerns.

Methods

A three-part survey was used to collect data for this study (see appendix A). The survey included basic information about all respondents – age, gender, race, religion, education, income, marital status, and sexual orientation. The survey asked a series of likert-type scaled questions about the subject's dating partner preferences, and in subsequent sections asked identical questions about sexual, and marriage partner preferences. The survey was distributed in all the Anthropology classes that were offered in the summer 2003 academic session at PSU, as well as some other disciplines, such as Biology, and Administration of Justice. Additional surveys were obtained from graduate students, members of Penn State's Department of Educational Equity, and undergraduates participating in PSU's summer research programs. A total of 317 surveys were distributed, and 248 completed surveys were collected. The data from these surveys were entered into an Excel file, and analyzed using Excel and SPSS.

Data

The sample obtained was almost evenly divided by gender (47.6% male, 52.4% female), with a mean age of 22 years (90% of the sample was 26 years of age or younger). The vast majority of subjects (93%) had earned a Bachelor's degree or less, and 80% of respondents earned less than $10, 000 annually. Of those responding, 96% reported

---

[2] Rarely is the alternative—an older woman with a younger man—seen. Some exceptions are Harold and Maude, and The Graduate.

being heterosexual, only 7% were married, and 95% of the subjects had no children. In terms of religion, 70.4% identified as being Christian, with Catholic (27.8%), Non-Denominational Christian (17.7%), and Protestant (16.5%) being the most frequent responses. Given the respondent pool, it isn't surprising that the sample was homogeneous. The exception was in terms of racial identification. While a majority of respondents (71%) identified as Caucasian or White, 29.5% identified as a racial minority (percentages do not total 100 because some respondents left the racial identity question blank). I found that 8% of subjects identified as Asian, 14% as Black or African-American, 7% as Hispanic, and .5% as Hawaiian or other Pacific Islander. This is more than double the typical percentage of the racial minority population of the entire Penn State- University Park campus. The most recent data available (2002) place Penn State-University Park's minority enrollment at 12% (Source: Office of Enrollment Management and Administration, Pennsylvania State University).

The tables show average scores and range between 1 and 5 – a 1 indicates strong disagreement, while a 5 indicates strong agreement. Tables 1, 2, and 3 show the racial preferences of respondents with respect to a potential dating, sex, or marriage partner. The tables illustrate that most respondents preferred members of their own race. Scores were typically higher—indicating agreement – for members of their own race on questions asking about who they were most likely to date, have sex with, and marry. Subjects were more selective about marriage partners than sex or dating partners, especially Asian respondents. This is revealed in the greater variability of average scores for marriage preferences, whereas average scores for sex partner preferences are less variable. For example, Table 2 shows that Asian respondents ranked potential sex partners from a high score of 4.17 to a low score of 3.00. These scores cover a smaller span than the Asian rankings for potential marriage partners on table 3, which range from a high of 4.37 to a low of 2.68. In other words, respondents are flexible about who they are willing to have sex with, but are particular about who they will and will not marry. One can see that 'Arab' is an unpopular category, particularly when it comes to marriage; which is quite telling since this runs congruent with the current American sentiment following September 11, 2001 and the war on Iraq.

Table 4 displays educational preferences, and is divided into male and female respondents. A preference for mates with higher education levels corresponds to higher scores. Both genders indicate that they are least particular about the education level of a sex partner, slightly choosier about a dating partner, and most selective about a marriage partner. For instance, the average scores for a potential mate with a G.E.D. show a high of 3.22 (female) for a sex partner, and reach a low of 2.28 (male) for a marriage partner. The four lowest education levels (did not finish High School, G.E.D., High School graduate, and attended college but did not graduate) all display the same pattern – highest scores for sex partners, slightly lower scores for dating partners, and lowest scores for marriage partners. This illustrates that subjects are more willing to have sex with someone who lacked a higher education than they are to marry someone from one of those categories. The scores for Bachelor's, Master's, and Doctoral degrees remain relatively constant across sex, dating, and marriage choices, indicating a willingness to engage in any of these activities with someone who held a degree.

Table 5 shows a similar pattern concerning the income of a potential mate. Again, people are least concerned about the income of a sex partner (as indicated by higher

193

scores in the lower income categories), more concerned about a dating partner's income, and most selective about the income of a marriage partner. This is especially true in the case of incomes lower than $40, 000 a year. For example, scores for potential sex partners range from a low of 3.33 for a person earning less than $10,000 a year, to a high of 4.31 for a person earning more than $200,000 a year. Contrast this narrow range of scores with that of potential marriage partners, which range from a low of 2.30 to a high of 4.43. In other words, subjects are willing to have sex with someone from any income level, but are selective about the incomes of potential spouses.

Tables 6, 7, and 8 show the religious preferences for potential mates. Once again, it is clear that subjects prefer members of their own religion as dating, sex, or marriage partners (this is indicated by the highlighted scores). The pattern concerning the type of mate also reappears. The highest scores for mates outside of the respondents' religion are for sex partners (see table 7). These scores become lower when respondents are asked about marriage partners (see table 8). This indicates that subjects are more willing to have sex with someone from a different religion than they are willing to marry someone from a different religion.

Discussion

There was a strong association between a person's background and the preferences that he or she expressed. If you consider that 71% of subjects were White, 74.8% Judeo-Christian, and 93% had either completed or were enrolled in a Bachelor degree program, it is not a surprise that, when taken as a group, respondents were most likely to choose a White, Judeo-Christian, college-educated partner. The highest ratings from the group as a whole reflect the composition of the sample, which indicates a preference for Assortative Mating. In some cases, informants adamantly expressed this preference. For example, a line was provided on some questions that was labeled 'other', in order to allow subjects to include certain groups that may not have been specifically listed on the survey question. Ten respondents identified themselves as Baptists, and disagreed with marrying, dating, or having sex with anyone from any of the groups listed. Five wrote in the name 'Baptist' on the line marked 'other' and indicated that they would strongly agree with selecting a member of this group as a marriage partner. This shows that some people are very particular indeed about marrying, dating, or having sex with someone like themselves.

The premise of Assortative Mating states that people are particular about the demographic characteristics of their prospective spouses; but the idea of Assortative Mating does not address the choices people make regarding dating or sex partners. This study, however, did consider dating or sex partners. The results indicate that subjects are most particular about who they would marry, less particular about who they would date, and least concerned with the characteristics of a potential sex partner. This difference in the level of concern people have about the characteristics of different types of mates is an important one, and one that previous research on Assortative Mating has overlooked.

Why would this difference exist? I believe it is due to the difference between cultural and biological drives behind the choice of a particular kind of mate, coupled with the relative force that Norms of Endogamy exert over culturally versus biologically driven behavior. Marriage is a cultural behavior, and as such follows prescribed rules and norms. Among these norms is the idea that marriage partners should ideally come

from similar demographic backgrounds. Marriages, and marriage partner choices, are subject to a great deal of public scrutiny, and people enter into these unions with a large amount of social pressure weighing on their decisions. Our marriage choices are subject to the approval of our parents and other family members, our friends and associates, and in many cases our clergy members. A great incentive exists for people to follow social norms when choosing a spouse.

The same norms exist concerning the choice of a sex partner, but there are critical differences. Sexual partnerships can be temporary unions, and they can serve quite limited functions in the lives of individuals. Although people frequently become very involved in the lives of their sex partners, this is generally an indication that the relationship has progressed beyond mere sexual pleasure. Strictly sexual relationships can be concealed from public scrutiny. Their private nature makes it very easy for people to behave in opposition to group norms without consequence. When individuals seek mates specifically for purposes of sexual gratification they are pursuing a personal and biological urge. When people are selecting marriage partners, on the other hand, they are obligated, to some degree, to take the values of the group-at-large into consideration. This reality accounts for the fact that respondents were most particular about the characteristics of a spouse, rather than a dating or a sex partner.

Problems

Some problems arose through the course of this project that would need to be addressed if this research was to be continued. Some issues pertained to the questionnaire itself. For instance, one survey question asks if the respondent would have sex with a homosexual. The problem is that the question does not specify a gender for the homosexual person. This ambiguity makes it difficult to interpret the responses. If a male participant responds that he 'strongly agrees' that he would have sex with a homosexual, does he mean a lesbian woman, a gay man, or both? This distinction is an important one. One other problem with the content of the survey was the use of the term 'asexual.' Participants were given 'asexual' as a choice on a list of possible sexual orientation identities. The most frequently asked question while administering the surveys concerned the meaning of the term 'asexual.' The term was intended to define the sexual orientation of an individual who had no intentions of engaging in sexual activity. Based on the fact that none of the 248 participants claimed to be asexual, perhaps its inclusion was not worth the confusion it caused.

The fact that no one identified as asexual brings up another important point. Were people being honest when answering the survey questions? The survey was anonymous so as to encourage honest responses, however most of the surveys were distributed and completed in class. Perhaps the lack of privacy influenced people's answers to questions about sensitive personal information, such as sexual orientation, marital status, and children. The fact that only four percent of the sample identified themselves as being of an alternative sexuality (homosexual or bisexual) may be a result of the public setting in which the surveys were administered.

Some of these problems may or may not have been avoided if interviews had been used instead of surveys. Face to face interviews allow for clarification on questions or answers that are unclear, but they are time-consuming and create problems of their own – not least of which is a lack of anonymity. The use of surveys allowed much more data to

be collected in a shorter period of time, and enabled a bigger sample. The drawback to using a completely anonymous survey is that ambiguities cannot be resolved because participants cannot be contacted after the survey has been completed. Overall, surveys were the most effective method in this particular study.

Conclusion

   Rules of exogamy are often viewed as facilitating relationships of exchange between groups (Levi-Strauss 1969). One group (A), forbidden from marrying within the group, must choose its spouses from another group (B). Likewise, exogamous group B must choose its spouses from outside of the group. These groups form an ongoing relationship based on the exchange of spouses; group A gets its mates from group B and vice versa. This process of ongoing exchange is known as reciprocity. Reciprocity implies a continuing relationship between two or more groups, and by the same token, refusing to participate in an exchange signifies a refusal to be involved in the relationship (Mauss 1990).

   Exogamy as a means to facilitate relationships with other groups has typically been observed in so-called 'simple' societies. These groups are normally foraging or horticulturalist populations. Characteristically, these populations are homogeneous – they share a common language and culture. Complex societies, such as our own, are comprised of many different groups. "[Where] not only social classes but ethnic groups or castes may be pieced together into larger systems . . . the component 'pieces' are usually endogamous" (Keesing 1975:53-4). Considering all of this information, one can deduce that the component parts of complex societies are refusing to carry on significant relationships with each other by refusing to exchange marriage partners. In other words, groups that hold rigid norms of endogamy are exclusive and isolated, and do not associate or form alliances with outside groups. Rules of endogamy serve to divide complex societies.

   Rigid norms of endogamy often indicate where important social divisions lie. For example, in America, people are not expected to choose a mate from within their state. Someone from New Jersey can marry someone from Ohio without social consequence. We do not hold norms of home-state endogamy. However, we do hold norms of racial, religious, and social class endogamy. This indicates that in America, racial, religious, and, social class boundaries are more socially cogent than geographic boundaries. "So with all of the other variables – the more significant to a society are the differences between two categories of people, the less they intermarry" (Udry 1971:183).

   It is clear that important social boundaries are revealed through mate choices. Mate choice behavior (that is who, in practice, do we marry) obviously reveals social boundaries, but mate choice *preferences* go further to reveal psychological boundaries. If people behave in accordance with social norms, but do not harbor strong preferences in accordance with those norms, this may indicate an imminent shift in the norms themselves. Although people are behaving as they are expected to, their lack of strong preferences for this behavior indicates a weakening of these values. Over time, these people will place less pressure on subsequent generations to choose mates according to norms of endogamy. On the other hand, if preferences are in accordance with norms then people accept and approve of the current distinctions and boundaries, and are likely to

pass these values on to younger generations. The result is that the categories and boundaries will remain stable.

What can be concluded then about social realities as revealed through mate choice preferences? It is clear that norms are quite powerful. The second half of the twentieth century was an era of great social change. Segregation was ended and Affirmative Action served to diversify the workplace. Overall, America made strides toward social and economic equality. Our workplaces, clubs, neighborhoods, and social scenes have all become more diverse – but not our families. Is marriage the final frontier? Segregation in a legal sense is long gone, but on the home front we remain <u>separate</u>, but (arguably) equal.

		Responses to the question, who would you most likely date?						
		White	Black	Am. Ind.	Asian	Hawaiian	Hispanic	Arab
Racial Identity of Respondents	White	4.69	3.62	3.71	3.40	3.80	3.81	3.14
	Black	3.84	4.88	4.28	3.48	4.12	4.48	3.00
	Asian	3.33	3.00	3.14	4.17	3.14	3.14	3.00
	Hispanic	4.46	4.31	4.38	4.00	4.23	4.77	3.77

Table 1: Average Scores for Dating Preferences of Four Main Groups

5 = Perfect Agreement   1 = No Agreement

		Responses to the question, who would you have sex with?						
		White	Black	Am. Ind.	Asian	Hawaiian	Hispanic	Arab
Racial Identity of Respondents	**White**	4.64	3.79	3.76	3.56	3.93	3.91	3.20
	Black	4.00	4.80	4.12	3.52	4.00	4.32	3.12
	Asian	3.50	3.33	3.17	4.17	3.17	3.00	3.00
	Hispanic	4.54	4.38	4.31	3.92	4.31	4.77	3.69

Table 2: Average Scores for Sex Preferences of Four Main Groups

5 = Perfect Agreement   1 = No Agreement

		White	Black	Am. Ind.	Asian	Hawaiian	Hispanic	Arab
Racial Identity of Respondents	**White**	4.62	3.30	3.41	3.25	3.66	3.56	2.90
	Black	3.48	4.79	3.88	2.94	3.67	4.24	2.70
	Asian	3.47	2.68	2.84	4.37	2.84	2.74	2.68
	Hispanic	4.44	4.06	3.94	3.44	3.61	4.72	3.28

Table 3: Average Scores for Marriage Preferences of Four Main Groups

5 = Perfect Agreement   1 = No Agreement

	No H.S.	GED	H.S. grad	Attended college	Assoc. degree	B.A.	M.A.	Ph.D.
Dating								
Male	2.14	2.81	3.42	3.72	4.13	4.57	4.59	4.54
Female	2.32	2.96	3.61	3.89	4.23	4.56	4.51	4.34
Sex								
Male	2.74	3.16	3.63	3.94	4.24	4.41	4.41	4.41
Female	2.76	3.22	3.77	4.08	4.35	4.50	4.53	4.47
Marriage								
Male	1.82	2.28	3.19	3.21	3.87	4.47	4.56	4.53
Female	1.94	2.35	2.95	3.25	3.82	4.22	4.27	4.25

Table 4: Average Scores for Education Preferences by Gender

5 = Perfect Agreement   1 = No Agreement

	Under 10K	10-20K	20-30K	30-40K	40-50K	50-75K	75-100K	100-200K	Over 200K
Dating									
Male	3.07	3.43	3.67	3.96	4.16	4.22	4.17	4.16	4.10
Female	3.06	3.44	3.81	4.08	4.29	4.33	4.32	4.32	4.29
Sex									
Male	3.33	3.62	3.96	4.12	4.19	4.26	4.28	4.29	4.28
Female	3.33	3.53	3.83	4.06	4.22	4.27	4.31	4.31	4.31
Marriage									
Male	2.30	2.56	3.27	3.76	4.16	4.32	4.39	4.38	4.38
Female	2.48	2.80	3.31	3.79	4.07	4.28	4.43	4.42	4.42

Table 5: Average Scores for Income Preferences by Gender

5 = Perfect Agreement   1 = No Agreement

		Average Scores for Responses to the question, who would you most likely date?							
		Muslim	Catholic	Protestant	Non-Den Ch	Jewish	Buddhist	Hindu	Atheist
Religious Identity of Respondents	Atheist	2.91	3.00	3.27	3.27	3.18	3.09	3.09	4.09
	Baptist	2.44	3.33	3.22	4.00	2.67	2.44	2.33	1.44
	Catholic	2.64	**4.56**	4.15	4.00	3.54	2.95	2.72	2.59
	Hindu	2.00	2.00	2.00	2.00	2.00	2.00	3.50	4.00
	Jewish	3.20	3.80	3.80	3.80	**4.20**	3.60	3.60	**4.20**
	Non-Den Ch	2.35	3.58	3.78	**4.15**	3.15	2.44	2.41	2.26
	None	**3.67**	3.67	3.67	3.56	3.67	**3.67**	**3.67**	3.78
	Protestant	2.72	3.72	**4.34**	4.00	3.38	2.76	2.66	2.41

Table 6: Dating Preferences by Religion

5 = Perfect Agreement   1 = No Agreement

			Average Scores for Responses to the question, who would you have sex with?							
			Muslim	Catholic	Protestant	Non-Den Ch	Jewish	Buddhist	Hindu	Atheist
Religious Identity of Respondents		Atheist	3.00	3.27	3.45	3.45	3.36	3.55	3.55	3.82
		Baptist	2.25	3.38	3.43	4.00	2.75	2.50	2.25	1.50
		Catholic	2.87	**4.41**	4.08	3.95	3.67	3.10	2.92	2.92
		Hindu	2.00	2.00	2.00	2.00	2.00	4.00	4.00	4.00
		Jewish	3.80	4.00	4.00	4.00	4.00	4.00	3.80	4.00
		Non-Den Ch	3.00	4.00	4.21	**4.44**	3.68	3.00	3.00	3.08
		None	**4.20**	4.20	4.20	4.20	**4.20**	**4.20**	**4.20**	**4.30**
		Protestant	2.83	3.79	**4.31**	4.00	3.41	2.76	2.76	2.66

Table 7: Sex Partner Preferences by Religion

5 = Perfect Agreement   1 = No Agreement

		Average Scores for Responses to the question, who would you most likely marry?							
		Muslim	Catholic	Protestant	Non-Den Ch	Jewish	Buddhist	Hindu	Atheist
Religious Identity of Respondents	Atheist	2.56	2.94	3.00	3.06	2.94	2.88	2.75	**3.56**
	Baptist	1.82	2.91	3.18	4.09	2.72	1.91	1.60	1.36
	Catholic	2.35	**4.62**	3.93	3.64	3.13	2.48	2.39	2.30
	Hindu	2.57	2.86	2.86	3.00	2.71	**3.71**	**4.43**	3.43
	Jewish	3.09	3.00	3.09	3.09	**4.18**	2.91	2.91	3.00
	Non-Den Ch	2.34	3.59	3.75	**4.34**	3.02	2.50	2.50	2.44
	None	**3.59**	3.53	3.59	3.71	3.59	**3.71**	3.59	3.53
	Protestant	2.09	3.68	**4.57**	4.18	3.02	2.20	2.16	2.23

Table 8: Marriage Preferences by Religion

5 = Perfect Agreement   1 = No Agreement

What is your:

Gender: ☐ Male ☐ Female

Age: _____

Race/ Ethnicity: (you may mark more than one)
☐ White
☐ Black or African American
☐ American Indian or Alaska Native
☐ Asian
☐ Native Hawaiian or other Pacific Islander
☐ Hispanic, Latino, Chicano, or Caribbean
☐ Middle Eastern or Arab
☐ Other, please specify _____

Religion:
☐ Muslim
☐ Catholic
☐ Protestant
☐ Non-denominational Christian
☐ Jewish
☐ Buddhist
☐ Hindu
☐ Atheist
☐ Other, please specify _____

Approximate Yearly Income:
☐ Under $10,000
☐ $10,000 - $20,000
☐ $20,000 - $30,000
☐ $30,000 - $40,000
☐ $40,000 - $50,000
☐ $50,000 - $75,000
☐ $75,000 - $100,000
☐ $100, 000 - $200,000
☐ Over $200, 000

Highest Education Level:
(if you are a student mark the degree you are currently working toward)
☐ did not complete High School
☐ G.E.D.
☐ high school graduate
☐ attended college, did not graduate
☐ associate's degree
☐ bachelor's degree
☐ master's degree
☐ PhD

Sexual Orientation:
☐ Heterosexual
☐ Homosexual
☐ Bisexual
☐ Transgendered
☐ Asexual

Marital Status:
☐ Single- currently dating
☐ Single- in a steady relationship
☐ Single- not dating, not in a steady relationship
☐ Married (please state number of years _____)
☐ Separated
☐ Divorced

Do you have children? ☐ Yes   or ☐ No

In the following section, please rank the listed characteristics according to the likelihood that you would DATE someone who fit that characteristic. For example, if you would be very likely to date someone in that category, mark *strongly agree*. If it is a possibility, mark *agree*. If you are indifferent to the characteristic, mark *neutral*. If it is an unlikely possibility, mark *disagree*. If you would be very unlikely to date someone from that category, mark *strongly disagree*.

