FOREWORD

Welcome to Volume 15 of the Inquiry Undergraduate Research Journal. The Inquiry journal was developed by the Teaching Academy of the University of Arkansas and is supported financially and conceptually by the offices of the Provost and the Vice Provost for Research and Economic Development. Inquiry provides a forum for sharing the research and creative endeavors of undergraduate students and their faculty mentors at the University of Arkansas. While the journal has traditionally been published once per year, given the increase in the number of outstanding submissions to the journal for possible publication, we have decided to publish two online volumes each year.

Volume 15 of the Inquiry Undergraduate Research Journal features the unique contributions of seven undergraduate student authors and their faculty mentors. Their research and creative endeavors span diverse fields at the University of Arkansas, including Architecture, Marketing, Agricultural Economics and Agribusiness, Mechanical Engineering and Chemistry/Biochemistry. For example, using the theories of Kevin Lynch and Jane Jacobs, Hannah Breshears measures, compares, and evaluates spatial performance of formal and informal neighborhood spaces in Istanbul. Turning to Marketing, Samantha Feng compares and contrasts consumer attitudes, purchase intentions, and product choices as they relate to a standardized sustainability index. Mike Norton, Agricultural Economics and Agribusiness, used 2010-2011 production data from Ghana to estimate the change in net present value associated with Cocoa Livelihoods Program training. Luke Osborn, Mechanical Engineering, examined the effects of targeted microstructures on the tribological performance of a surface in an effort to examine their ability to withstand considerable wear. In his study, Keaton Piper developed and examined whether low-cost titanate nanofiber-bioscaffolds compensate for the longstanding problem of implantable biomaterial’s poor stability and versatility.

I would like to extend a special thank you to the many faculty members who volunteered their time and expertise to provide comprehensive reviews of student manuscripts. While we are unable to publish all of the submitted manuscripts, we want to thank the students and faculty mentors for their diligent efforts. Please join me in congratulating our authors; I hope that you enjoy this edition of the journal as much as we have.

We plan to publish Volume 16 of the Inquiry journal in March 2014. I encourage undergraduate students and faculty mentors to consider the Inquiry Undergraduate Research Journal for future publication.

Marcia A. Shobe, Editor
TABLE OF CONTENTS

Inquiry Undergraduate Research Journal of the University of Arkansas

Volume 15          Fall 2013

1  Editor’s Foreword

2  Table of Contents

3  ARCHITECTURE: Breshears, H. (In)formal Distinction in Urban Istanbul: Evaluating Spatial Performance

24  MARKETING. Feng, S. Communicating Product Sustainability: Consumer Responses to Sustainability Labeling in a Retail Laboratory Environment

45  AGRICULTURAL ECONOMICS & AGRIBUSINESS: Norton, M. Cost-Benefit Analysis of Farmer Training in Ghanaian Cocoa Farming

69  MECHANICAL ENGINEERING: Osborn, L. Enhanced Tribological Properties of Surfaces Patterned with SU8/DLC Microstructures

87  CHEMISTRY & BIOCHEMISTRY: Piper, K. Preparation of New Versatile and Implantable Titanate Nanofiber-Bioscaffolds via Efficient Cation Exchanges

Cover image by Rob Byrd Designs
(IN)FORMAL DISTINCTION IN URBAN ISTANBUL:
EVALUATING SPATIAL PERFORMANCE

By Hannah A. Breshears
Department of Architecture

Faculty Mentor: Dr. Noah Billig
Department of Landscape Architecture

ABSTRACT

As globalization continues to draw the world into closer economic and intellectual dependence, massive tracts of informally designed communities in Istanbul are being cleared to accommodate the growing infrastructure of the modern, tourism-driven city. This attempt to purge the city of its ‘squatter’ heritage is startling and raises questions of cultural integrity in urban development. Istanbul’s desire for expanded global investment is particularly apparent in the object of this study, the blended district of Kartal. This study measures, compares, and evaluates spatial performance of formal and informal neighborhood spaces, but makes no formal attempt to draw normative prescriptive conclusions. The theories of Kevin Lynch and Jane Jacobs are synthesized in order to analyze three constructs of spatial performance: density, grain and access. As such, this study has not only produced a more rigorous tool for remote analysis, but one that can be applied to other urban settlements in the future.
1. Understanding Istanbul & Urban Performance

As globalization continues to draw the cities of the world into closer economic and intellectual dependence, Istanbul stands as a bridge between two continents and a city on the brink of a total urban transformation. International planning theory has become an increasingly influential part of Istanbul’s governmental agenda since the fall of the Ottoman Empire; as a result the city has renewed its drive for large-scale redevelopment of its districts (Wende 27). Massive tracts of informally designed communities are being cleared to accommodate the structure of the modern, tourism-driven city. The attempt to purge some areas of Istanbul of its ‘squatter’ heritage raises questions of cultural and architectural integrity in urban development. For example, what benefits do user-generated design and construction bring to Istanbul, if any, and how do the spaces produced by informal, formal, and blended development compare?

Turkish gecekondular, the informally designed squatter settlements prevalent in Istanbul since the 1950s, provide an environment for study unfettered by traditional design standards and regulations, and as such garner further scholarly inquiry. This study focuses on the public spaces of the blended district of Kartal (kahr-tahl), where formally and informally designed areas reside and merge, and attempts to determine the performance or quality of the urban fabric in each type of settlement, as well as the district as a whole (Figure 1). Well documented in urban planning literature, Kartal is a coastal settlement located on the Asian side of Istanbul, just below the former administrative boundary of the city (Urban Age 26).

Fig.1: Administrative boundaries of Kartal. Source: Author.

The study area is limited to this district, where the physical configurations of public space in informal neighborhoods are extracted from available aerial imagery and compared to the analogous configurations of formally designed public space within the district. The data under study have been limited to imagery from the last 50 years and representative sites that demonstrate potentially fruitful integrations of local and global norms for study have been selected. Special attention has been paid to the formal planning schemes made and executed for
this area of the city within the allocated time period, with the intent to determine how the developmental patterns identified in informal, formal, and blended areas compare.

According to urban theorist Kevin Lynch, public spaces are “all those regions in the environment which are open to the freely chosen and spontaneous activities of people”; this definition encompasses not only the designated fields and parks traditionally identified as “green space” in a city, but also unfenced vacant lots, streets, alleys, and abandoned waterfronts (Lynch, Banerjee, and Southworth 396). Public space is designed to satisfy the whims of the user, to extend his or her knowledge of self and the environment, and to provide space for growth and change, not merely to exhibit aesthetic charms. Lynch even champions the role of “derelict and waste lands” in the public space system for their ability to provide satisfactions distinct from those afforded by “ordered” space (Lynch, Banerjee, and Southworth 400). He argues that “where open space is not highly manicured, and the social investment is low, the individual has a chance to demonstrate mastery, to meet challenges, and participate actively in a way usually denied him in the protected and expansive city environment” (Lynch, Banerjee, and Southworth 397).

Lynch’s position does not ‘pigeonhole’ informal space and its use as a detriment to a city, but acknowledges the potential benefits such development could bring to urban design and urban experiences. Hulya Ertas similarly advocates for the use of “unprogrammed space” in the city as a purveyor of urban success or failure (57). In her research, Ertas identifies three characteristics deemed most responsible for the success of informal, unregulated space relative to other formal space: openness to transformation and change, encouragement for heterogeneous use, and propensity for privatization (53).

Gecekondu citizens attempt to beautify the streets and gardens close to their homes by adopting public spaces as their own. According to Ertas, this lends a life and vibrancy to the public spaces of informal settlements that are not evident in mass housing projects. The highly mobile vendors of Turkey’s informal markets similarly transform space by claiming their place in open forums and along streets, and drawing large crowds of people to unregulated areas. Though both Madanipour and Ertas recognize the influence that urban planners and managers have on the collective domain of the city, they also understand the role of citizen-architects in spatial design and use. While a number of methodological precedents can be identified for urban analysis, there are few theories that provide usable criteria for determining spatial quality or value. Even fewer systems have been widely recognized and applied, though the theories of Jane Jacobs and Kevin Lynch seem to have had the most influence in the last 50 years. While a handful of data-driven systems have emerged post-Lynch, his methods continue to have considerable authority in the fields of urban design and theory and thus form the basis of my own methodology.