(1)Age: Would you date someone in the age-range of:

	strongly agree	agree	neutral	disagree	strongly disagree
(1a) 16-18	☐	☐	☐	☐	☐
(1b) 18-21	☐	☐	☐	☐	☐
(1c) 21-25	☐	☐	☐	☐	☐
(1d) 25-30	☐	☐	☐	☐	☐
(1e) 30-35	☐	☐	☐	☐	☐
(1f) 35-40	☐	☐	☐	☐	☐
(1g) 40-50	☐	☐	☐	☐	☐
(1h) 50-60	☐	☐	☐	☐	☐
(1i) 60-70	☐	☐	☐	☐	☐
(1j) 70-80	☐	☐	☐	☐	☐
(1k) 80-90	☐	☐	☐	☐	☐

(2) Nationality: Would you date someone with a nationality that was:

	strongly agree	agree	neutral	disagree	strongly disagree
(2a) the same as yours	☐	☐	☐	☐	☐
(2b) different from yours	☐	☐	☐	☐	☐

(3) Education Level: Would you date someone with the education level of:

	strongly agree	agree	neutral	disagree	strongly disagree
(3a) Did not complete High School	☐	☐	☐	☐	☐
(3b) G.E.D.	☐	☐	☐	☐	☐
(3c) High School graduate	☐	☐	☐	☐	☐
(3d) Attended college, did not graduate	☐	☐	☐	☐	☐
(3e) Associate's degree	☐	☐	☐	☐	☐
(3f) Bachelor's degree	☐	☐	☐	☐	☐
(3g) Master's degree	☐	☐	☐	☐	☐
(3h) PhD	☐	☐	☐	☐	☐

(4) Religion: Would you date someone whose religion was:

	strongly agree	agree	neutral	disagree	strongly disagree
(4a) Muslim	☐	☐	☐	☐	☐
(4b) Catholic	☐	☐	☐	☐	☐
(4c) Protestant	☐	☐	☐	☐	☐
(4d) Non-denominational Christian	☐	☐	☐	☐	☐
(4e) Jewish	☐	☐	☐	☐	☐
(4f) Buddhist	☐	☐	☐	☐	☐
(4g) Hindu	☐	☐	☐	☐	☐
(4h) Atheist	☐	☐	☐	☐	☐
(4i) Other, please specify	☐	☐	☐	☐	☐

_____

(5) <u>Race/ Ethnicity</u>: Would you date someone with the race or ethnicity of:

	strongly agree	agree	neutral	disagree	strongly disagree
(5a) White	☐	☐	☐	☐	☐
(5b) Black or African American	☐	☐	☐	☐	☐
(5c) American Indian or Alaska Native	☐	☐	☐	☐	☐
(5d) Asian	☐	☐	☐	☐	☐
(5e) Hawaiian or other Pacific Islander	☐	☐	☐	☐	☐
(5f) Hispanic, Latino, Chicano, or Caribbean	☐	☐	☐	☐	☐
(5g) Middle Eastern or Arab	☐	☐	☐	☐	☐
(5h) Other, please specify	☐	☐	☐	☐	☐

_____

(6) <u>Income (per year)</u>: Would you date someone whose yearly income was:

	strongly agree	agree	neutral	disagree	strongly disagree
(6a) under $10,000	☐	☐	☐	☐	☐
(6b) $10,000 - $20,000	☐	☐	☐	☐	☐
(6c) $20,000 - $30,000	☐	☐	☐	☐	☐
(6d) $30,000 - $40,000	☐	☐	☐	☐	☐
(6e) $40,000 - $50,000	☐	☐	☐	☐	☐
(6f) $50,000 - $75,000	☐	☐	☐	☐	☐
(6g) $75,000 - $100,000	☐	☐	☐	☐	☐
(6h) $100, 000 - $200,000	☐	☐	☐	☐	☐

(6i) over $200, 000 ☐-------------☐-------------☐-------------☐-------------☐

(7) <u>Sexual Orientation:</u> Would you date someone whose sexual orientation was:

	strongly agree	agree	neutral	disagree	strongly disagree
(7a) heterosexual	☐	☐	☐	☐	☐
(7b) homosexual	☐	☐	☐	☐	☐
(7c) bisexual	☐	☐	☐	☐	☐
(7d) transgendered	☐	☐	☐	☐	☐
(7e) asexual	☐	☐	☐	☐	☐

(8) <u>Marital Status:</u> Would you date someone with the marital status of:

	strongly agree	agree	neutral	disagree	strongly disagree
(8a) never been married	☐	☐	☐	☐	☐
(8b) separated	☐	☐	☐	☐	☐
(8c) married	☐	☐	☐	☐	☐
(8d) divorced once	☐	☐	☐	☐	☐
(8e) divorced twice	☐	☐	☐	☐	☐
(8f) divorced more than two times	☐	☐	☐	☐	☐

(9) <u>Children:</u> Would you date a person that:

	strongly agree	agree	neutral	disagree	strongly disagree
(9a) has no children	☐	☐	☐	☐	☐
(9b) has children out of wedlock	☐	☐	☐	☐	☐
(9c) has children with an ex-spouse	☐	☐	☐	☐	☐
(9d) has children with a current spouse	☐	☐	☐	☐	☐

In the following section, please rank the listed characteristics according to the likelihood that you would have SEX with someone who fit that characteristic. For example, if you would be very likely to have sex with someone in that category, mark *strongly agree*. If it is a possibility, mark *agree*. If you are indifferent to the characteristic, mark *neutral*. If it is an unlikely possibility, mark *disagree*. If you would be very unlikely to have sex with someone from that category, mark *strongly disagree*.

(10) Age: Would you have sex with someone in the age-range of:

	strongly agree	agree	neutral	disagree	strongly disagree
(10a) 16-18	☐	☐	☐	☐	☐
(10b) 18-21	☐	☐	☐	☐	☐
(10c) 21-25	☐	☐	☐	☐	☐
(10d) 25-30	☐	☐	☐	☐	☐
(10e) 30-35	☐	☐	☐	☐	☐
(10f) 35-40	☐	☐	☐	☐	☐
(10g) 40-50	☐	☐	☐	☐	☐
(10h) 50-60	☐	☐	☐	☐	☐
(10i) 60-70	☐	☐	☐	☐	☐
(10j) 70-80	☐	☐	☐	☐	☐
(10k) 80-90	☐	☐	☐	☐	☐

(11) Nationality: Would you have sex with someone whose nationality was:

	strongly agree	agree	neutral	disagree	strongly disagree
(11a) the same as yours	☐	☐	☐	☐	☐
(11b) different from yours	☐	☐	☐	☐	☐

(12) Education Level: Would you have sex with someone with the education level of:

	strongly agree	agree	neutral	disagree	strongly disagree
(12a) Did not complete High School	☐	☐	☐	☐	☐
(12b) G.E.D.	☐	☐	☐	☐	☐
(12c) High School graduate	☐	☐	☐	☐	☐
(12d) Attended college, did not graduate	☐	☐	☐	☐	☐
(12e) Associate's degree	☐	☐	☐	☐	☐
(12f) Bachelor's degree	☐	☐	☐	☐	☐
(12g) Master's degree	☐	☐	☐	☐	☐
(12h) PhD	☐	☐	☐	☐	☐

(13) Religion: Would you have sex with someone whose religion was:

	strongly agree	agree	neutral	disagree	strongly disagree
(13a) Muslim	☐	☐	☐	☐	☐
(13b) Catholic	☐	☐	☐	☐	☐
(13c) Protestant	☐	☐	☐	☐	☐
(13d) Non-denominational Christian	☐	☐	☐	☐	☐
(13e) Jewish	☐	☐	☐	☐	☐
(13f) Buddhist	☐	☐	☐	☐	☐
(13g) Hindu	☐	☐	☐	☐	☐
(13h) Atheist	☐	☐	☐	☐	☐
(13i) Other,	☐	☐	☐	☐	☐

please specify _____

(14) Race/ Ethnicity: Would you have sex with someone whose race or ethnicity was:

	strongly agree	agree	neutral	disagree	strongly disagree
(14a) White	☐	☐	☐	☐	☐
(14b) Black or African American	☐	☐	☐	☐	☐
(14c) American Indian or Alaska Native	☐	☐	☐	☐	☐
(14d) Asian	☐	☐	☐	☐	☐
(14e) Hawaiian or other Pacific Islander	☐	☐	☐	☐	☐
(14f) Hispanic, Latino, Chicano, or Caribbean	☐	☐	☐	☐	☐
(14g) Middle Eastern or Arab	☐	☐	☐	☐	☐
(14h) Other, please specify	☐	☐	☐	☐	☐

_____

(15) Income (per year): Would you have sex with someone whose yearly income was:

	strongly agree	agree	neutral	disagree	strongly disagree
(15a) under $10,000	☐	☐	☐	☐	☐
(15b) $10,000 - $20,000	☐	☐	☐	☐	☐
(15c) $20,000 - $30,000	☐	☐	☐	☐	☐
(15d) $30,000 - $40,000	☐	☐	☐	☐	☐
(15e) $40,000 - $50,000	☐	☐	☐	☐	☐
(15f) $50,000 - $75,000	☐	☐	☐	☐	☐
(15g) $75,000 - $100,000	☐	☐	☐	☐	☐
(15h) $100, 000 - $200,000	☐	☐	☐	☐	☐
(15i) over $200, 000	☐	☐	☐	☐	☐

(16) <u>Sexual Orientation:</u> Would you have sex with someone whose sexual orientation was:

	strongly agree	agree	neutral	disagree	strongly disagree
(16a) heterosexual	☐	☐	☐	☐	☐
(16b) homosexual	☐	☐	☐	☐	☐
(16c) bisexual	☐	☐	☐	☐	☐
(16d) transgendered	☐	☐	☐	☐	☐
(16e) asexual	☐	☐	☐	☐	☐

(17) <u>Marital Status</u>: Would you have sex with someone with the marital status of:

	strongly agree	agree	neutral	disagree	strongly disagree
(17a) never been married	☐	☐	☐	☐	☐
(17b) separated	☐	☐	☐	☐	☐
(17c) married	☐	☐	☐	☐	☐
(17d) divorced once	☐	☐	☐	☐	☐
(17e) divorced twice	☐	☐	☐	☐	☐
(17f) divorced more than two times	☐	☐	☐	☐	☐

(18) <u>Children:</u> Would you have sex with a person that:

	strongly agree	agree	neutral	disagree	strongly disagree
(18a) has no children	☐	☐	☐	☐	☐
(18b) has children out of wedlock	☐	☐	☐	☐	☐
(18c) has children with an ex-spouse	☐	☐	☐	☐	☐
(18d) has children with a current spouse	☐	☐	☐	☐	☐

In the following section, please rank the listed characteristics according to the likelihood that you would MARRY someone who fit that characteristic. For example, if you would be very likely to marry someone in that category, mark *strongly agree*. If it is a possibility, mark *agree*. If you are indifferent to the characteristic, mark *neutral*. If it is an unlikely possibility, mark *disagree*. If you would be very unlikely to marry someone from that category, mark *strongly disagree*.

(19) Age: Would you marry someone in the age-range of:

	strongly agree	agree	neutral	disagree	strongly disagree
(19a) 16-18	☐	☐	☐	☐	☐
(19b) 18-21	☐	☐	☐	☐	☐
(19c) 21-25	☐	☐	☐	☐	☐
(19d) 25-30	☐	☐	☐	☐	☐
(19e) 30-35	☐	☐	☐	☐	☐
(19f) 35-40	☐	☐	☐	☐	☐
(19g) 40-50	☐	☐	☐	☐	☐
(19h) 50-60	☐	☐	☐	☐	☐
(19i) 60-70	☐	☐	☐	☐	☐
(19j) 70-80	☐	☐	☐	☐	☐
(19k) 80-90	☐	☐	☐	☐	☐

(20) Nationality: Would you marry someone with a nationality that was:

	strongly agree	agree	neutral	disagree	strongly disagree
(20a) the same as yours	☐	☐	☐	☐	☐
(20b) different from yours	☐	☐	☐	☐	☐

(21) Education Level: Would you marry someone with the education level of:

	strongly agree	agree	neutral	disagree	strongly disagree
(21a) Did not complete High School	☐	☐	☐	☐	☐
(21b) G.E.D.	☐	☐	☐	☐	☐
(21c) High School graduate	☐	☐	☐	☐	☐
(21d) Attended college, did not graduate	☐	☐	☐	☐	☐
(21e) Associate's degree	☐	☐	☐	☐	☐
(21f) Bachelor's degree	☐	☐	☐	☐	☐
(21g) Master's degree	☐	☐	☐	☐	☐
(21h) PhD	☐	☐	☐	☐	☐

(22) Religion: Would you marry someone whose religion was:

	strongly agree	agree	neutral	disagree	strongly disagree
(22a) Muslim	☐	☐	☐	☐	☐
(22b) Catholic	☐	☐	☐	☐	☐
(22c) Protestant	☐	☐	☐	☐	☐
(22d) Non-denominational Christian	☐	☐	☐	☐	☐
(22e) Jewish	☐	☐	☐	☐	☐
(22f) Buddhist	☐	☐	☐	☐	☐
(22g) Hindu	☐	☐	☐	☐	☐
(22h) Atheist	☐	☐	☐	☐	☐
(22i) Other, please specify	☐	☐	☐	☐	☐

_____

(23) <u>Race/ Ethnicity</u>: Would you marry someone with the race or ethnicity of:

	strongly agree	agree	neutral	disagree	strongly disagree
(23a) White	☐	☐	☐	☐	☐
(23b) Black or African American	☐	☐	☐	☐	☐
(23c) American Indian or Alaska Native	☐	☐	☐	☐	☐
(23d) Asian	☐	☐	☐	☐	☐
(23e) Hawaiian or other Pacific Islander	☐	☐	☐	☐	☐
(23f) Hispanic, Latino, Chicano, or Caribbean	☐	☐	☐	☐	☐
(23g) Middle Eastern or Arab	☐	☐	☐	☐	☐
(23h) Other, please specify	☐	☐	☐	☐	☐

_____

(24) <u>Income (per year)</u>: Would you marry someone whose yearly income was:

	strongly agree	agree	neutral	disagree	strongly disagree
(24a) under $10,000	☐	☐	☐	☐	☐
(24b) $10,000 - $20,000	☐	☐	☐	☐	☐
(24c) $20,000 - $30,000	☐	☐	☐	☐	☐
(24d) $30,000 - $40,000	☐	☐	☐	☐	☐
(24e) $40,000 - $50,000	☐	☐	☐	☐	☐
(24f) $50,000 - $75,000	☐	☐	☐	☐	☐
(24g) $75,000 - $100,000	☐	☐	☐	☐	☐
(24h) $100, 000 - $200,000	☐	☐	☐	☐	☐
(24i) over $200, 000	☐	☐	☐	☐	☐

(25) <u>Sexual Orientation:</u> Would you marry someone whose sexual orientation was:

	strongly agree	agree	neutral	disagree	strongly disagree
(25a) heterosexual	☐	☐	☐	☐	☐
(25b) homosexual	☐	☐	☐	☐	☐
(25c) bisexual	☐	☐	☐	☐	☐
(25d) transgendered	☐	☐	☐	☐	☐
(25e) asexual	☐	☐	☐	☐	☐

(26) <u>Marital Status</u>: Would you marry someone with the marital status of:

	strongly agree	agree	neutral	disagree	strongly disagree
(26a) never been married	☐	☐	☐	☐	☐
(26b) separated	☐	☐	☐	☐	☐
(26c) married	☐	☐	☐	☐	☐
(26d) divorced once	☐	☐	☐	☐	☐
(26e) divorced twice	☐	☐	☐	☐	☐
(26f) divorced more than two times	☐	☐	☐	☐	☐

(27) <u>Children:</u> Would you marry a person that:

	strongly agree	agree	neutral	disagree	strongly disagree
(27a) has no children	☐	☐	☐	☐	☐
(27b) has children out of wedlock	☐	☐	☐	☐	☐
(27c) has children with an ex-spouse	☐	☐	☐	☐	☐
(27d) has children with a current spouse	☐	☐	☐	☐	☐

Works Cited

Atkinson, Maxine P., and Becky L. Glass
  1985 Marital Age Heterogamy and Homogamy, 1900 to 1980. Journal of Marriage and
        the Family 47:685-691.

Barbara, Augustin
  1989 Marriage Across Frontiers. David E. Kennard, trans. Philadelphia: Multilingual
        Matters Ltd.

Boas, Franz
  1928 Anthropology and Modern Life. New York: W.W. Norton.

Brossard, James
  1932 Residential Propinquity as a Factor in Marriage Selection. American Journal of
        Sociology, September:288-294.

Clarke, Alfred C.
  1952 An Examination of the Operation of Residential Propinquity as a Factor in Mate
        Selection. American Sociological Review, February:17-22.

Davie, Maurice R., and Ruby J. Reeves
  1939 Propinquity of Residence Before Marriage. American Journal of Sociology,
        January:510-517.

Glenn, Norval D.
  1982 Interreligious Marriage in the United States: Patterns and Recent Trends. Journal
        of Marriage and the Family 44:555-566.

Gordon, Albert I.
  1964 Intermarriage. Boston: Beacon Press.

Johnson, Robert A.
  1980 Religious Assortative Marriage in the United States. New York: Academic Press.

Katz, Alvin, and Reuben Hill
  1958 Residential Propinquity and Marital Selection: A Review of Theory, Method, and
        Fact. Marriage and Family Living, February:27-35.

Keesing, Roger M.
  1975 Kin Groups and Social Structure. New York: Holt, Rinehart and Winston.

Kephart, William M.
  1961 The Family, Society, and the Individual. Boston: Houghton Mifflin Company.

Labov, Teresa, and Jerry A. Jacobs
  1986 Intermarriage in Hawaii, 1950-1983. Journal of Marriage and the Family 48:79-88.

Levi-Strauss, Claude
  1949 [1969] The Elementary Structures of Kinship. James H. Bell and John R. von Sturmer, trans. Boston: Beacon Press.

Mauss, Marcel
  1925 [1990] The Gift: The Form and Reason for Exchange in Archaic Societies. W.D. Halls, trans. New York: W.W. Norton.

Rockwell, Richard C.
  1976 Historical Trends and Variations in Educational Homogamy. Journal of Marriage and the Family 38:83-95.

Tucker, M. Belinda, and Claudia Mitchell-Kernan
  1990 New Trends in Black American Interracial Marriage: The Social Structural Context. Journal of Marriage and the Family 52:209-218.

Udry, J. Richard
  1971 The Social Context of Marriage, 2nd edition. Philadelphia: J.B. Lippincott Company.

Vera, Hernan, Donna H. Berardo, and Felix M. Berardo
  1985 Age Heterogamy in Marriage. Journal of Marriage and the Family 47:553-566.

Waddlington, Walter
  1966 The Loving Case: Virginia's Historical Perspective. Virginia Law Review 52(June-Dec.):1189-1223.

Warren, Bruce L.
  1968 A Multiple Variable Approach to the Assortative Mating Phenomenon. Eugenics Quarterly 13:285-290.
  1970 Socioeconomic Achievement and Religion: The American Case. *In* Social Stratification: Research and Theory for the 1970s. Edward O. Laumann, ed. Pp. 130-155. Indianapolis: The Bobbs-Merrill Company, Inc.

Design and Demonstration of a Ferroelectric Optical Switch

Luisa F. Soaterna, McNair Scholar, Virginia State University

Faculty Research Adviser
Dr. Venkatraman Gopalan
Assistant Professor of Material Science
The Eberly College of Science
Pennsylvania State University

Graduate and Post-Doctoral Students
Sungwon Kim
Natalia Malkova
ChengNan
David Scrymgeour
Alok Sharan
Lin Lili Tian
Pennsylvania State University

Abstract: From the telegraph to satellites, the field of telecommunications has grown at an incredible rate. Allowing for huge amounts of data to be transferred at faster speeds than the blink of an eye, thanks to the use of light as the preferred mode of transportation. This paper will examine the interaction between light and a special class of materials called ferroelectrics. Ferroelectrics are materials that have built-in electrical dipoles inside their crystal structure. This report will examine the basic properties of light and of ferroelectrics. Such properties include polarization, total internal reflection, and the electro-optic effect. Then an optical switch will be designed and fabricated based on ferroelectric lithium tantalate, so that it modulates light using a combination of total internal reflection and the electro-optic effect. The performance of this device was tested and compared with theoretical predictions.

INTRODUCTION

Historical Perspective

In the year 1837 Samuel Morse invented the telegraph[1]. This invention only allowed the transmission of data at a very slow rate, only a few Hertz. However, the telegraph itself spread all over the world at a very fast rate. In 1878, Alexander Graham Bell invented the telephone[1], which was handled by an operator. At the same time James C. Maxwell made a very important contribution to the world with his ingenious and simple way of combining four equations into a set of equations called Maxwell's equations. These equations have been applied into many of today's creative inventions. They also

explained clearly how light could be described as an electromagnetic wave. As a result in 1895 Guglielmo Marconi discovered that electromagnetic waves did not need wires to travel, they could easily travel through air, thus came the invention of the radio[1]. In 1940 the use of telecommunications increased exponentially due to the invention of satellite communications. Satellite communications allowed the transfer of data at very fast rates and over a very long range. They improved the way in which people could communicate globally. Finally in 1960, "the laser", a bombshell hit the telecommunications industry[1]. This invention has allowed for the improvement of many of today's communication systems. It allows the use of light as a pathway for communication around the world.

Specific Aims

The specific aims of my research are to understand the interaction between light and materials. Also to learn and manipulate properties of light such as Snell's law, index of refraction, total internal reflection and the electro-optic effect in order to design and fabricate an optical device. Once the optical device is fabricated it will be tested in order to verify and gain support from theory.

Significance

Since 1960, when the first laser was developed the field of optics has grown at an incredible rate. This revolution in optics has had a powerful impact on society; thus, making the use of optical devices an every day necessity. Improvements have been made in global communication, health care service, and manufacturing[2]. In the field of telecommunications, optics has played an enormous role in developing systems around the world that include the use of fiber optics to transmit voice, data and video.
Fiber optics are cables made of glass, each wire is the size of a hair strand capable of carrying one million channels of data. Fiber optics cables have advantages over coaxial cables because they do not erode; they also allow large transmission of data and less electromagnetic interference. The greatest advantage of using fiber optic cables over coaxial cables is the fact that light can be used in order to transmit data. However, once the data reaches its final destination electrical components are used and these slow down the entire process significantly by creating bottlenecks. It is thus believed that such electrical components have reached their full potential and it is time for new devices to play an active role in the telecommunications industry. Some of these devices include optical scanners, switches, modulators and telescopes. These devices will allow the use of light to control and manipulate light.

BACKGROUND

Light

A sinusoidal or harmonic wave is composed by wavelength, frequency and amplitude. Wavelength is the length of a single wave or period, frequency is the number of waves per unit time and amplitude is the maximum disturbance, height, of the wave.

Light has been proven to be an electromagnetic wave with an Electric Field, **E**, and a Magnetic Field, **B**, propagating in a direction perpendicular to both of these fields[3] as shown in <u>Figure 1.1.</u> Electric field is force per unit charge, $E = \dfrac{F}{Q}$ and the magnetic field is force times velocity per unit charge. The following equation best describes the electric field in light in terms of its angular frequency,

$$E = E_0 \cos(kz - \omega t + \delta)n \hspace{3cm} Equation\ 1.1$$

where k represents the direction in which the wave is moving, ω is the angular frequency, t is the time, δ is the direct displacement of the wave and n is the unit vector.
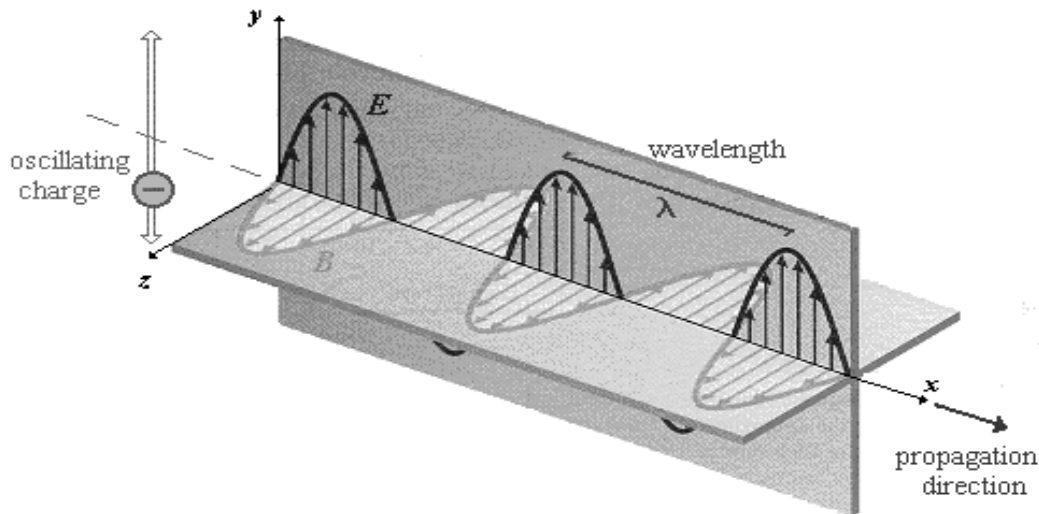

Figure 1.1: The electric field, **E**, travels along the y-axis. The magnetic field, **B**, travels perpendicularly to the **E**, it travels along the z-axis. The wave itself travels along the x-axis; perpendicular to both the **E** and **B**.[4] Light is a transverse wave. The Polarization of the wave is in the direction of the **E.**

Polarizers

Polarization of light is defined as the direction of oscillation of the electric field in the electromagnetic wave. Polarizers are made up of polymers. The polymers are placed in such a way that they allow only certain component of light to pass through. The polymers absorb the other components of light, which do not pass through. If the polarization axis between two polarizers is parallel, certain components of light will be allowed to pass through. If the polarization axis between the two polarizers is perpendicular, all of the light will be absorbed and none of the light will be allowed to pass.

An experiment was conducted with two polarizers (see Figure 1.2). The first polarizer had its optical axis at maximum intensity, the second polarizer or as commonly referred to as the analyzer had a variable optical axis. The optical axis of the analyzer was rotated every ten degrees.
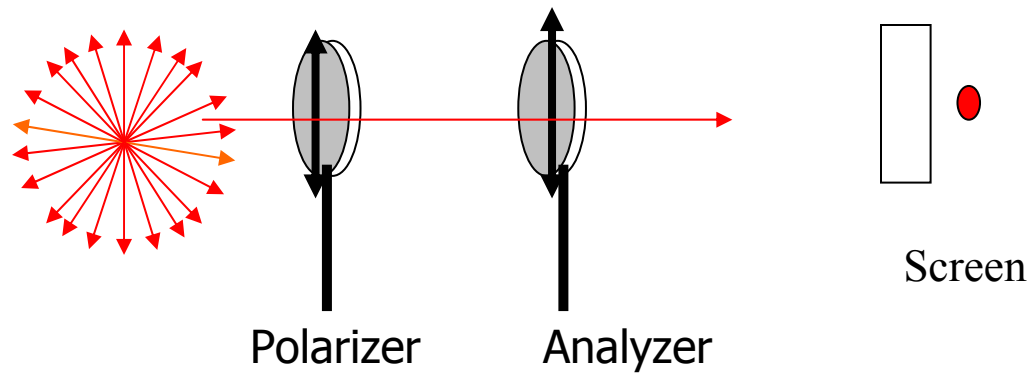
Figure 1.2(a): Experimental set up for experiment between two polarizers, axes are parallel
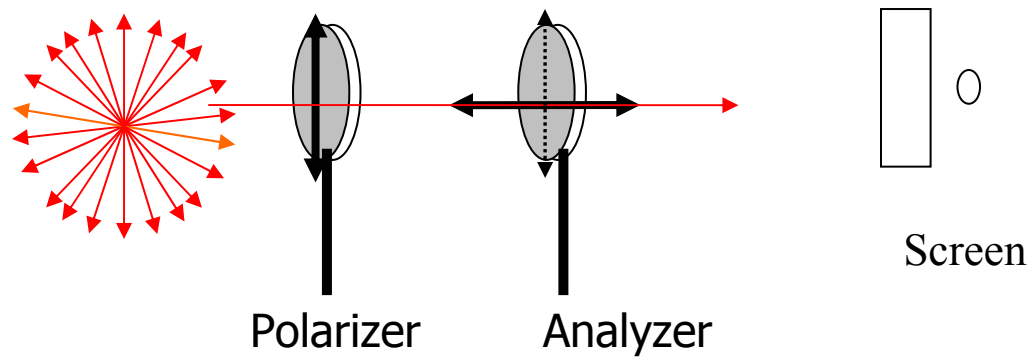

Figure 1.2(b): Experimental set up for experiment between two polarizers, axes are perpendicular.

In <u>Figure 1.2</u> the relationship between the polarizer and the analyzer is clearly seen. When the optical axes of the two polarizers are parallel (Figure 1.2(a)) maximum intensity is attained, and when the optical axes of the two polarizers are perpendicular
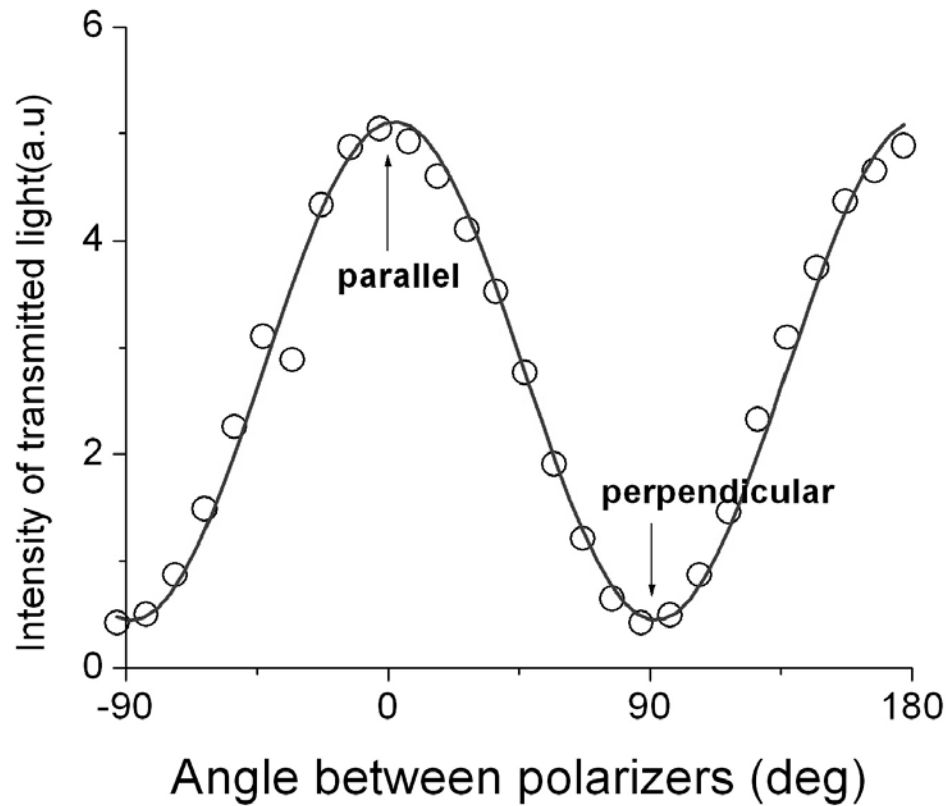
(Figure 1.2(b)) minimum intensity is attained.

Figure 1.3: This graph represents the relationship between two polarizers and light. Circles are experimental points and the solid line is a theoretical fit.

Figure 1.3 depicts the relationship between the angle of the axis of polarization and the intensity or brightness of the light. It is clearly understood that when the axes of polarization for the two polarizers are at zero degrees with respect to one another or parallel, the intensity attained is the highest. Once the axes between two polarizers forms 90 degrees, the intensity attained is the lowest. However, in the figure the intensity never reaches zero completely and this can be attributed to other stray sources of light in the room.

Total Internal Reflection (TIR)

Total internal reflection is a process that occurs when the incidence angle is equal or greater than the critical angle, thus preventing light from being refracted to being completely reflected. According to Snell's Law, there is a relationship between the angle of incidence and the angle of transmittance. This relationship can be best described by the following equation.

$$n_i \sin \Theta_i = n_t \sin \Theta_t \qquad \qquad \text{Equation 1.2}$$

However, under TIR, the angle of incidence $\Theta_i = 90^0$, thus $n_i = n_t \sin \Theta_t$. Solving for Θ gives rise to Equation 1.3 or the equation for the critical angle.

$$\Theta_c = \sin^{-1}\left(\frac{n_i}{n_t}\right)$$

Equation 1.3

NO transmitted light!

$$\theta_c = \sin^{-1}\left(\frac{n_{air}}{n_{water}}\right)$$

$n_{air}=1$ Air

Water

$n_{water}=1.33$

Reflected Ray

$\theta \geq \theta_c$

Incident Ray
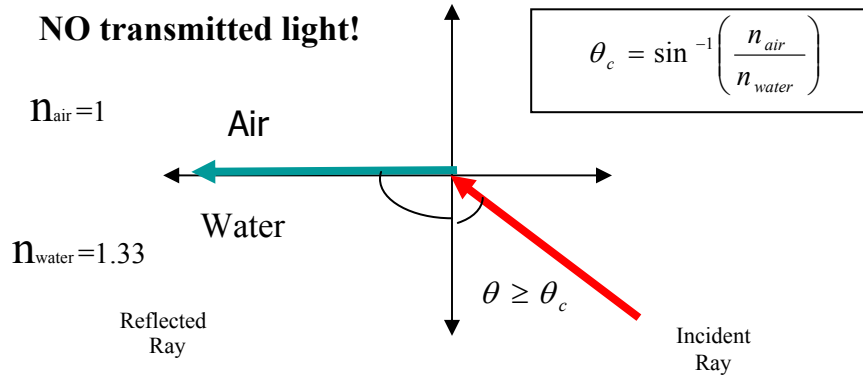
Figure 1.4: Diagram of TIR. Any beams of light equal or greater than the critical angle will not be transmitted.

In <u>Figure 1.4</u> the angle of incidence happens to be equal to the critical angle. Therefore, the beam of light is refracted parallel to the wafer surface. If the angle of incidence is greater than the critical angle then the light will be completely reflected back into the water medium.

Electro-Optic Effect

The electro-optic effect deals with the interaction of an applied electric field, E, and the change of polarity within a crystal after the field is applied, as a result the index of refraction will change. Equation 1.3 can be expanded into the following;

$$\Theta_t = \sin^{-1}\left(\frac{n_t}{n_i}\right) = \sin^{-1}\left(\frac{n_3 - \Delta n_3}{n_3 + \Delta n_3}\right)$$

Equation 1.4

However according to the electro optic effect

Equation 1.5

$$\Delta n_3 = \frac{-1}{2}n_3^3 r_{33} E_3$$

Where n is the refractive index, r is the electro optic coefficient and **E** is the applied electric field. Subscript 3 refers to a specific direction inside the crystal. Therefore (equation 1.3) becomes

$$\Theta_t = \sin^{-1}\left(\frac{2n - n^3 r_{33} E}{2n + n^3 r_{33} E}\right).$$
Equation 1.6
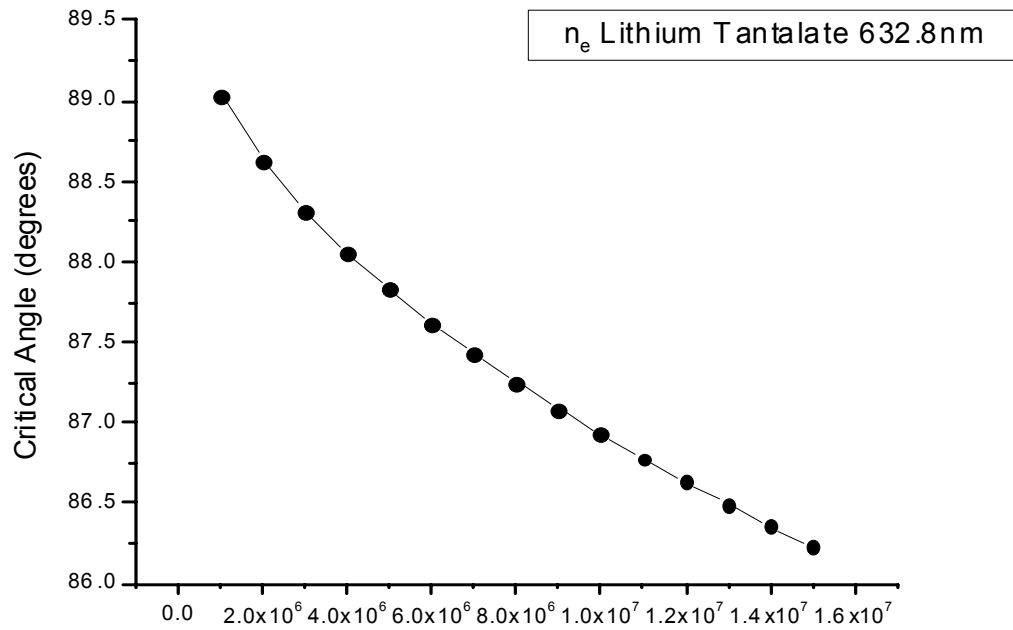
Figure 1.5: Graph describing the relationship between the electric field, **E**, and the critical angle, according to the electro-optic effect in single crystal lithium tantalate.