According to Jane Jacobs, intense zoning laws and use-legislation that strictly limit the style and pattern of development in urban areas have an enormous impact on contemporary design and renewal schemes. Jacobs heavily criticizes such policies for creating unnatural, and as such, unused urban space (J. Jacobs 151). She outlines four conditions necessary to create “healthy” spaces and promotes the chaos and innovation demonstrated in unregulated communities over the order and efficiency upheld by modern planners. Jacobs calls for mixed-use neighborhoods, short blocks to allow high pedestrian permeability, population and structural density, and provisions for buildings of various levels of age and repair (J. Jacobs 151). These characteristics, deemed “generators of diversity,” are inherent to gecekondu design, though the success of such neighborhoods has not been adequately assessed. This study proposes an analysis
of the informally designed public space of gecekondular that will not only contribute to our knowledge of user-generated spaces, but will form a basis for comparison with formal space that can begin to evaluate a number of Jacobs’ criteria as measures of good urban space.

Though Lynch disagrees with Jacobs’ understanding of diversity, convinced that it could not be measured “until one knows how people perceive difference, and in which features variety is important,” his spatial dimensions align with Jacobs’ own theories (A Theory 192). Together, the theories of Jacobs and Lynch form an analytical base by which quality and performance can be studied, if not measured. Kevin Lynch identifies the notions of perceived vitality, spatial sense, environmental fit, access, and control as “dimensions of performance” that create good urban form (193). In Lynch’s system, vitality refers to the capacity of a place to support biological functions—to provide food, water, energy, air, and waste disposal for human inhabitants (A Theory 122). A vital city not only shelters from hazard and disease, but also regulates external stimuli to the satisfaction of its citizens.

Spatial sense and access refer to the resident’s perception of urban form and its ability to accommodate desired action. The access dimension is measured in a more quantitative way, for it depends heavily on the rate of occupation and the use of city streets, sidewalks, and alleys. It is important to note, however, that social barriers formed by gentrified neighborhoods and government sanctioned commerce as opposed to the informal shops and markets of the gecekondular can limit access as effectively as physical barriers (Lynch, Banerjee, and Southworth 401). Spatial sense is a measure of the city’s formal structure and the perceived coherence of its linkages. Lynch’s system lends itself to quantifiable measurement, even remote analysis in the case of vitality, access, and possibly spatial sense, though the nature of fit and control in the city requires close interaction with its structures and residents.

These measures of urban form were at least partially derived from Lynch’s graphic method. Lynch devised and defined five elements with which he was able to distill the key features of virtually any urban area into a legible diagram. A series of paths, edges, districts, nodes, and landmarks were identified for each city case study and used to compare and contrast form as a measurable entity, without auxiliary influences. Though Lynch recognized the city as a “multi purpose, shifting organization…raised by many hands,” his method is limited to the physical appearance and structure of a given area and can only measure social or cultural meaning abstractly (K. Lynch 91).

Lynch’s five elements of path, edge, district, node, and landmark, though useful in their own right, are incorporated into the notion of grain. Grain seems to function as a more comprehensive analytical device that acknowledges Jane Jacobs’ cry for diversity in urban space and is deemed, “critical to the goodness of a place” (A Theory 266). In these later writings, Lynch determined that density, grain, and access form the internal texture of a city, and that by measuring these characteristics, its performance might be judged. Density and access have been addressed in various forms already, but grain requires further elaboration. For Lynch, grain refers to the “way in which the various elements of a settlement are mixed together in space;” grain also has multiple strengths or ratings (A Theory 265).

Despite the evolution of technologically advanced systems, such as the Global Positioning System (GPS), Google Earth, and smartphones, novice and professional researchers continue to study and employ Lynch’s method of imaging. The imaging technique remains relevant in modern urban analysis for its ability to create intuitive maps that capture the way cities “feel” according to land use blogger Roger Valdez (Valdez 1). Lynch’s method is widely accepted as one that can be applied at the micro- and macro-levels of urban analysis; it is also
useful for this study in the sense that it can be easily produced for the analysis of multiple sampled areas, and it allows for extremely complex spatial data to be distilled into concise graphic representation.