According to Figure 1.5 it is clearly seen that the applied electric field, E, is inversely proportional to the critical angle. In other words the greater the applied electric field the smaller the critical angle will be. By substituting the applied voltage into equation 1.6, a theoretical prediction can be attained for the critical angle. Then this value can be compared to the experimental value.

Optical Switch

Multiplexing is when telephone calls or data channels are transmitted in a simultaneous process[2]. Optical switches are devices that have been invented in order to perform multiplexing at very fast speeds and with less delay than by using the customary electronic signals. This will allow the processing of millions of signals at a speed of

terahertz. An optical switch will be created using the electro-optic effect, total internal reflection (TIR) and a ferroelectric crystal with a single domain wall.
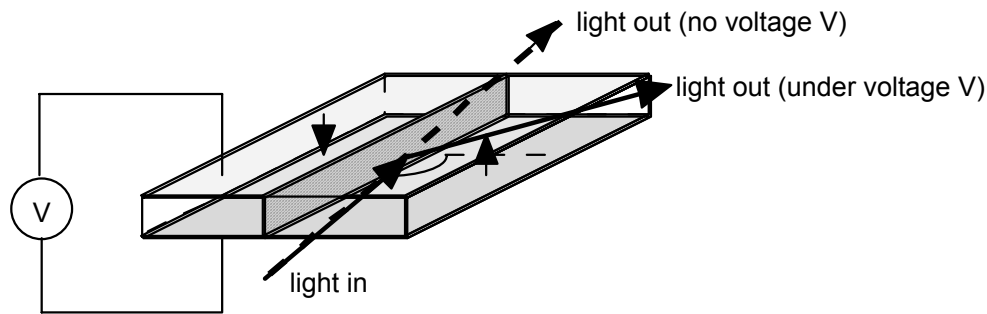

Figure 1.6[6]: Schematic of an optical switch based on a single domain wall interface inside a ferroelectric across which field indiced index change occurs. This is used to totally internally reflect a light beam on the same side of the wall.

In Figure 1.6 the domain wall will be created by applying the necessary voltage in order to create the electro-optic effect and therefore creating total internal reflection within the optical switch.

Ferroelectric Materials

In nature most materials are neutrally charged, meaning that their positive and negative particles are placed in such a way that there is no charge within the actual molecule. However, there are some materials such as ferroelectric materials that are structured in such a way that their positive and negative particles are slightly off centered, therefore creating a dipole or as commonly referred to as a frozen dipole. Optical switches are made with ferroelectric materials because the electro-optic effect is large and fast (GHz). In addition, by applying a voltage, one can change the direction of the polarity thus allowing the creation of a domain wall within the crystal.
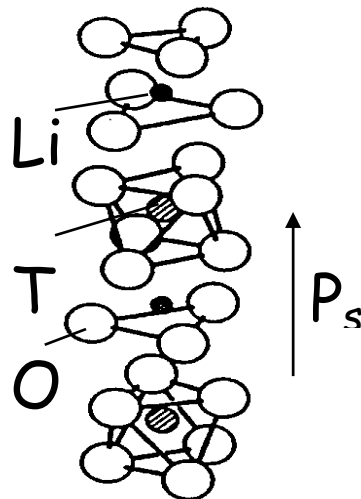

Figure 1.7: Lithium Tantalate, $LiTaO_3$

Figure 1.7 is a picture of the ferroelectric material that will be used for the device, Lithium Tantalate (LiTaO$_3$). In this molecule the Lithium and Tantalum atoms are positively charged and the Oxygen atom is negatively charged. The Oxygen atoms sit slightly lower than the other atoms in the molecule, therefore creating a dipole in the upward direction. Once voltage is applied, the Oxygen atoms will be pushed upward and they will sit slightly higher than other atoms, hence changing the direction of the dipole.

FABRICATION

Small pieces of crystal are cleaned with acetone. In the clean room the samples are carefully placed on top of silicon wafers, which are slightly larger than the sample itself. Following this step both, the wafer and the crystal, are placed on a spinner and three to four drops of 1811 photoresist are added to the sample. Photoresist is a polymer film which breaks up some structural bonds as ultra violet light is applied. This part allows for a very thin and uniform film of photoresist to be placed on the sample as the wafers spins around. Then the sample and wafer are placed on a hot plate and baked at 100° C for two to three minutes. This process is called soft bake. Soft bake is needed in order to dry off any solvent from the photoresist; it improves adhesion and uniformity and it optimizes the light absorbance characteristic of the photoresist[2].
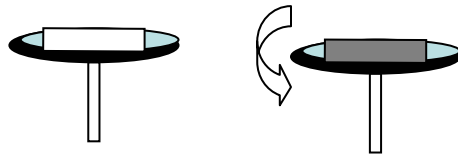
Spin photoresist

Figure 2.1: (a) Sample placed on wafer, photoresist added. (b) Picture of spinner.

Figure 2.1 shows a schematic of how the photoresist was added onto the sample. It also shows a picture of the actual spinner. Now the samples undergo photolithography. In this process the pattern of a mask is developed onto the surface of the sample with the help of ultraviolet light in this case the source is Mercury light. However, the mask does not touch the sample. The sample is placed on a stage, which is then raised to a very finite distance from the mask. This distance is so minimal that the design from the mask is transferred onto the sample with very close detail. This technique is referred to as soft contact. The sample is now developed for 20 seconds in developer and 40 seconds in

deionized water. After all samples have been exposed and developed they are ready for sputtering.

In order to deposit Tantalum on to samples, they are mounted on a silicon wafer and taped with vacuum tape. The tape is placed along the edges of each sample without covering the design that has already been exposed on to the sample during the lithography. 1000 Amperes of Tantalum are deposited on to the sample surface through radio frequency magnetic sputtering.

In order to remove unwanted Tantalum and uncover the design, the samples are soaked on acetone. This step is very delicate because some particles of the Tantalum will lift off by themselves others have to be cleaned by applying gentle pressure with cotton swabs. However, one has to be very careful while rubbing in order to avoid scratching the surface of the actual design. A layer of photoresist in applied once again to the surface of the sample. Now comes the critical step, poling.

During poling a voltage is applied to the surface with the electrode, or layer of Tantalum. This voltage allows for a change to take place in the two areas of the device and facilitating the creation of a domain wall in the middle. A voltage of ~21KV/mm was applied to create the domain wall.

In Figure 2.2 voltage is applied so that the direction of the dipole is changed in the two areas where the electrode was deposit. In the area in the middle no voltage is applied therefore the dipole continues to be in the upward direction. This is how the domain wall was created in the crystal.
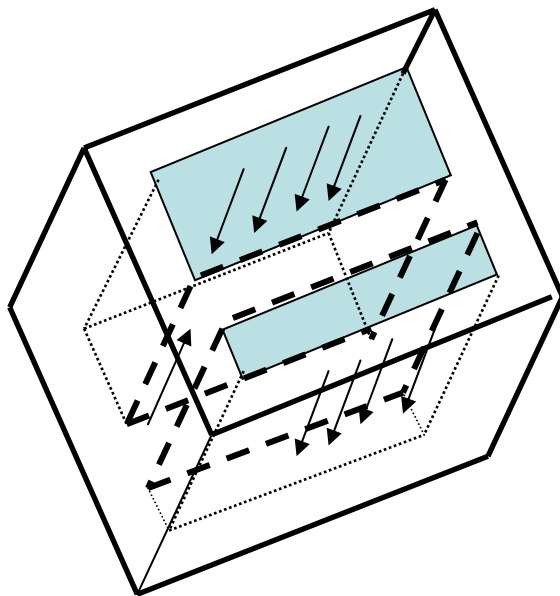
Figure 2.2: Poling of Lithium Tantalate Crystal

Next comes polishing, which is the one step that requires a lot of patience. The samples are placed in a sandwich between pieces of glass and rubber. Each piece of glass has to be cut slightly larger than the actual sample size. Then the entire stack is placed in a

holder. The edges of the sample are polished with diamond polish of 14.5 um down to 0.5 um by moving the holder in the shape of an eight in order to attain a uniform polish.

After the sample has been poled, photoresist in once again applied to both surfaces of the sample and then it is packaged in a holder. Carbon leads are attached to each side of the electrode by using silicone glue, and pulled up through each side of the holder so that it will facilitate for voltage to be added to the device during testing.

Figure 2.3: Sample in holder.

Figure 2.3 shows a picture of the actual holder. The device sits in the middle and each of the carbon leads goes up through the respective sides.

Testing and Conclusion

The device was tested by using a Helium Neon Laser with a wavelength of 632.8 nm and by applying 300V/0.27mm. It took some time to determine the best value for the voltage so that optimal results could be attained. The set up was similar to that of the polarizers, and the holder was placed behind the polarizers so that the laser could go through the device. Then a camera was placed directly behind the holder so that the results could be recorded easily.

The expected value for the critical angle when applying 300 Volts is 87.2538 degrees. Using basic geometry the value of the critical angle came out to be 86.8391 degrees $\pm$ 0.48% error. Therefore our theoretical value comes very close to the experimental value. In <u>Figure 3.1</u> it is clearly seen that our experimental value fits well with our theoretical data.
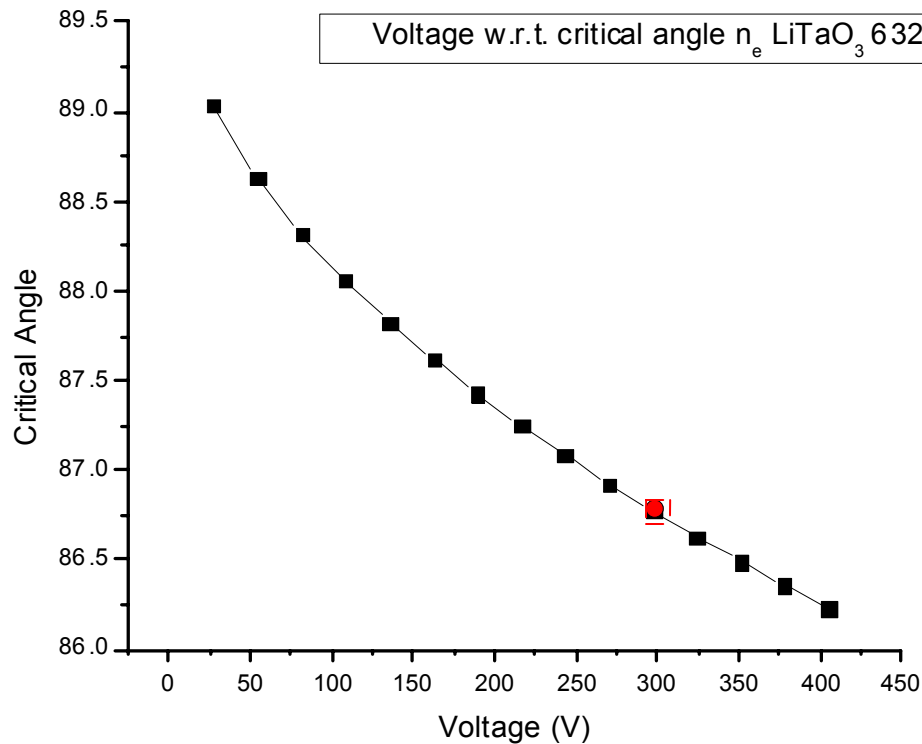
Figure 3.1: Graph describing the relationship between the applied voltage and the critical angle, according to the electro-optic effect in single crystal lithium tantalate. The red circle represents the experimental value surrounded by its' respective error bars.

Only one experimental value was plotted in Figure 3.1 because this was the only value that was accurately measured.
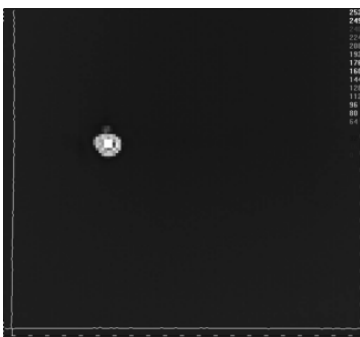
Figure 3.2(a): 0 V

Figure 3.2(b): 300 V

Figures 3.2 (a) and (b) are pictures from video clip of device in motion.

In Figure 3.2(a) the voltage applied was 0 V, meaning that the device is allowing for the light to be transmitted straight through. In Figure 3.2(b) the voltage applied was 300 V, in this case the electro-optic effect took place thus allowing for TIR to occur. In conclusion a device that obeyed optical properties was fabricated and tested. The results attained came fairly close to theoretical predictions and the optical switch behaved the way it was expected to.

233

REFERENCES

1. Pollock, Clifford R.. (1995) <u>Fundamentals of Optoelectronics</u>. Pp 4, Irwin, Chicago.

2. National Research Council. (1998) <u>Harnessing Light,</u> National Academy Press, Washington, D.C. 1998.

3. Griffiths, David J. (1999) <u>Introduction to Electrodynamics 3rd Ed.</u>, Prentice Hall, Upper Sadle River, New Jersey.

4. www.cbu.edu/~jvarrian/122/nsci122.html

5. Hecht, Eugene. (1998) <u>Optics 3rd Ed.</u> Addison-Wesley, Reading, Massachusetts.

6. Gopalan, Venkat. <u>Light in Nonlinear Optical Media.</u> Pennsylvania State University. p. 19, 2001.

Multicultural Education: Using Penn State's Teacher Education Performance Framework to Help Education Students Develop a Multicultural Perspective

Kashra Taliaferro, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. Murry Nelson, Professor of Education
Department Head of Curriculum and Instruction
Pennsylvania State University

Abstract

The make-up of classrooms is changing. According to the National Center for Educational Statistics 37%, or one out of every three students, are considered to be of a racial or ethnic minority background (National Center for Educational Statistics [NCES], 2000). However, nine out every ten teachers are white. By 2020, minority students will make up 44% of students enrolled in public schools. By 2050, minority students will make up 54% of that same population. Enrollment for minority students in both elementary and secondary public schools has increased 73% over the last 25 years, compared to 19% of white students. Minorities enrolled in teacher education programs nationwide, however, account for only 15% of all students, while whites make up over 80%. As recently as Fall 2003, minorities comprise approximately 10% of the 3,210 students enrolled in the College of Education at Penn State. Many education researchers see this as an alarming trend that has harmful implications for the quality of teaching and the teacher shortage in elementary and secondary schools.   Therefore, it is becoming increasingly important to help teachers, regardless of background, to acquire the appropriate attitudes, knowledge, and dispositions to work effectively with students that come from different cultures or class backgrounds (Tiedt and Tiedt, 2002).

Introduction

Background and Motivation

Prospective teachers should be prepared to work with diverse groups of learners. The concept of a multicultural education is one of the major approaches in recent years that have been proposed to address some of the disparities in classroom learning for all students, particularly those of a racial or ethnic minority. It is still the responsibility of teacher education programs to help prepare prospective minority teachers as well to help them draw on their own unique experiences, as they are applicable to classroom teaching and learning.

The *Teacher Education Performance Framework* for the College of Education at the Pennsylvania State University outlines the individual expectations for all teacher education candidates graduating from its program. As a university composed of

predominately white, monolingual students, it is imperative that the College helps their students to develop a more multicultural perspective early in the program at the university.

Challenge Five of the "Assessment of Progress Towards Implementing A Framework to Foster Diversity at PSU: 1998-2003" listed the freshman seminar as one of the college's efforts to support a multicultural curriculum, however each freshman seminar varies widely in the structure in which the material is presented.

Although the language of the framework and the college, itself, do express some means of support for addressing issues of diversity, it is the issue of implementation that raises some concerns. The major goal of this research is to study possible means of further implementing aspects of multiculturalism into the teacher education program by developing a unit on multiculturalism in the freshman seminar course that is required of all students entering the College of Education at Penn State.

Explanation of Terms

Multicultural Education: An educational reform movement whose major goal is to restructure curricula and educational institutions so that students from diverse backgrounds will experience equal educational experiences.

Culture: The set of beliefs, attitudes, and values that define a person or group.

Cultural Responsiveness: The skills, beliefs, and concepts that guide student-teacher interactions in the classroom setting that views cultural diversity as a resource rather than a problem.

Research Questions

The following are the questions that guided this research and that it seeks to answer:
1. How does the performance framework address issue of diversity in terms of teacher preparation?
2. If there is evidence of language regarding issues of diversity within the framework, how is it implemented throughout the curriculum?
3. What possible models exist that can be used to improve the curriculum to address multicultural issues?
4. What approach could best be used to introduce multicultural issues early in the teacher education program?

Current State of Teachers In Terms of Diverse Make-Up

It is clear that the low percentage of minority teacher education students enrolled in prospective teacher programs should be alarming. This disparity between high percentages of students of color and teachers illustrates potential for drastic changes in national education policy. Minority-serving institutions (MSI's), which include historically black, Hispanic-serving, and Tribal Colleges, graduate nearly half of all minorities with teaching degrees (www.ihep.com/Org.php?parm=Press/pr21.htm, 2003).

In this press release from the report "Educating the Emerging Majority," it states that although MSI's get 36% less funding than other universities, they award 46% of teacher education bachelor's degree for African American students, 49% for Hispanics, and 12% for American Indians.

One of the major issues that has added to the low numbers of minority students enrolled in teacher education programs nationwide have to do with low teacher pay. Raising the salaries of teachers to levels of other professions that make invaluable contributions to society was one recommendation that the report stated to help raise the numbers. Another recommendation was to create a national study to identify specific factors that lead to changes in enrollment among minorities in teacher education programs.