2. Measuring Urban Performance

Jane Jacobs and Kevin Lynch, both well-respected advocates of micro-level urbanism, champion form analysis as a viable method of study. Jacobs dismisses the solitary application of use-by-use analysis of cities as studies that yield overall pictures “about as useful as the picture assembled by the blind men who felt the elephant and pooled their findings” (143). For Jacobs, spatial and architectural diversity are the keys to “good” urbanism. Jacobs suggests that to understand the life-giving complexity of a thriving city, “we have to deal outright with combinations as essential phenomena” (J. Jacobs 144). Though micro-level studies are important to Jacobs’ work, she advocates the form analysis method as a means of comprehending and creating “good” urbanism only when used in conjunction with macro-level analysis.

Kevin Lynch makes a case for form analysis by moving away from the basic units of city form described in *Image of the City* and postulating about the method itself. According to Lynch, good urbanism cannot be achieved by a collection of small, but “handsomely designed” areas, for “every physical whole is affected not only by the quality of its parts, but also by their total organization and arrangement” (Lynch, Banerjee, and Southworth 358).

For the purposes of this study, we combine the two dominant forms of investigation, micro-level and macro-level, to provide a comprehensive analysis of formal and informal design performance in Kartal. It should be noted that the following methodology is based on the work of predominantly western theories of analysis and that certain biases in that respect can be assumed. However, as Anne Vernez Moudon notes in her catalogue of design theory, these methods have been applied in research and planning endeavors across the globe and are readily accepted in the diverse disciplines of planning, architecture, and geography as a viable means of analysis (363).

This study is an explanatory single-case study with three embedded units. Singleton, Straits, and Straits outline three major categories of study that are relevant to the design of this project: exploratory, descriptive, and explanatory research. While exploratory studies simply investigate a phenomenon about which little is known, descriptive and explanatory research are considerably more structured. According to Singleton, descriptive research is normally a “fact-finding enterprise… focused on relatively few dimensions of a well-defined entity” (68). This study moves beyond the singular phenomenon of informal development to analyze the relationships *density*, *grain*, and *access* form in the urban context so that the relative performance of formal, informal, and blended spaces might be compared and judged.

Research of this type requires rigorous analysis of each variable and an understanding of their impact on the formation of good urban space that cannot be dismissed as mere description. Data are analyzed within each embedded unit and between them in order to draw conclusions about the main unit (Yin 86). The purpose of this study is to answer the following research questions, understood as stages of analysis conducted chronologically:

1) Can informally and formally designed regions of Kartal be identified and mapped?
2) What spatial typologies can be identified in the fabric of each?
3) What structural and spatial densities can be observed in the formal and informal regions?
4) How do the spatial and architectural grains of these regions compare?
5) What degrees of pedestrian, vehicular, and transit-oriented access can be observed in the informal and formal regions of Kartal?

The objects of study are analyzed through the operationalization of three constructs: 1) density, 2) grain, and 3) access. These constructs are combined through data and methods triangulation to assess the larger construct of performance in selectively sampled regions of Kartal.

There is one main unit of analysis defined for this study, with three embedded units. The singular district of Kartal in Istanbul, Turkey, has been selected as the main unit of analysis and is bound geographically by the Omerli Reserve to the north and the Sea of Marmara to the south. Kartal is divided into four quadrants by the D-100 and Samandıra Kartal Bağlantısı highways and has been targeted by the Istanbul Municipality Planning Commission as the site of an urban revitalization scheme submitted by architect Zaha Hadid (Ayatac 8). Kartal’s history as an informally designed district has been amended by formal government intervention. The region can now be considered a blended district where formally and informally designed areas co-exist. This blended status provides the conditions necessary for a single researcher to analyze each type of development within a short period of time; it was thus chosen as the basis for this study.