Regardless of background, it has become imperative that teachers are prepared to teach racially, ethnically, economically, and linguistically diverse groups of students (Lucas and Villegas, 2002). Lucas and Villegas' proposed approach to creating culturally responsive teachers states that most teachers teach how they were taught. In addition, most of them spend only one semester in a classroom with one teacher before completing their pre-service education. This, in Lucas and Villegas' opinion, is not sufficient in giving prospective teachers the preparation they need to work effectively with diverse groups of learners. It is this issue that drives the vital concept that teacher education programs prepare their teachers, regardless of racial or ethnic background, to develop multicultural perspectives for the classroom.

Their approach also states that universities as institutions should in addition to creating culturally responsive teachers, they should be committed to diversifying the teacher force. Mission statements alone do not bring about change. Therefore, it has become important that colleges and universities make efforts to start rethinking general education courses and electives to help their teacher education students develop a more multicultural perspective (Lucas and Villegas, 2002).

Table I
Distribution of the Student Population and the Teaching Force in Public Schools by Race/Ethnicity, 1995-96. (Villegas and Lucas, 2002)

	White	Of Color
Students	64.8%	35.2%
Teachers	90.7%	9.3%

Source: NCES, 1997b

Lack of Multicultural Perspectives Impact on Public Schools

The educational system, particularly in the area of teacher education, must become more responsive to the needs of the growing population of students in order to stay in line with democratic principles that this nation was built on (Lucas and Villegas, 2002).

The impact of a multicultural curriculum on this population of students can be showed in the many areas that constitute an achievement gap between minority and majority students. These areas include standardized test scores, the overrepresentation of minority students in special education programs, low-achieving instruction group designations, vocational curricular tracks, elevated high school drop-out rates, and low enrollment to postsecondary educational institutions (Educational Testing Service, 1994; Ekstrom and Villegas, 1991; Meier, Stewart, and England, 1989; Oakes, 1985; NCES, 1997c, 1999d, 1999b, 1999c; Tomas Rivera Center [TRC], 1993).

Because the majority of teachers are white, middle class, and English monolingual, it creates a challenge for teachers to design instruction that builds on the background experiences and capitalize on the background experiences of their students (Fuller, 1992; NCES, 1993a; Zimpher and Ashburn, 1992). In support of this idea, *Teaching in the Secondary School* states, "if you are a member of the white majority, you may well live in a world where perspectives of your own group are so dominant that you may fail to recognize that you have a worldview that may differ from that of members of other cultural and ethnic groups" (Armstrong and Savage, 2002).

Table II
Current and Projected Distributions for Five to Nineteen Year Olds in the United States by Race/Ethnicity (Villegas and Lucas, 2002)

Race	1995	2035	2050
White	67.6 %	49.5%	43.9%
Black	14.7%	15.8%	16.0%
American Indian	1 0%	1.1%	1.1%
Hispanic	13.1%	25.8%	30.25
Asian	3/6%	7.8%	8.8%

Source: U.S. Department of Commerce, 1996. Calculations by authors.

The Penn State Teacher Education Program

In *Challenge One* of the "Framework to Foster Diversity: 1998-2003" the college was asked to provide its own description of diversity and how that understanding is demonstrated within the college. It described it as "fostering and respecting diversity in ethnic, racial, and cultural identities as well as in philosophies and points of view along with sensitivity to issues surrounding gender identity, age difference, and challenges associated with disabilities of various kinds. It also stated that it had developed several courses that explicitly addressed issues of diversity.

In *Challenge Two* of the "Framework to Foster Diversity: 1998-2003" the college addresses the question of what visible support it had for diversity. It listed the 80 hours educational work experience requirement, student teaching placements, Urban and Middle Childhood Education degree, the Institution to Foster Multicultural Understanding and other urban programming as some concerted efforts to address the need to understand multiculturalism in the area of education.

The 80 hours of education work experience is required of all teacher education students before entering the college in their junior year. Half of those hours must be with learners of a different cultural, social, or ethnic background than the student. Although this introduces the student to the college's commitment to diversity, the student has a lot of flexibility in how to fulfill that requirement. Upon completion, the student is not given any specific guidelines on how to analyze their experiences in terms of possible future implications for teachers.

Student teaching is another opportunity to expose teacher education students to diversity outside of the university setting. However, universities as institutions should be aware that this portion of the pre-service teacher program has a limited effect on the perspectives of their students who have little or no sensitivity to diversity (Lucas and Villegas, 2002).

The freshman seminar, required by all first-year education students, is a fairly new aspect of the teacher education program initiated in the summer of 1999. The primary goal of the course is to guide education students through the transition from high school to college. The secondary goal is to acquaint first-year students interested in education with the college itself, including contemporary issues in education, multiculturalism on campus, in education, and in the society at-large, in addition to other careers in the education field. The seminar is composed of several sections, usually around seven or eight, which are each offered by different instructors. The instructors have their own unique abilities and expectations for the course, which in terms of multiculturalism presents some additional challenges.

Current Definitions and Challenges of Multicultural Education

"When we open ourselves to learning about the historical perspectives and cultural experiences of other races in America, much of what we discover is incompatible with our image of a free and democratic nation." (Howard, 2002)

As earlier defined, multicultural education is a movement in education reform that is concerned with changing the way schools teach students to include the diverse perspectives that our nation represents. In his third edition of *An Introduction to Multicultural Education* Professor James A. Banks at the University of Washington, Seattle describes this reform movement as the efforts to enable an increased amount of interactions between student and teacher that builds upon the idea of cultural diversity as an enriching educational tool (Banks, 2000).

In his book, *Studies in Comparative Education: Teacher Training and Multiculturalism*, Raul Gagliardi discusses the varying perspectives of multicultural education from around the world. It is not only the United States that is facing issues of diversity in the classroom as we move further into the millennium. He believes that teachers play a critical role in one of the more powerful educational strategies by promoting dialogue between cultures and obstructing the appearance of the phenomena of cultural intolerance (Gagliardi, 1995). Besides media and family, teachers are one of the most important aspect of any child's socialization process. Therefore, it is vital that teachers take this role seriously by helping to create responsible citizens that respect differences.

Understanding pupil's learning difficulties is a fundamental element in any educational activity, particularly when teachers come from a different community than the students (Gagliardi, 1995). Learning style is inherent within multicultural education. Students learn differently. Whether they are hands-on, auditory, or visual learners, it is up to the teacher to make those types of connections with their students. Cultural differences also bring about some new issues. Teachers that misunderstand a student's cultural behavioral style may underestimate that student's intellectual potential and unknowingly misplace, mislabel and mistreat them (Bennett, 2003). The focus should be on how students learn rather than assuming whether or not they are capable of learning.

Historically, there has been little concrete evidence that serious work has been done to address issues of race/ethnicity, class, and language in the teacher education curriculum (Villegas and Lucas, 2002). In some ways, this widens the gap between the white majority, middle-class, English-speaking teachers and their poor or minority students. The large-scale academic underachievement is evidence that the educational, social, economic, and political systems are not working as they should.

Cultural Responsiveness

Often placed under strict curriculum guidelines, teachers get stuck in the mind-set of teaching to a test or teaching to a set of standards. Although these tests and standards have certain expectations, there is still much room for flexibility in terms of instruction. Cultural responsiveness seeks to provide teachers with some guidelines that they can work within to integrate ideas of diversity while using it as a resource rather than a hindrance in the classroom. In Villegas and Lucas' book, they present six strands that are inherently connected which illustrate some concrete ways that teachers can begin to explore some of these concepts. The first three strands examine any assumptions about schools and their relationships to society, poor students of color, and the work of teachers. The second three strands are practical examples of implementing aspects of culturally responsive teaching.

Strand One: Gaining Sociocultural Consciousness
The idea that teacher education has the responsibility of developing specific strategies for learning about student individual and cultural background knowledge and experiences helps teachers to achieve a consciousness that is socially and culturally aware.

Strand Two: Develop an Affirming Attitude Toward Students from Culturally Diverse Backgrounds
This is important in allowing prospective teachers to examine some of their own assumptions and perceptions about other groups to create an awareness of how these may potentially harm their ability to teach students effectively.

Strand Three: Develop a Commitment and Skills to Act As Agents of Change
The main goal of this strand is to get teachers to develop a clear vision of their role and goals as educators. In many ways teachers can be seen as moral actors, but it is

up to the individual to decide how much of a moral actor that he or she chooses to be in the classroom.

Strand Four: Embracing Constructivist Foundations of Culturally Responsive Teaching
Constructivist foundations are composed of the transmission and the constructivist views. The transmission view sees school knowledge as a collection of facts, concepts, principles, and theories that compose the curriculum. The constructivist view perceives beliefs, assumptions, and theories as being shaped by individual experiences.

Strand Five: Learning About Students and Their Communities
To help promote student motivation in the classroom teachers should be aware of their students' lives, how they perceive school knowledge, prior experiences with the subject matter, and their community lives.

Strand Six: Cultivating the Practice of Culturally Responsive Teaching
In a multicultural society, it is the responsibility of educators continuously tailor instruction to individual children in particular contexts.

Although these strands provide some specific contexts within which prospective teachers can be trained, without the support of the institution, they have no bearing. Traditionally, colleges and universities were not designed to promote values associated with diversity or serve diverse populations (Villegas and Lucas, 2002). However, in a diverse society it is vital that prospective teachers gain these types of experiences before entering a classroom of their own where they can put some of them to practice.

The Concept of Self- and Social Construction

As *Strand Two* states, it is important for prospective teachers to grasp their own perceptions and assumptions of other groups to begin the process of self-reflection. Self-exploration is essential to the practice of becoming a culturally responsive teacher. In addition, teachers need opportunities for critical reflection for the variety of pre-service experiences (Villegas and Lucas, 2002). Becoming multicultural can be defined as the process whereby an individual develops competencies of perceiving, evaluating, believing, and doing in multiple ways (Dillard and Ford, 1996). This is the idea behind the process of self- and social construction. These four phases are the driving forces behind the conception of the unit that follows. However, in terms of self-reflection, I will be examining the first two phases.

Phase I: Construction of Self
Constructions of self are the set of basic motivations, values, beliefs, and identities that makes up each individual person. Each person brings a unique history with her or him that influences her or his perceptions of self. These unique histories help create the social interactions in both learning and school context.

Phase II: Deconstruction Through Critical Self-Reflection

The basic idea supporting deconstruction is that self-knowledge is crucial for understanding self as the subject and involves a combination of self-reflection on past events as well as the critical questioning of those experiences in the present. To aid this process, time and place for reflection in addition to the response of others are essential.

This idea supports the dialogic process that takes place in two arenas that include 1) the self and others and 2) self and the reflective self. It is through dialogue with others who may perceive the world differently from ourselves that inner dialogue can reshape those types of generalizations. Sharing personal histories with others enables the recognition of similarities and opens the door for understanding. However, reconstruction of the self begins with the feedback one gets when one acts from one's own personal perspective. Continuing these interactions with others followed up by internal reflection on those interactions cycles of instability and changing perspectives may ensue.

Designing the "Multicultural Unit"

The multicultural unit utilizes the concepts of self-construction, deconstruction, and reconstruction to allow the prospective teacher the opportunity for self-exploration primarily through open dialogue within the setting of the small group discussion and a variety of assessment tools. The approach proposed by Villegas and Lucas on culturally responsive teaching is also used to contribute to the set of small group discussion with guided questions in addition to several classroom activities. The assessments that are used are based on the language within the teacher education performance framework that stresses the need for its students to understand student diversity and how to tailor instruction to the needs of their students.

The freshman seminar course lasts approximately fifteen weeks allowing for great flexibility for the seminar instructor. Because the concept of multiculturalism is already included within the syllabus topic list for the course, it provides easy opportunity for expansion. The unit itself is designed to last about six days. Depending on the flow of the class and what days the course is offered, the unit may last up to two weeks. A list of resources is included in the references section for instructor to consult for further information. In addition, it is encouraged that the instructor would also use the students as resources. The assessment and guided questions come from *Teaching in the Secondary School: An Introduction* by Armstrong and Savage.

"The Multicultural Unit"

Day One: Define several terms relating to multiculturalism (race, ethnicity, diversity, prejudice, culture). Students will also receive a multicultural inventory assessment to complete individually.

Explanation: Depending upon the size of the course, the class may be split into four or five small groups. Together, they will come up with a group definition for each term. This idea is based on the second phase in which receiving feedback from other students and witnessing their reactions aids in the 'awakening' process. There are several inventories that allow students to examine some specific aspects of multiculturalism as they relate to them. For the remainder of the unit, these

242

definitions and inventories will allow the students to reflect on them and their process from the beginning to the end of the unit.

Day Two: Sharing first awareness or a major awareness of difference.

Explanation: Having an experience and processing an experience are two very different things. This activity will also take place in the same small groups that the students were in the first day in which they will continue for the rest of the unit. This will contribute to a comfort level between the members of the group. The conscious awareness of these experiences and the feedback received from other members of the group will aid in the processing of those experiences and what that may mean for them as a future teacher.

Assessment: The students will be given this writing prompt to answer and bring in for the next class period (How are my ideas influenced by the community I lived in and the schools I attended when I was a elementary/high school student?)

Day Three: Present the first three strands of Villegas' and Lucas' approach to culturally responsive teaching as it applies to the importance of doing self-reflection. The brief presentation will be followed by small group discussion. The groups will be given the following questions to answer:
1) How do I think individuals learn in the classroom?
2) To what do I attribute lack of school success of students who are members of certain groups?
3) Where did I get my ideas about good educational practice?

Explanation: These are not easy questions for anyone to answer. The goal of these questions is to get some positive interactions and allow the students to draw on their own personal experiences to better understand themselves.

Assessment: The students will be given the second writing prompt question that will relate to the day's discussion. They will be asked to describe:
1) Two examples of instruction that they believe went really well and why.
2) Their picture of the teacher as a change agent. Do you believe that teachers should take this role upon themselves?

Day Four: Share past learning experiences in which they learned about the "other."

Explanation: Students will discuss in small groups the variety of learning experiences that they have had in their respective elementary or high schools in which they learned about a culture different from their own.
1) What was this topic related to?
2) Was the teacher enthusiastic about the topic?
3) What impression did learning about the topic have on you as a student?

Assessment: Students will be given a third writing prompt which will ask them to:
1) Describe what views they have about what the curriculum (based on subject area) should be?
2) Consider what would they as the teacher could possibly do differently to teach students about the same topic?

Day Five: Students will begin learning about and discussing specific techniques that they

can use to start putting some of the things they have learned into practice. Students will each describe the meaning of their first and last names.

Explanation: This will show students that there is diversity within themselves even if It may not seem obvious. This will also be a good exercise that they can utilize in their own classrooms.

Assessment: Students will be required to find an outside activity that they can use in the classroom to illustrate concepts of multiculturalism that can be easily integrated into their classroom instruction.

Day Six: The students will bring in their examples of an outside activity on the last day of the unit and introduce it to the class. After a two to three minute presentation, the other students will be given the opportunity to ask questions. During the presentation they will have to include the answers to these three questions:
1) What age group is this activity intended for?
2) How is this activity administrated?
3) In what subject areas could this activity be used?

Explanation: The students will be able to use each other as resources. In addition they will see practical applications that they can use.

Assessment: Students will prepare a written multicultural philosophy regarding their role as a teacher in presenting material from various perspectives.

Conclusion

The key to making the teacher education performance framework successful is in its implementation and reinforcement in other courses within the College of Education. Although the language of the framework expresses support for teacher understanding of diversity as it relates to classroom success, there needs to be further concrete support within the college to prepare teachers to understand its importance. It is vital that these types of experiences begin early in the student's experiences within the college. They not only help to reaffirm the college's commitment to diversity, but also show that the college values it as being important to teaching.

The purpose of designing the unit for the freshman seminar course is to provide a practical example of implementation that would help introduce material to education students that they would be able to reflect back on throughout their preparation as teachers in the College of Education at Penn State. However, there are many challenges to implementing these changes. The misconceptions that are often associated with the notion of "teaching diversity" create a major barrier to bringing about institutional changes. Diversity is not meant to be something that should be taught. It should be something that people understand and come to appreciate for its values. As teachers, recognition of diversity, in its many forms, is essential to teaching and preparing students to be fully functional as democratic citizens, therefore this focus on diversity needs the support of its institution. Preparing instructors to incorporate the concepts of the proposed unit into the class curriculum is another major challenge. These are all areas of future research that would be needed before actual implementation can take place.

The potential impact of the unit on the students can have many widespread implications. The opportunity for self-reflection is at the heart of the unit. As Dillard

and Ford noted, the process of self-reflection is vital to contributing to self-awareness before opening the door for possible change. Teaching is a very challenging job that requires teachers to understand their own strengths and weaknesses. If teachers do not take time out to evaluate their own perceptions of students and possibly other teachers, it could greatly hinder their ability to help their students reach their maximum potential in their classrooms. Interacting with other students in education courses will allow these prospective teachers the opportunity to learn from each other. Regardless of whether there is obvious racial or ethnic diversity in a classroom, students may share a wide spectrum of experiences throughout their own education and pre-service teaching experiences that have shaped them in various ways. The specific techniques in the class are also things that they can incorporate into their own classrooms to show that these concepts are very relevant to schools. With the demand placed on teachers to fulfill curriculum guideline, state and national standards, incoming teachers may feel the pressure to stick to strict classroom instruction to maintain their teaching positions. By providing an early introduction to multicultural concepts, teacher education students may develop an early appreciation for diversity that will give them a new perspective as life-long learners. The vision that I have of the unit would be unfulfilled without the development of a curiosity and desire of the students to continue the internal awareness of the self as a life-long learner.

References

Banks, James A. (2002). An Introduction to Multicultural Education. (3rd ed.). Boston, MA: Pearson Education Company.

Bennett, Christine I. (2003). Comprehensive Multicultural Education: Theory and Practice. (5th ed.). Boston, MA: Pearson Education Company.

Children's Defense Fund. (2000). The State of America's Children: Yearbook 2000. Washington, D.C.

Educational Testing Service. (1992). Trends in Academic Progress. Washington, D.C.: National Center for Educational Statistics, Government Printing Office.

Ekstrom, R.B. and A.M. Villegas. (1991). "Ability Grouping in Middle Grades Mathematics: Processes and Consequences." Research in Middle Level Education 15, no 1:1-20.

Grant, C. A. & Sleeter, C. (2003). Making Choices for Multicultural Education: Five Approaches to Race, Class, and Gender. (4th ed). New York, NY: Hermitage Publishing Services.

Kiros, Teodros. (1998). Self-Construction and the Formation of Human Values: Truth, Language, and Desire. Westport, CT: Greenwood Publishing Group, Inc.

Korn, Carol & Bursztyn, Alberto. (2002). <u>Rethinking Multicultural Education: Case Studies in Cultural Transition</u>. Westport, CT: Greenwood Publishing Group, Inc.

Meier, K.J., J. Stewart, and R.E. England. (1989). <u>Race, Class, and Education: The Politics of Second-Generation Discrimination</u>. Madison, WI: University of Wisconsin Press.

National Center for Educational Statistics [NCES]. (2000). <u>The Condition of Education</u>. Washington, D.C.: U.S. Government Printing Office.

NCES. (1997c). <u>NAEP 1996 Mathematics Report Card for the Nation and the States</u>. Washington, D.C.: U.S. Government Printing Office.