The three embedded units of analysis identified for this study include the formal, informal, and blended regions of Kartal; the regions can be understood as distinct units within the greater district, but also as characteristics representative of the district as a whole. As such, the embedded units have an interdependent relationship that contributes to the spatial and architectural performance of the main unit, the blended district of Kartal. Developmental divisions within the main unit of analysis were inferred and mapped through a critical reading of planning and housing legislation effective in the area since 1950, as well as the analytical reports of environmental studies made in the area since that time (Unsal). Formal and informal regions of the same district were selected to ensure that the legal and cultural factors involved in each area’s development were analogous (Figure 2).
The overarching construct measured by this study is that of *performance*, or *quality* as it is called by Jane Jacobs, of formal and informal fabric in the urban environment of Kartal. Both Jacobs and Lynch recognize that the quality of a place is not simply determined by its physical form, but by the behavior and intent of the society that occupies it. Lynch writes quite extensively about this interdependence, though he contends that a certain degree of goodness can be measured solely by reference to spatial form through the application of *performance dimensions*.

According to Lynch, a city with satisfactory performance is one that is “vital, sensible, well fitted, accessible, and well controlled,” such that all characteristics are “achieved with just and internal efficiency” (*A Theory* 235). As a number of these dimensions have been discussed in detail already, in more general terms, good urbanism is found in places that are “continuous, well-connected, open, and conducive to development” (*A Theory* 235). Jacobs agrees that there is no single element or “kingpin” in a city that serves to clarify all, but that “the mixture [of elements] itself is the kingpin, and its mutual support is order” (J. Jacobs 376). She writes that “only intricacy and vitality of use give, to the parts of the city, appropriate structure and shape,” and has defined measurable characteristics of diversity to that end in a way much similar to Lynch (377). Jacobs calls for mixed-use districts, pedestrian access, varied structural age, and dense concentrations of streets and intersections to ensure diversity.

The *performance* element is broken into three sub-con structs by Kevin Lynch, who declares that “density, grain, and the access system—the internal texture of a city—are the principal features by which we may judge its performance” (*A Theory* 274). As mixed-use and structural age fall easily under the construct of grain, and pedestrian access and density are even more obviously accounted for, Jacobs’ criteria can be analytically subsumed by Lynch’s constructs. Thus, the sub-constructs of *density*, *grain*, and *access*, as defined below, form the analytical base of this study by which the main construct of *performance* is judged.

**Density**

*Density*, called “concentration” by Jane Jacobs, refers to the quantity of something per unit of measure, particularly a unit of land. As such, several types of density can be observed in a given area, reflective only of the characteristic being measured. Jane Jacobs defined three of these types in her description of urban diversity, each of which play a role in the “goodness” of an urban area and can be quantified for study: structural or built density, dwelling density, and use density, of which use refers to the type and number of amenities available in a given area (200).

**Grain**

The construct of *grain* refers to the “typical local interrelations between similar or dissimilar elements, without reference to total pattern” described by Kevin Lynch (*Sense* 362). A more accessible definition reveals that grain is “the way in which the various elements of a settlement are mixed together in space…be they activities, building types, persons, or other features” (*A Theory* 265).

**Access**

The construct of *access* is described as the ability of a resident or visitor to move freely toward desired people, goods, or settings within an urban area, and speaks particularly to the number and quality of paths, edges, and intersections in a given place. According to Allan Jacobs, *access* not only privileges open-space design over structural form, but also begins to measure the comfort of open-space in dimension and movement, elements necessary to the design of “good” urban fabric (A. Jacobs 302).
Table 1 illustrates the study’s constructs, with definitions and means of operationalization for each. The construct of *density* is operationalized through the construction of figure-ground diagrams and square-mile studies described (A. Jacobs 302). Figure-ground diagrams are used to create a visual register of occupancy that can be qualitatively described or quantified as a percentage for comparison. The square-mile studies advocated by Allan Jacobs are the limiting framework for these diagrams, designed to showcase street and block patterns that “permit some dimensioning and measuring of differences and similarities and of how those seem to change over time and distance” (268). This type of measurement is applied to the embedded units of analysis (i.e., formal, informal, and blended regions) in order to draw conclusions about the main unit of analysis (i.e., the district as a whole).

<table>
<thead>
<tr>
<th>Element</th>
<th>Definition</th>
<th>Operationalization</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Density</strong>&lt;br&gt;-built&lt;br&gt;-dwelling&lt;br&gt;-use</td>
<td>Quantity of something per unit of measure, typically per unit of land</td>
<td>Analyzed through figure ground diagrams and square mile maps from <em>Great Streets</em>, grid units coded by development type and concentration, summarized as a percentage for comparison</td>
</tr>
<tr>
<td><strong>Grain</strong></td>
<td>Way in which the various elements of a settlement are mixed together in space</td>
<td>Evaluated using Lynch’s diagramming and photo-grid techniques, identified using Google Earth photos, 2- and 3-D Geodata</td>
</tr>
<tr>
<td><strong>Access</strong>&lt;br&gt;-transit&lt;br&gt;-vehicular&lt;br&gt;-pedestrian</td>
<td>Ability of a resident or visitor to move freely toward desired people, goods, or settings within an urban area, particularly the number or quality of paths, edges, or intersections</td>
<td>Analyzed using characteristics of diversity from Jacobs, documented and compared using path network diagrams similar to those found in <em>Urban Age</em></td>
</tr>
</tbody>
</table>

The *grain* construct is operationalized using Lynch’s diagramming and photo-grid techniques, created from aerial- and street-level imagery (Lynch, Banerjee, and Southworth 266). Once a visual database is established, iconographic diagrams are drawn utilizing the five elements of urban fabric (paths, edges, districts, nodes, and landmarks) designed by Lynch to create an “image” of the region for analysis and comparison (*A Theory* 47). This type of measurement is applied at a small scale to the embedded units of analysis (i.e., sampled formal and informal regions) in order to understand the main unit of analysis, the entire district of Kartal.

*Access* is operationalized using Jane Jacobs’ four *characteristics of diversity* as a guide for quantitative measurements of the pedestrian and vehicular path networks observed in each sampled area (150). Once calculated, these measurements are reproduced as a series of path diagrams that can be used to interpret the strengths and weaknesses of the *access* system in each sampled region. This type of measurement is applied to the embedded units of analysis (i.e., formal, informal, and blended regions). All analyses are combined in a cross-case synthesis at the main unit of analysis, Kartal. This is achieved by creating a photo-grid of the entire area.
from which observations informed by Kevin Lynch’s five elements are made and “imaged” through a series of iconographic diagrams. The process of creating each of these tools is described in the next section.

**Photo-grid**

To create an urban photo-grid in the style of Lynch, a grid is laid over the base map of the study area, at a scale analogous to the fabric of the region (Figure 3). Once this grid has been established, the nearest accessible point to each grid intersection is found and documented to produce a “complete sampling of the visual character of the area” that can be used in further analysis (Lynch, Banerjee, and Southworth 266).

![Fig.3: Example photo grid, Gumuspinar. Source: Author.](image)

**Iconographic diagrams**

Using the spatial characteristics observed in the imagery of the photo-grid, the five elements of urban fabric conceived by Kevin Lynch are catalogued as observed in the sampled area. Once the elements are coded for diagrammatic reproduction using Lynch’s iconographic key, they are re-applied to the map of the area of study to create a symbolic “image” of the urban form (Figure 4). This method of analysis is used to distill complex spatial data into a diagram that can be used to identify and compare form qualities within the city or district fabric and to illuminate the role of physical form in spatial *performance*.
The embedded units of analysis are evaluated for the constructs of density, grain, and access at the scale of the sampled region. This is achieved using the previously described iconographic diagrams of Lynch, as well as the figure ground method of analysis proposed by Allan Jacobs and measurements guided by the diversity criteria proposed by Jane Jacobs. The sampled areas addressed in this study have been limited to one square mile in accordance with Allan Jacobs’ method. This limitation was adopted to ensure that multiple units could be documented and analyzed under the time constraints of the study using the two techniques described in the next sections.