NCES. (1997d). <u>NAEP 1996 Science Report Card for the Nation and the States</u>. Washington, D.C.: U.S. Government Printing Office.

NCES. (1999b). "Dropout Rates in the United States, 1996: Table 13—High School Completion Rates and Methods of Completion of 18-through 25-year-olds not currently enrolled in high school or below, by race-ethnicity: October 1998 through October 1999." www.nces.ed.gov/pbus98/dropout/ch06t13a.html.

Oakes, J. (1985). <u>Keeping Track: How Schools Structure Inequality</u>. New Haven, CT: Yale University Press.

Tiedt, Pamela and Iris. (2002). <u>Multicultural Teaching: A Handbook of Activities, Information, and Resources</u>. (6th ed). Boston, MA: Pearson Education Company.

Villegas, A.M. and Lucas, T. (2002). <u>Educating Culturally Responsive Teachers: A Coherent Approach</u>. Albany, NY: State University of New York.

Examining the function of the residue C207 in the cysteine desulfurase reaction of an essential NifS-like protein from Synechocystis PCC6803

Anthony Paul Trace, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Dr. J. Martin Bollinger, Associate Professor
Department of Biochemistry and Molecular Biology
The Eberly College of Science
Pennsylvania State University

I.   Abstract

Biochemical analysis showed the NifS protein is essential to the production of many metalloenzymes, specifically iron-sulfur clusters. Slr0077 is thought to play a vital role in the shuttling of sulfur, through a two-step persulfide cleavage involving several key sites elucidated by comparative analysis of the crystal structures of the WT versus variants of the following genes: Δ 0077, Δ 0387, Δ0704.   It has been shown that the essential Δ0077 mutant contains a unique site – C207. A site directed variant has been made to analyze its role in enzymatic activity. Preliminary characterization using steady state analysis showed a complex behavior of C207S toward cysteine and DTT. C207S binds L-cysteine generating Cys aldimine/Cys ketimine intermediate state and the spectrum slowly returns to that of the native enzyme. Steady state analysis of WT CD-0077 does not show substrate inhibition under the conditions tested. Preliminary data shows that C207S, on the other hand, exhibits a very different behavior, indicating that this residue might contribute to the cysteine desulfurase activity of WT CD-077. Further experiments will investigate the specific role of C207S in the formation and cleavage of the persulfide.

II.   Introduction

Cysteine desulfurases were first discovered by Zheng et. al[1]. These enzymes were found to be involved in incorporating sulfur into iron-sulfur (Fe-S) clusters and other biologically important compounds (***Figure 1***). Homologs of these proteins exist both in prokaryotes and eukaryotes, some of which have been characterized.

The cyanobacterium *Synechocystis* sp. PCC 6803 encodes three NifS-like proteins: Slr0077, Slr0387, and Sll0704. Genetic experiments, performed by the Bollinger lab group and others, have shown that both *slr0387* and *sll0704* can be inactivated without altering the growth of the cell; whereas, only merodiploids can be obtained when attempting to inactivate *slr0077*, implying the essentiality of Slr0077.

Crystallographic studies, performed in conjunction with colleagues from Dennan Lab at MIT, have allowed for the identification and study of the active site of Slr0077. The active site has the cofactor pyridoxal-5'-phosphate (PLP) surrounded by the following residues: H128, K 231C207, K231, H370, and C372. Sequence alignments with characterized cysteine desulfurases suggested that C372 was the nucleophile that generates the persulfide intermediate.

This work will focus on investigating a site-directed variant protein – C207S. It should be noted that the C207 position was selected because of its' uniqueness, having only been identified in Slr0077. Over-expression, purification, and characterization have been performed and reported in this study. The preliminary results suggest C207S might indeed have, as yet, an unidentified role in cysteine desulfurase activity.

III.   Materials and Methods

L-cysteine was purchased from Sigma. Tris hydrochloride, magnesium chloride, and glycerol were also obtained from Sigma. All restriction enzymes were purchased from New England Biolabs (*EcoR1* and *Nde1*). BL21 (DE3) Star cells were purchased from Invitrogen.

B. Overexpression of C207S:

C207S plasmid was transformed into BL21 (DE3) Star cells. Starter cultures were started using individual colonies that were then used to inoculate 100mL LB Rich media with appropriate antibiotic, Ampilicillin (10μL of 150mg/mL stock). Cultures were grown in LB Rich broth (3.5% tryptone, 2.0% yeast extract, and 0.5% sodium chloride) containing 150mg/L ampicillin until an OD_{600} of 0.5-0.8, and then induced by addition of 200μM IPTG overnight (20-24 hours) at 16-20°C. Cells were spun down (6000rpm for 15 minutes) and stored in liquid nitrogen until further use.

C. Protein Purification of C207S:

Cells were resuspended in 5mL/g of 50mL Tris/HCl buffer (pH=7.8) with 0.005% PMSF (protease inhibitor) and allowed to thaw on ice. To lyse, we then French Press cells at 16,000psi and cellular debris was collected via centrifugation (16,000g for 15 minutes). A final centrifugation (12,000g for 10 minutes) yielded a protein rich supernatant, which was pelleted after ammonium sulfate precipitation. The pellet was resuspened in 50mL Tris/HCl buffer (pH=7.8) and transferred to a MWCO dialysis bag of 12-14 kDa and dialyzed for 4-6 hours. Final concentration of C207S protein was calculated based on final volume of solution in the dialysis bag.

MonoQ column (dimensions 10/10 1cm x 10cm, 8mL total volume) was equilibrated with 50mL Tris/HCl buffer (pH=7.8) and 4mL of protein solution was eluted at a time using FPLC. Fractions were collected and analyzed for (protein absorbance) by scanning from 240-600nm, specifically looking for absorbances at 280nm (protein peak) and 425nm (PLP cofactor peak).

D. Spectral analysis of the C207S upon addition of L-cysteine:

Varying concentrations of L-cysteine were added to known concentrations of C207S and the complexing ability was analyzed. Spectral data were collected and recorded using the HP8453 spectrophotometer as a function of absorbance versus wavelength (*Figure 2*). Using this data, the rate (v/Et) was determined and plotted as a function versus cysteine and DTT concentration (*Figure 3*).

E. Assay to Measure Cysteine Desulfurase Activity:

An assay was developed in order to quantify sulfide production by the action of cysteine desulfurase via the radioactive isotope S-35* incorporated into the cysteine substrate ($[^{35}S]$-cysteine).

Using 500μL Sodium Hydroxide (NaOH) in a cylindrical tube, a reaction mix (250mM cys, 100mM Hepes buffer – pH=7.8) is combined in a separate cup with varied concentrations of DTT (0.05,0.1, 0.5, 1, 2, 4, 5, 10, 50, 100mM) and the appropriate enzyme (C207S). The enzymes is allowed to react for 1 minute, after which, the reaction is quenched with 0.72N sulfuric acid (H_2SO_4) thereby halting any further enzymatic activity. The reaction chambers are left for approximately 1 hour with gentle shaking in order for the H_2S gas, produced as a result of enzyme activity, to react with the NaOH in the collection tube by the reaction. The apparatus is pictured in *Figure 4*.

$$H_2S + NaOH \rightarrow NaS^* + H_2O$$
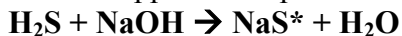
A Scintillation counter was then used to quantify the ^{35}S present in unreacted substrate and resultant product in each of the two chambers: reaction chamber (unreacted sulfur substrate) and collection chamber (reacted sulfur product. Using these numbers, the enzyme's reaction rate is calculated. Reaction rates, variant versus wild type, are compared to determine role, if any, of C207 in cysteine desulfurase activity.

IV.   Results

A. Expression and purification of C207S:

Over-expression of C207S was achieved using the BL21 (DE3) Star cells. The protein was purified by a series of three columns: Q-Sepharose, Mono Q, and finally Superose-12. The protein purification was monitored throughout, via SDS gel electrophoresis (*Figure 5*). The C207S variant protein appears as a 46.8KD protein band on SDS-PAGE gel. Approximately 3 mg of pure C207S was obtained per gram of cells, after the final column purification.

B. PLP-Binding analysis:

Pyridoxal-5'-phosphate (PLP) cofactor has a characteristic bright yellow color that allowed for easy identification of our protein throughout the purification process. Spectral analysis of C207S showed two main peaks: a protein absorption maxima at 280nm and the PLP cofactor absorption maxima at 425nm (*Figure 6*).

C. Spectral changes associated with addition of L-cysteine to C207S:

The internal aldimine of the C207S absorbs at 425nm, while the spectral peak at 342nm indicates the complex formed upon addition of the substrate, L-cysteine, which results in the formation of a cysteine ketimine intermediate. It should also be noted a dramatic decrease in aldimine is observed initially upon addition of L-cysteine but the peak slowly returns to native state. Addition of stoichiometric amounts of L-cysteine results in the formation and eventual decay of the complex noted at 342nm concomitant with reduction of subsequent increase in the 425nm peak. Owing to this complex behavior determination of Kd for cysteine was not possible.

D. Steady State behavior of C207S:

WT CD-0077 does not demonstrate typical steady-state kinetic behavior. Preliminary data with C207S shows that this variant shows a typical steady state behavior pattern with varying cysteine concentration and DTT concentration. C207S has a turnover number of 5min^{-1} with cysteine as substrate and 3.5min^{-1} with DTT as substrate.

V.   Discussion

The protein analyzed within this study is a member of the NifS family of proteins that is involved in sulfur metabolism. This supports the hypothesis that the enzyme complex serves mainly as a sulfur shuttle, providing necessary sulfur to be transferred from the dietary amino acid cysteine to other essential molecules, such as: biotin, lipoic acid, molybpterin, 4-uridine, and iron-sulfur clusters (*Figure 1*).

WT CD-0077, as mentioned earlier, does not demonstrate typical steady state kinetics. We propose that the reductant performs the characteristic C-S bond cleavage (in the cys-ketimine intermediate) at higher cysteine concentrations. It appears in the C207S however, that the cys-ketimine becomes inaccessible to the reductant.

The role of C207S is still unclear. However, the results argue that the residue might have an important contribution towards the altered mechanism in WT CD-0077 setting it apart from other characterized cysteine desulfurases. Further investigations addressing the role of C207 in cysteine desulfurase activity is being presently conducted.

VI. References

1.) Zheng L, White RH, Cash VL, Jack RF, Dean DR. Cysteine desulfurase activity indicates a role for NIFS in metallocluster biosynthesis. Proc Natl Acad Sci USA. 1993 Apr 1;90(7):2754-8.

2.) Zheng L, White RH, Cash VL, Dean DR. Mechanism for the desulfurization of L-cysteine catalyzed by the nifS gene product. Biochemistry. 1994 Apr 19;33(15):4714-20.

3.) Flint DH. Escherichia coli contains a protein that is homologous in function and N-terminal sequence to the protein encoded by the nifS gene of the Azotobacter vinelandii and that can participate in the synthesis of the Fe-S cluster of dihydroxy-acid dehydratase.

4.) Jaschkowitz K, Seidler A. Role of a NifS-like protein from the cyanobacterium Synechocystis PCC 6803 in the maturation of FeS proteins. Biochemistry. 2000 Mar 28;39(12):3416-23.

5.) Mihara H, Kurihara T, Yoshimura T, Esaki N. Kinetic and maturational studies of three NifS homologs from Escherichia coli: mechanistic difference between L-cysteine desulfurase and L-selenocysteine lyase reactions. J Biochem (Tokyo). 2000 Apr;127(4):559-67.

6.) Urbina HD, Silberg JJ, Hoff KG, Vickery LE. Transfer of sulfur from IscS to IscU during Fe/S cluster assembly. J Biol Chem. 2001 Nov 30'276(48)44521-6.

Figure 1:

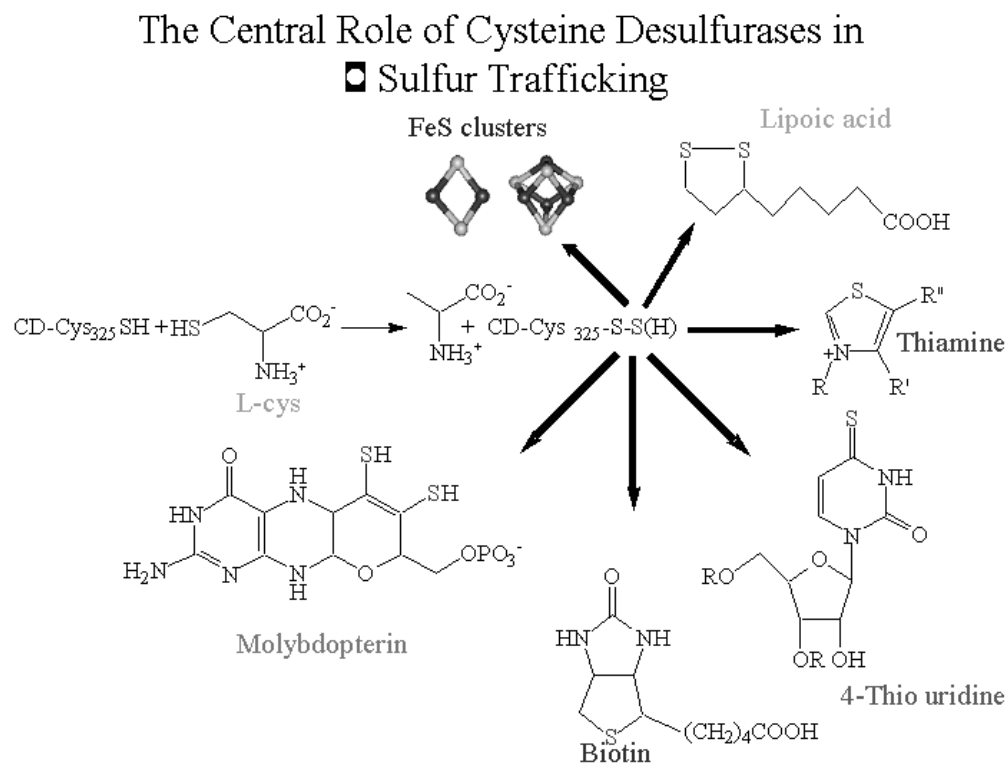The Central Role of Cysteine Desulfurases in Sulfur Trafficking

Figure 1 – This figure illustrates the sulfur trafficking in the body. This relates the possible functions of the Nif-like enzymes, defining their significance in the body.
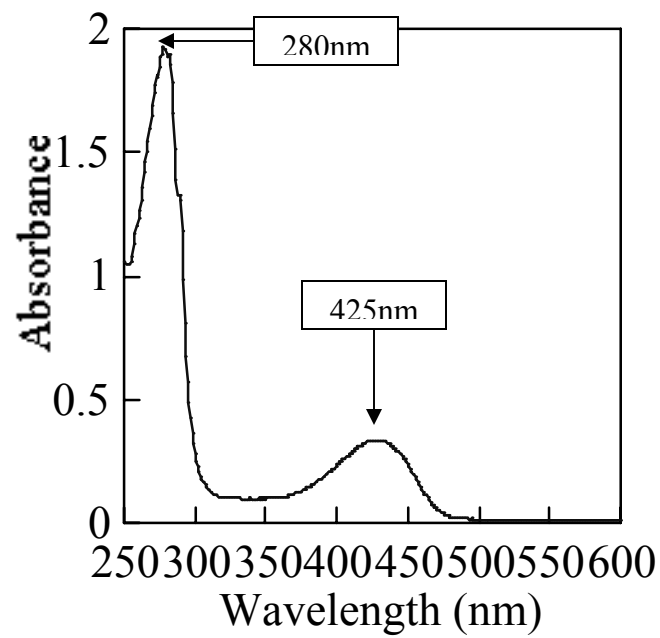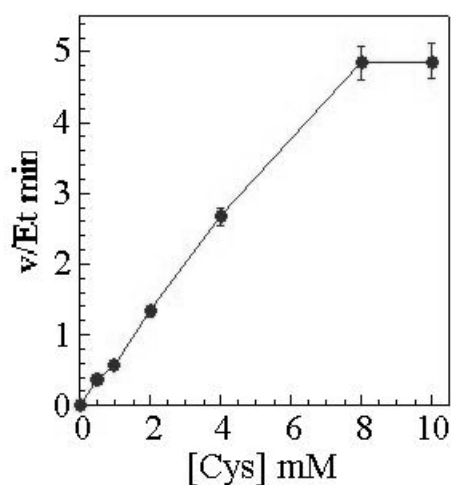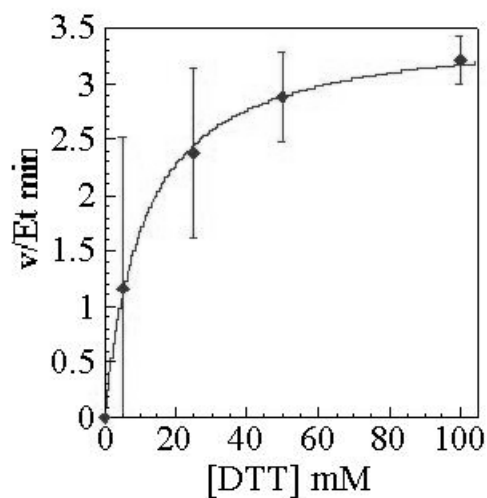
Figure 2:

Figure 2 – This figure illustrates the purification of the C207S protein. Notice the protein peak (~280nm) and the PLP cofactor peak (~425nm).

Figure 3:

Dependence of Cysteine desulfurase Activity of C207S on Varying Concentrations of Cysteine and DTT

This figures illustrates the cysteine desulfurase activity with a maximum turnover number of 5/minute.