**Figure Ground**

Figure ground studies are simple diagrams that give order to the complicated relationship of structure and space in an urban area (Figure 5). Figure ground images “help visualize the microstructure of urban neighborhoods, how buildings (in black) and open spaces (in white) come together to create an integrated urban whole” (*Urban Age* 36).
Path Network Diagrams

To create a path network diagram, lines of varying thickness and hue are laid over a map of the sampled area to represent the size, direction, and density of accessible paths (Figure 6). This method is used to distill the complex network of transportation and mobility into a simple graphic representation that can be compared across and between sampled areas (**Urban Age** 30).
The sampling frame is a map of the district of Kartal, Istanbul, as defined by the administrative boundaries created in 2004. The sampling design uses stratified random sampling to select sample sites for the embedded units of analysis, as well as the main unit, Kartal, for the singular construct of grain (Singleton, Straits, and Straits 124). The administrative boundaries of the district’s individual neighborhoods, called mahalleler, divide the area of study into 21 representative units. These administrative divisions are a more efficient means of understanding the district than conventional sampling techniques and have been used to select representative samples of the area in the place of a traditional polygonal grid.

For the embedded units of analysis, the sampling size is limited to a single neighborhood, or mahalle. Three spatial strata were identified during preliminary analyses of the district: formal, informal, and blended space. A single mahalle was selected for each stratum, of which 75% of the land area within the mahalle conforms to the stratum identity. Purposive sampling strategies were employed to ensure analogous conditions were present in each of the sample units, particularly in the study of housing types/densities and access to major highways (Singleton, Straits, and Straits 133).

Each sampled unit was restricted to a square mile of land in accordance with the analytical technique defined by Allen Jacobs (260). Observation sites were again selected using stratified random sampling of selected grid units from within the square mile sample in order to measure the pedestrian strata of analysis for the construct of access (Figure 7). Purposive techniques, as defined by Dixon, were used to ensure that grid units representative of the urban fabric were chosen for analysis (33). The stratified method of analysis utilized in this study is summarized graphically in Figure 8 to clarify the diagramming techniques employed throughout the study. Though each mahalle was subjected to the full range of analysis and documentation described previously in this chapter, only the blended unit of Cevizli is represented in Stages 3 through 5 of the figure.

![Fig. 7: Pedestrian units, Gumuspinar. Source: Author.](image-url)
Once the mahalleler of Esentepe, Gumuspinar, and Cevizli were identified for sampling, a preliminary analysis of each mahalle was conducted to identify the form typologies described by Lynch. Direct observations from satellite imagery were recorded using hand-sketching techniques. No more than four form typologies were identified in any one mahalle, and only the characteristic of dominance was present in each. In the formally designed mahalle of Esentepe, the dominance and singularity types were observed. The informal mahalle of Gumuspinar was clearly dominated by small-scale residential fabric, though a large area of open farmland in the northeast corner of the sample area formed a singularity clearly far removed from the dense housing blocks. The blended mahalle of Cevizli exhibited only the characteristics of dominance and continuity.

INQUIRY: Volume 15
3.1 Density Analysis

Square mile maps and figure ground diagrams were used in the first stage of analysis to distill complex housing and land use data. These diagrams were constructed for each mahalle using area maps and building information supplied by the Istanbul Buyuksehir Belediyesi and the Istanbul Electricity, Tramway, and Tunnel General Management Department. This data was cross-referenced with satellite imagery from Google Earth taken as late as December 2012 to produce square mile diagrams of each sample area using AutoCad 2012 software. The gridded system developed for photo documentation was employed in each analysis to divide the sample area into describable units. As such, specific units have been referenced by their grid designation in the remainder of this chapter (A1, B2, etc.).

The built density, dwelling density, and use density, as described by Jane Jacobs, were evaluated in each sample and are included in Table 2 as percentages for easy comparison between cases (200). For the purposes of this analysis, built density refers to the percentage of land covered by buildings, regardless of use. Dwelling density is more selective, referring specifically to the percentage of land covered by residential buildings, regardless of scale, though an attempt to differentiate between mass housing and low-density housing was made. Use density refers to the percentage of land employed as industrial, residential, and blended space, each represented as a percentage of the total land area. Units with no structural presence were labeled open space, and thus included as a fourth category of use. All density types were evaluated on a unit-by-unit basis and calculated as a percentage of the total number of units observed in the sample grid. Only the final percentages for each density type are included in Table 2.