This figure illustrates the dependence of cysteine desulfurase activity of C207S variant on varying DTT concentrations. DTT (dithiothreitol) acts as a reductant that cleaves the persulfide.
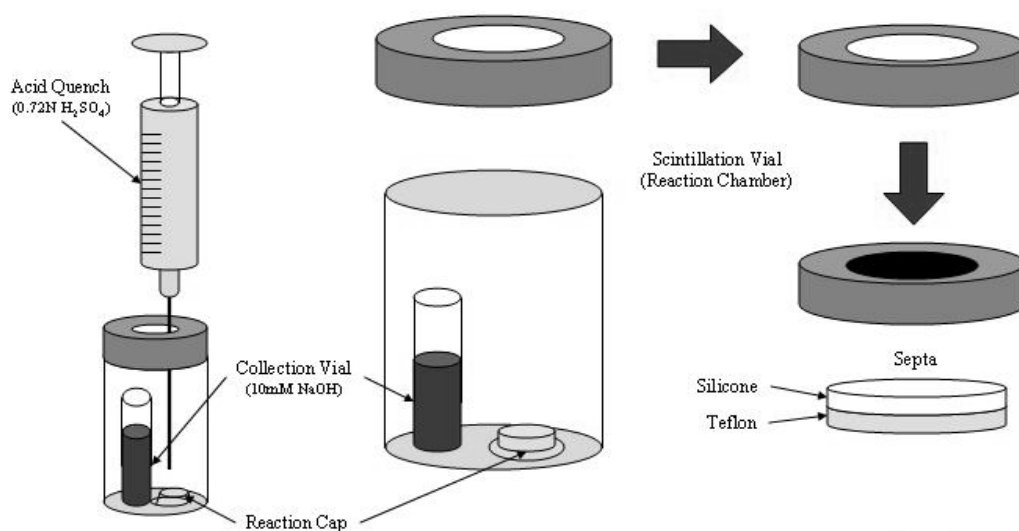
Figure 4:

Novel Cysteine Desulfurase Activity Assay

Figure 4 – This figure illustrates the novel cysteine assay used to monitor the desulfurase reaction and activity.
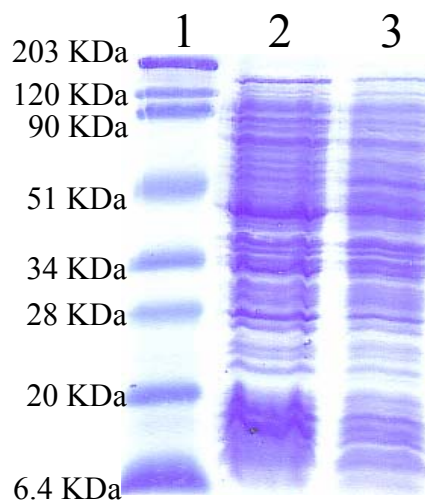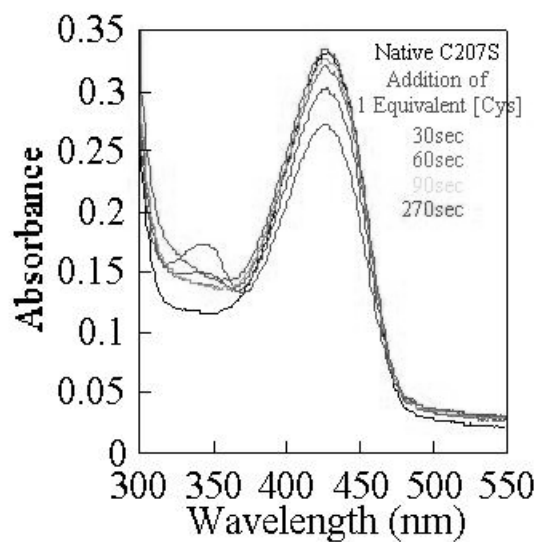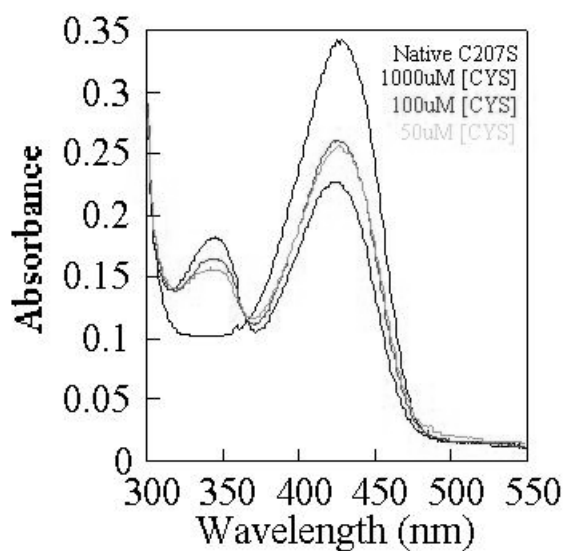
Figure 5:

Figure 5 – SDS PAGE illustrates induction. The molecular weight marker (1), the post-induced (2), and the pre-induced (3) samples allow the comparison and identification of the C207S protein (~46.8 KDa).

Figure 6:

Effect of L-cysteine on Absorption Spectrum of C207S

• Spectral changes with time upon addition of one equivalent L-cysteine to C207S.

• Spectral Changes upon addition of increasing concentrations of L-cysteine.

Figure 7:

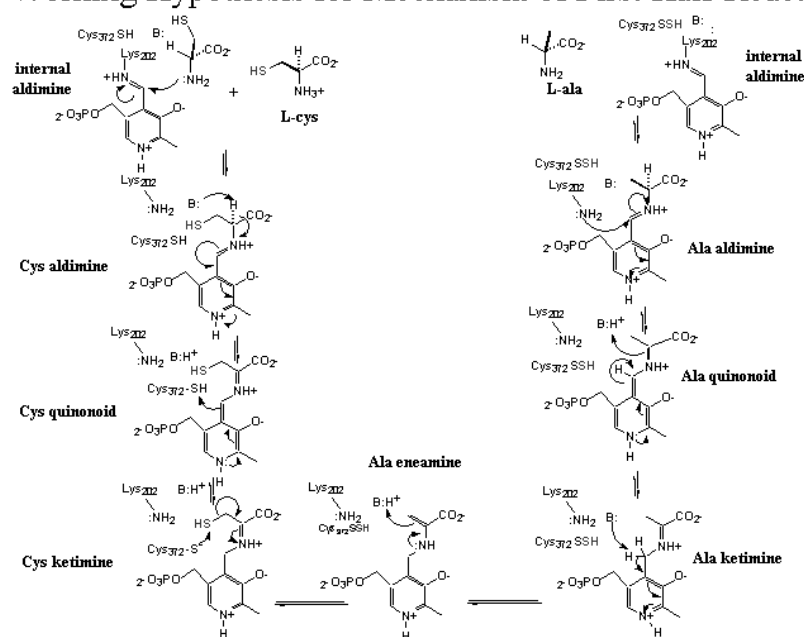Working Hypothesis for Mechanism of First Half-Reaction

Figure 8 – This figure illustrates the proposed chemical reaction of cysteine desulfurase activity thought to be similar, or identical, to the pathway of the NifS mechanism.

Optimized Registration for Computer Assisted Total Knee Arthroplasty

Jon-Michael Tucker, McNair Scholar, Pennsylvania State University

Faculty Research Adviser
Stephen Piazza, PhD, Assistant Professor
Departments of Kinesiology, Mechanical Engineering, Bioengineering,
Orthopedics and Rehabilitation
Pennsylvania State University

Introduction:

Total knee arthroplasty, (TKA), also known as total knee replacement, (TKR), is a surgical process that involves the removal of diseased bone and cartilage from the distal end of the femur and from the proximal end of the tibia. Once those cuts have been accomplished, a cement epoxy is then used to firmly secure the mechanical components that will operate as "new knees." Two to three hundred thousand United States residents receive knee replacements each year. The major reason for TKA is Osteoarthritis. Osteoarthritis is the stiffening of joints and bones due to a depletion of cartilage in those areas. Osteoarthritis can occur in the hands, feet, spine, knees, and hip. Severe Osteoarthritis in the knee joint involves the loss of all or almost all cartilage cushioning. Those cases can lead to bone rubbing on bone, which is extremely painful and causes rapid bone deterioration. The illness of Osteoarthritis can cause unhealthy changes in the body. People who suffer from knee joint pain are less prone to walk or exercise, which has an effect on their cardiovascular and respiratory systems. If persons experiencing pain do walk, their gait may alter as to make their walking more bearable.

This changes in their gait, which can have a crippling affect on foot angle, hip movement, and the lower back. These crippling affects can lead to obesity, alternate joint complications, and lower back pain.

The conventional procedure for total knee arthroplasty is by no means a simple or trivial process. This process is still strongly influenced by human error. There are many considerations that are necessary to perform a good TKA. Pre-operative planning is a helpful tool in TKA, but currently only has applications for choosing the proper size replacement components and choosing appropriate size jigs and guides for the patient. (Mantas & Bloebaum, 1995) Alignment issues of the femoral component are a large problem with conventional TKA. Knee alignment is based on three axes, the anteroposterior axis, the posterior axis, and the epicondylar axis. These axes represent the line or axis that the femur would rest upon if lying on a table, the line or axis that splits the femur in halves in the vertical direction, and the line or axis that goes through both epicondyles. During this procedure key anatomical landmarks for locating these

axes are removed. Then it is left to the surgeon's accuracy, pre-operative planning and judgments based on X-rays, and the precision of the surgeon's tools in order to insure proper alignment. While surgeons can be extremely precise during the operation, femoral misalignment of as little as $2°$-$3°$ or 2mm - 3mm in translation will cause multiple problems for the patient post operation. If the misalignment is too severe, a revision surgery will have to take place (Amira & Whiteside, 1992).

With the incorporation of computers in the operating room, more judgments can be made based on pure math. The perception and experience of the surgeon will still be helpful but not always necessary. By using a computer to assist in TKA, the patient, the doctor, and the hospital can all be assured that the accuracy of this surgery will be near perfect each-and-every-time.

Computer assisted surgery can only be achieved through registration. Registration is a procedure where a model such as a bone is located in 3-D space, then used as a reference guide to the x, y, and z directions for the computer. Even with computers in the operating room and proper registration achieved, the precision of the cutting devices and the machining of the knee components will still limit TKA results. However, the results will still be far better than human surgeons alone can achieve.

The purpose of this study is to provide a template for future revisions of TKA that will help produce a better and more accurate TKA result. By adding the assistance of computer software and robotic arms, not only will surgeons allow less room for error in the incisions and bone removal, but also in component design specifications.
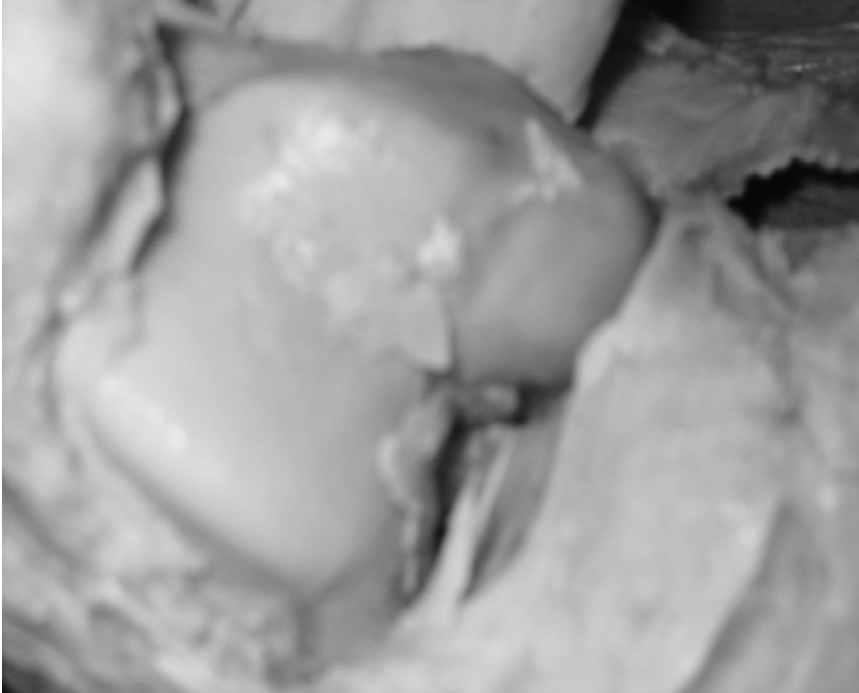

Figure 1: Frontal view of the distal end of the femur. The same view the surgeon views before bone removal takes place.

Because the actual knee components will fit on the bones more precisely, the patients will experience less pain and a greater range of motion in their knees following the operation, allowing them to enjoy a more full-filling life.

Literary Review:

Methods other than Optimized Registration for Computer Assisted TKA have been proposed and are being used. These other methods involve the use of fiduciaries and similar computer assisted set-ups. A fiduciary method involves a preliminary surgical procedure to place pins or markers on certain anatomical positions based on CT scans and/or X-ray images to help register, or locate, specific locations during the major surgery (Kienzle & Stulberg). While using the fiduciary method and computer technology, these markers/pins are located on the bones of the patient. Then, through registration and computer programming, a robot arm assists in the cutting and alignment that takes place during TKA. The fiduciary surgical procedure is not minimally invasive. It involves multiple operations on the patient while achieving only a single benefit, a new set of replacement knees (Abdel-Malek & McGowan, 1997).

Computer assisted integrated surgery became more prevalent in the world as the search for a more accurate convention for TKA continued. A program that could recognize a bone structure in virtual space and accurately register that same structure in real space was the next step in the design process. There are many algorithms used now to perform the described task, such as: segmentation, voxel, optimization, singular value decomposition, orthronormal and eigenvalue systems, unit and dual quaternion, (Eggert, 1997 and Maintz & Viergever).

The Iterative Closest Point, (ICP), proved to be the most efficient algorithm given the specific bone registration problem and operating room time constraints. The ICP algorithm is one of the most popular algorithms for image registration. Its popularity within the scientific world is accredited to its accuracy, robustness, and usability. Various papers by authors from all over the world describe and use the ICP algorithm for similar registration processes.

Besl and McKay provided the framework and the backbone for the ICP algorithm. They were the first to describe and use the ICP algorithm. The ICP algorithm that Besl and McKay produced is capable of handling different types of 3-D shapes. The algorithm uses points, lines, curves and triangles to match two sets of data to one another. The ICP algorithm always converges to the nearest local minimum. With the proper initial positioning executed, this algorithm can converge on the most complex of shapes.

The ICP algorithm registers a model set of data to an actual set of data by completing a number of very distinct steps. First the algorithm finds the closest set of data points, and then computes a registration built from quaternions. The quaternion matrix built for this registration is built from the centroids of each data set and the computation of a cross-covariance matrix. The next step involves applying the registration to the chosen data set. Lastly, the algorithm will stop transforming data

points through this repetition once a certain threshold error is met or the distance between original points and the new set of points is no longer decreasing. The threshold is a value chosen by the operator, which is built from the mean error of the original points in comparison to the new set of "closest points". (Besl & McKay, 1992).

There are variations of the ICP algorithm and certain methodologies that can be performed with the ICP algorithm, which raises ICP performance exponentially. While K-d tree and Elias methods are alternate forms of registration, they suggest ideas for applications that could work conjuctively with the ICP algorithm (Greenspan, Godin, & Talbot). By building triangular neighborhoods or special reference systems, the K-d tree and Elias methods provide a guide and a network throughout the registration process.

Research similar to, and in other cases surpassing, this experiment is already being done. Yet there is room for further validity in methodology and more consistency in results. In the Laboritoio di Biomeccanica, located in Bologna, Italy, researchers have used a registration procedure similar to the "Optimized Registration for Computer Assisted TKA", but they have gone further by performing this form of computer-assisted surgery on an actual cadaver (CAOS website). CT scans produced virtual models of the bones, the ICP algorithm registered the bones, and a robotic arm performed the proper cuts and incisions for the TKA.   The researchers have also performed an analogous experiment involving a unicompartmental arthroplasty. They used the ICP algorithm for registration and a similar set-up as previously described. The set-up differed only in the types of user interfaces and tools used (Marcacci & Tonet).

In Fluente and Glozman's articles, both propose and use methodologies that not only register a femur, but also provide comparative results on how a femoral registration should be achieved. These papers help to better define the discrepancy between a true fit versus an accurate registration. Meaning errors could be low, considering the threshold tolerances and initial orientation, while not necessarily attaining a near perfect registration. Through this experiment and its results, a better guide is provided that assures a near perfect femoral registration.


ICP Methods:

The ICP algorithm prescribed by Besl and McKay is a key tool in the analysis of a "set of best registration points."   This algorithm drives two sets of points closer to each other, minimizing the mean distance of all points in one set to the corresponding "closest point" in the other set. This algorithm was applied to match two identical surfaces: 1) being the whole object or surface model, 2) being part of that object surface or the digitized sample, and then place these two surfaces in 3-D space in their proper orientation with one another.

Because this proposed solution was designed for a specific problem for a specific application, the ICP algorithm was enhanced with certain features such as a "point to line" addition, a "point to plane" addition, a "triangle neighborhood" addition and lastly, a "Good Start Position" addition.

The point to line and the point to plane addition were both added to the algorithm to help its convergence on a global minimum. Because Computer Aided Design (CAD), surfaces are defined by many triangles, being able to define lines and planes in space to represent these triangles gave the algorithm a closer look at the actual surface being investigated, instead of just a cloud of points. Using parametric equations of lines and planes and certain geometrical relationships, this algorithm was enhanced to check each aspect of the surface, from points on a surface, to lines on a surface, to triangular planes on a surface in order to find the "closest point" or "best match."

The triangular neighborhood addition is based on vertex connectivity of triangles on the CAD, and its purpose was to make the algorithm run faster. Each point on the digitized sample has either a prescribed or a correlating "closest point" found on the CAD model. From that prescribed or correlating "closest point" on the CAD model, the vertexes of the immediate connecting triangles and those of the secondary level of connectivity were found and saved. This enabled the algorithm to check a lesser number of triangles, lines, and points for the location of the next "closest point" for every prescribed or correlating "closest point."
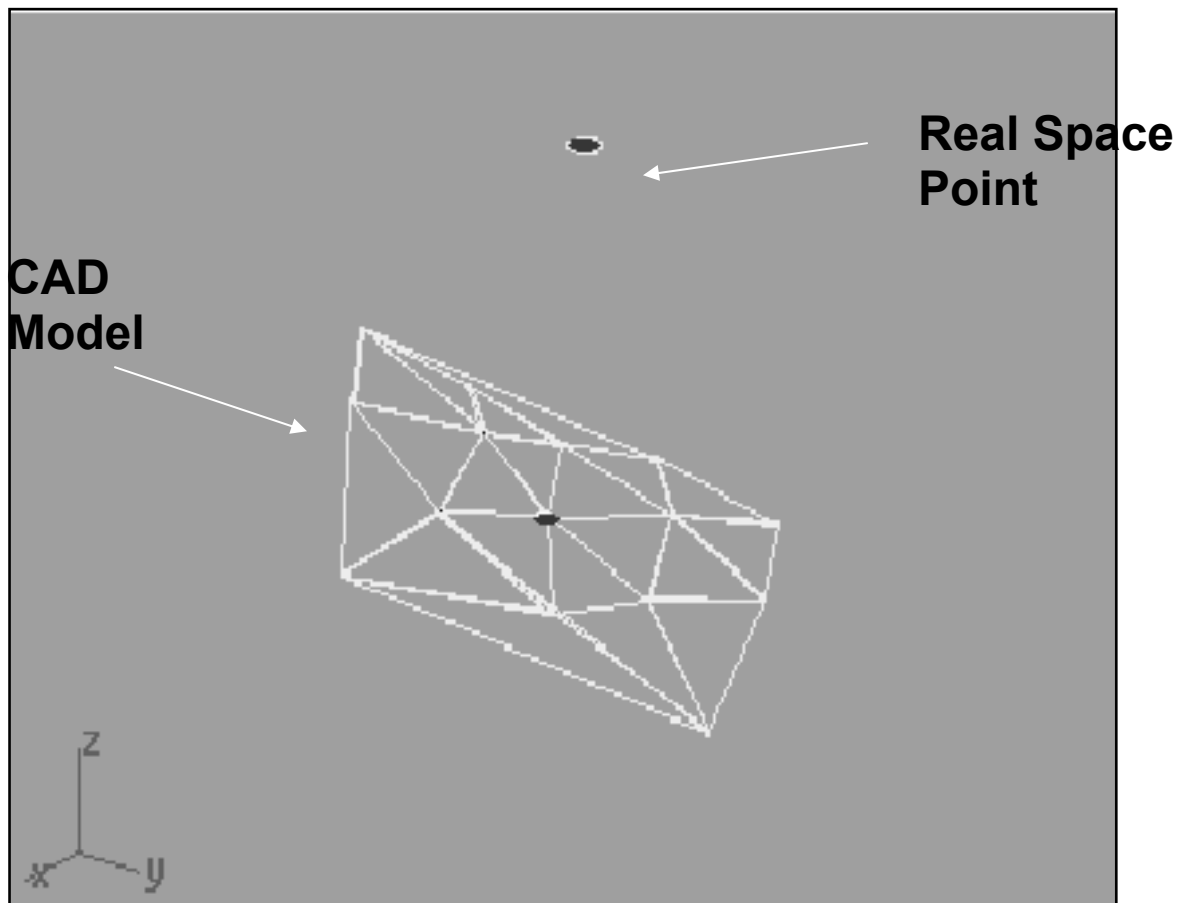
Routine developed by Dr. John Challis. This smaller algorithm was applied to give the ICP a proper place to begin since it is known that the absolute convergence of the ICP algorithm is based primarily on the starting positions of the two samples in question.

This Least Squares routine was applied because there were three points of reference, which were defined on the CAD model and again on the real-space model. Those three points were the greater trocanter, the lateral epicondyle, and the medial epicondyle.

Experimental Methods:

The tools used during this experiment consisted of a CAD model from http://www.cineca.it/hosted/LTM-IOR/back2net/ISB_mesh/mesh_list.html. Sawbones.com suggested this site, and is also the company from which the 3-D foam model of the left femur was acquired. A Microscribe digitizing arm from Immersion Corporation was used to accurately register points into software packages such as MatLab, (Mathworks, Inc.; Natick, MA), and Rhino 3-D modeling software (Rhinosceros; Seattle, WA). A Dell with Pentium 4 processor was used for these computer analyses.
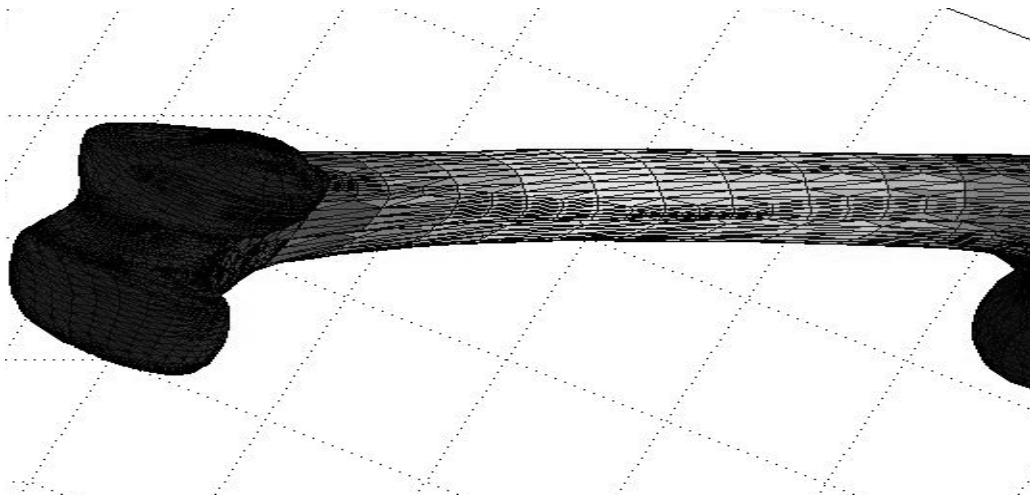

Figure 3: Computer Aided Design depiction of the left femur.

After previewing footage and diagrams of total knee arthroplasty, a defined foundational view was built in Rhino based on how much of the femur that is actually displayed during TKA, (Figures 1 & 3). The foundational view/ femoral display was divided into quadrants and an array was formed over the surface. These quadrants were built based on the prescribed view and the three axes used to align femoral components conventionally. The posterior condylar axis was used to bound the bottom of the box. The epicondylar axis was approximated to be the centerline of the box in the horizontal direction and the anterior-posterior axis was approximated as the centerline of the box in the vertical direction. The array of points consisted of 624 points on this surface.

Figure 4: Distal Femur surface chosen for analysis.

The optimization portion of this analysis consisted in taking the four quadrants and running 624 tests in order to find the "best registration points". In order to find the "best registration points" in each quadrant, three of the four quadrants were held constant as a base set of points. Then, with each test, one point from the quadrant being tested was added until all the points in that quadrant were individually tested using the set of base points for that quadrant. This allowed each point to have a weighted error factor. Each quadrant tested approximately 156 points and each quadrant took over 20 hours to complete.
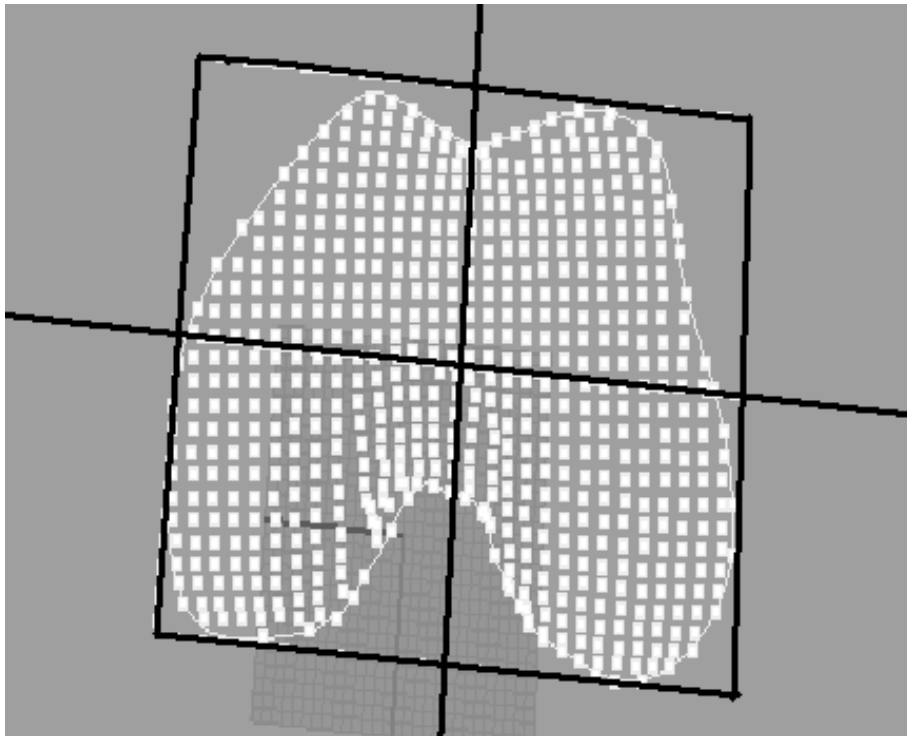

Figure 5: Top view of 3-D depiction of Four-quadrant set-up with bounding box.

The 12 points with the least weighted error were found for each femoral quadrant and saved. These 48 points were tested for accuracy and efficiency given the application. These tests involved moving the 48 points away from the CAD model and then allowing the ICP algorithm to place them back in their proper positions. From 4 different tests consisting of only translations, 7 outliers were identified and removed from the best registration point set, (Table 1 rows 1-4).

To test the "best registration points" against alternate registration points, 9 more tests were run. These tests were run with the Good Start algorithm added so starting positions became of no consequence. The first 3 tests were of the "best registration" points registered by hand using the Microscribe digitizing arm. The second 3 tests were again digitized by hand. Approximately 10 arbitrary registration points were discriminately chosen in areas other than those near the "best registration" points. The final three trials entailed approximately 10 registration points digitized by hand and representing "best registration" points as accurately as possible.

Results:

Table 1 - Data found before the implementation of the Good Start Position-

Trial	Number of Points	Point Type	Translation	Total Mean Error(mm)	Maximum Mean Error(mm)
1	624	Array Over Surface	[-.3 .2 .1]	0.2847	5.8226
2	48	Initial Best Regis	[-.3 .2 .1]	0.1922	0.8599
3	48	Initial Best Regis	[-3 2 1]	0.3836	1.3459
4	48	Initial Best Regis	[30 20 10]	0.6756	2.6006
5	41	Outlier Removal	[-.3 .2 .1]	0.1594	0.5455
6	41	Outlier Removal	[-3 2 1]	0.1921	0.9753
7	41	Outlier Removal	[30 20 10]	2.8055	7.3868
8	41	Outlier Removal	None	0.1303	0.4595

Table 1 - Continued

Trial	Number of Outliers	Time(sec)
1	51	559.69
2	7	75.69
3	9	153.17
4	8	261.69
5	9	101.39
6	5	152.21
7	3	208.16
8	8	60.168

Table 1 shows a range of tests. Trial 1 tested the 624-point array on the femoral surface. Trials 2 – 4 tested the initial set of defined "best registration" points and trials 5 – 7 tested the 41 "best registration" points. Based on the outliers found in trials 2 – 4, seven points were removed to make the 41 "best registration points". Trial 8 gives a base error and a base time that should be expected.
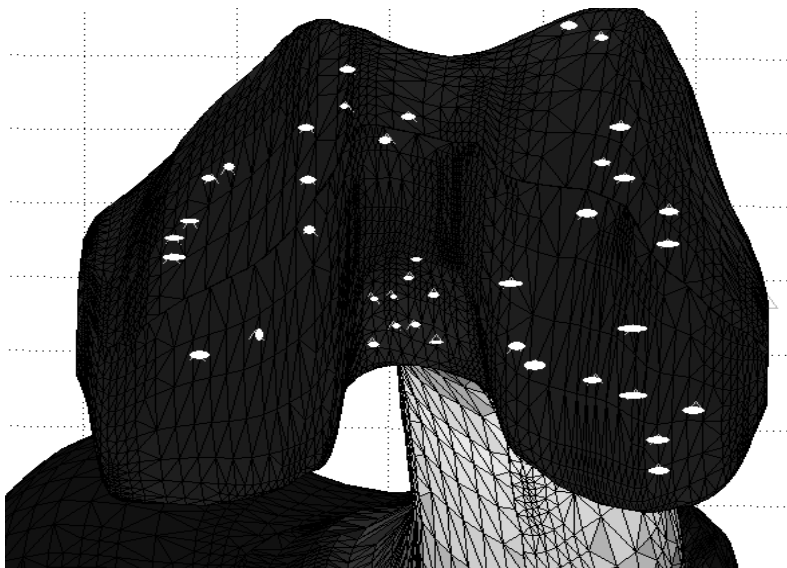

Figure 6: Showing a Matlab view of the 41 "Best registration points."

Table 2 -Data found after the implementation of Good Start Position-

Trial	Number of Points	Point Type	Total Mean Error(mm)	Maximum Mean Error(mm)	Number of Outliers	Time(sec)
1	44	Best Regis	0.245	1.3473	4	400.37
2	46	Best Regis	0.2055	1.0243	9	451
3	45	Best Regis	0.4868	2.7565	4	74.03
4	10	Arbitrary	0.1027	0.2187	3	116.89
5	10	Arbitrary	0.4562	2.0971	1	48.53
6	11	Arbitrary	0.1109	0.2547	3	210.78
7	10	Abbrev Best Regis	0.1876	0.3745	0	23.7
8	10	Abbrev Best Regis	0.0671	0.1828	2	24.02
9	11	Abbrev Best Regis	0.1143	0.2616	2	57.58

Table 3 – Averages of the Different types of trials

Trial	Number of Points	Point Type	Total Mean Error(mm)	Maximum Mean Error(mm)	Number of Outliers	Time(sec)
AVG(1-3)	45	Best Regis	0.3124	1.70930	5.667	308.47
AVG(4-6)	45.000	Arbitrary	0.22320	0.85680	2.333	125.40
AVG(7-9)	10.333	Abbrev Best Regis	0.1230	0.27290	1.333	35.10

Table 2 shows the results of tests of three different types of point sets from actual digitized samples. Trials 1 – 3 tested the accuracies and stabilities of the 41 "best registration points" with three reference points added. Because these points were digitized by hand, an extra point, or maybe a repeated point, was digitized. Trials 4 – 6 depict points designated as arbitrary because they were randomly chosen points on the femoral surface. Trials 7 – 9 tested an abbreviated set of "best registration" points. Table 3 is a color-coded chart of the corresponding averages in the previously discussed 9 trials in Table 2.

Discussion:

Some have suggested having a large array of points on a surface that, when applied to an algorithm, would better insure a proper registration. However, after examining the results of a direct comparison of the 624 points with the 41 "best registration points" and the "abbreviated best registration points", it was noted that this is not necessarily the case. The results show that there is a strong correlation between acquiring minimal errors and the choosing proper registration points. In addition to choosing proper registration points, convergence of the algorithm on a global minimum is completed more quickly than if one were to chose many registration points.

After examining the results for the "best registration" points, it appears the computer is attempting to draw the femoral view almost as an artist would make an outline sketch. This means that it looks as though the computer is trying to find "best registration" points that lay on the defining contours of the surface, in effect producing a line along the border where two surfaces meet. In addition the right condyle is relatively flat compared to the left condyle and no points were chosen in the lower right region. As in a drawing, a flat region can only be defined by shading and/or depth, but because the computer was not programmed to recognize depth nor shading, points were just not chosen in flat regions.

The intercondylar region of the femur is an extremely influential region for this type of analysis. From the 41 "best registration" points, approximately 25% of these points are in that region. When viewed by the human eye, the intercondylar region is the distinct region, which stands out from the distal portion of the femur. Indeed, without this distinction, the condyles would not be called condyles. This very distinction is what allows a proper fit or orientation to be achieved.

The acquisition of outliers is almost inevitable. Due to this phenomena, it is strongly recommended that 10 – 15 of the "best registration" points be chosen in addition to the three reference points for this application. So, if necessary, removing 2 – 3 outliers will not cause a flawed or improper registration. Outliers are a major concern in choosing a number of registration points, but with orientation also being a key component of this analysis, as more registration points are chosen, there is a greater probability that a proper orientation will be found within the specified tolerance.

Limitations:

Several limitations in this study should be noted. First, the variability of this study is low because testing was only done on one femur. Second, the registration errors could have been lower if more CAD models were available during the time of experiment. Research shows that a better registration can be found when meshes or polyhedrals of different qualities are tested in sequence from coarse to fine. The polyhedral used in this experiment differed from the foam 3-D model on a magnitude of 1.59mm to 0.79mm. This error is similar to actual procedures that take place in hospitals due to the accuracy of CT scanning, which is around 1mm at best. The final limitation of this study is that, while MatLab software carries out operations to 10 significant figures, the "best registration" points could only be approximated by the digitizing arm and operator.

Future:

This type of research will influence ACL and PCL repair procedures, and other joint replacement procedures such as hip joint replacements. Registration and computer assisted surgery has multiple applications for future procedures in spinal complications and brain surgeries. The concepts of perfect fit and improved range of motion based on proper registration are readily applicable for more efficient prosthesis.

This small study shows, that with enough research and the correct methodology, an improved TKA can be achieved. If computer assisted surgery for TKA was incorporated into the main stream of modern medicine, it would increase the success rate of TKA around the world. Not only would patients have more fulfilling lives due to their knees working as they did in the past, but research doctors, medical doctors, medical students, and others would have more time to find solutions for Osteoarthritis and other devastating illnesses that affect mankind.

References:

Glozman, D., Sholam, M., and Fischer, A. (2001) A surface matching technique for robot-assisted registration. *Computer Aided Surgery*, 6, 259 – 269.

Fluente, M., Lavallee, S., and Remi, J. (1999) Incorporating a statistically based shape model into a system for computer-assisted anterior cruciate ligament surgery. *Medical Image Analysis*, 3, 209 – 222.

Marcacci, M., Tonet, O., Megali, G., Dario, P., Carrozza, M., Nofrini, L., La Palombara, P. (2000) A navigation system for computer assisted unicompartmental arthroplasty. *Lecture Notes in Computer Science, S. L. Delp, A. M. Di Gioia and B. Jaramaz Eds., Springer*, vol. 1935, 1153-1157.

Glozman, D., Fischer, A., and Shoham, M. (1999) Registration of 3-D objects for Computer Intergrated Surgery. *M.Sc. Thesis, Mechanical Engineering, Technion*

Laboratorio di Biomeccanica – http://www.ior.it/biomec/caos.htm

Rusinkiewicz, S., and Levoy, M.(2001) Efficient variants of ICP algorithm. *Presented at the Third International Conference on 3D Digital Imaging and Modeling (3DIM).*

Besl PJ., and McKay, N. (1992) A method for registration of 3-D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2), 239 – 256.

Sunday, Dan. (2001) "About Planes and Distance of a Point to a Plane." http://www.geometry algorithms.com

Challis, J. (1995) A procedure for determining rigid body transformation parameters. *J. Biomechanics,* 28(6), 733 – 737.

Greenspan, M., Godin, G., and Talbot, J. (2000) Acceleration of binning nearest neighbor methods. *In Proceedings of Vision Interface 2000, Montreal, Canada*, 337-344.

Greenspan, M., Godin, G., and Talbot, J. (2001) A nearest neighbor method for efficient ICP. *Third International Conference on 3-D Digital Imaging and Modeling (3DIM '01) Quebec City, Canada*

Maintz, JB., and Viergever. (1998) An overview of medical image registration methods. *Image Sciences Institute, Utrecht University. Med cal Image Analysis* 2(1), 1 – 37.

Lavallee, S. (1996) Registration for computer-intergrated surgery: Methodology, state of the art. *Computer Intergrated Surgery Technology and Clinical Applications*. 77 – 97.

Eggert, DW., Lorusso, A., and Fisher, RB. (1997) Estimating 3-D rigid body transformations: a comparison of four major algorithms. *Machine Vision and Applications*, 9, 272 – 290.

Kienzle, T., Stulberg, SD., Peshkin, M., Quaid, A., Lea, J., Goswami, A., and Wu, C. A computer-assisted knee replacment surgical system using a calibrated robot.

Abdel-Malek, K., McGowan, D., Goel, V.K., Kowalski, D., and Smith, S. (1997), "Bone registration method for robot-assisted surgery: Pedicle screw insertion*. ImechE Journal of Engineering in Medicine,* 211(H), 221 – 233.

Mantas, J., Bloebaum, R., Skedros, J., and Hoffmann, A. (1992) Implications of reference axes used for rotational alignment of the femoral component in primary and revision knee arthroplasty. *The Journal of Arthroplasty*, 7(4), 531 – 535.

Arima, J., Whiteside, L., McCarthy, D., and White, S. (1995) Femoral rotational alignment, based on the anteroposterior axis, in total knee arthroplasty in a valgus knee. *The Journal of Bone and Joint Surgery*, 77(9), 1331 – 1334.

PENN STATE

Making Life Better™