Table 2: Density Measurements

<table>
<thead>
<tr>
<th>Mahalle</th>
<th>Built density</th>
<th>Dwelling Density</th>
<th>Use Density</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low: 34.7%</td>
<td>Mass housing: 4.1%</td>
<td>Industrial: 38.8%</td>
</tr>
<tr>
<td></td>
<td>High: 63.3%</td>
<td>Low-density: 38.8%</td>
<td>Residential: 30.6%</td>
</tr>
<tr>
<td></td>
<td>None: 2.0%</td>
<td>Combined: 16.3%</td>
<td>Blended: 28.6%</td>
</tr>
<tr>
<td>Esentepe</td>
<td></td>
<td>Total housing units: 59.2%</td>
<td>Open: 2.0%</td>
</tr>
<tr>
<td></td>
<td>Low: 30.6%</td>
<td>Mass housing: 6.1%</td>
<td>Industrial: 8.2%</td>
</tr>
<tr>
<td></td>
<td>High: 57.1%</td>
<td>Low-density: 38.8%</td>
<td>Residential: 83.7%</td>
</tr>
<tr>
<td></td>
<td>None: 8.2%</td>
<td>Combined: 36.7%</td>
<td>Blended: 0.0%</td>
</tr>
<tr>
<td>Gumuspinar</td>
<td></td>
<td>Total housing units: 81.6%</td>
<td>Open: 8.2%</td>
</tr>
<tr>
<td></td>
<td>Low: 38.8%</td>
<td>Mass housing: 8.2%</td>
<td>Industrial: 20.4%</td>
</tr>
<tr>
<td></td>
<td>High: 59.2%</td>
<td>Low-density: 51.0%</td>
<td>Residential: 44.9%</td>
</tr>
<tr>
<td></td>
<td>None: 2.0%</td>
<td>Combined: 12.2%</td>
<td>Blended: 32.7%</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Total housing units: 71.4%</td>
<td>Open: 2.0%</td>
</tr>
</tbody>
</table>

The formally designed mahalle of Esentepe exhibited a high built density overall, with few undeveloped units in the sample area. Low development areas occur along the edges of the industrial districts and near a pond observed in the sample area, creating wide gaps between residential and industrial fabric. Units with purely industrial fabric are most prevalent in the formally designed mahalle, forming nearly 40 percent of the total land area. Residential and blended units, those exhibiting both residential and industrial fabric, comprise 30 percent and 28 percent of the settlement respectively. Only one unit was determined to be open space, exhibiting no structural development whatsoever.
The informally designed mahalle of Gumuspinar also exhibited a fairly high built density, though a larger number of grid units had not been developed at all. Low development areas are found along the edges of this large undeveloped area and in the grid units nearest to the major highways. Highly developed units dominate the sample, comprising 50 percent of the total land area. Both mass housing and low-density housing are present in this mahalle, though units with only low-density housing outnumber the units with both types by only two percent. More than 80 percent of the units in the mahalle contained some type of housing.

The percentage of highly developed areas in the blended settlement of Cevizli was similar to that of the other two settlements, at 59 percent. Low development areas comprised 39 percent of the sample area and were most concentrated in the industrial areas near the D-100 highway. Just over 70 percent of the settlement contained some type of housing, with more than 50 percent of the units devoted to low-density housing alone. Eight percent of the units contained only mass housing. Residential fabric was most prevalent in the blended settlement, with nearly 45 percent of the land area consumed by low-density housing.

3.2 Grain Analysis

Photo grids were constructed for each of the sampled mahalleler using the user-generated Panoramio™ photos available on Google Earth. Each mahalle was restricted to a square mile sample, which was then divided into a 49-square grid. From the available imagery, photos were selected from within each square for inclusion in the photo grid. An effort was made to verify the location of each geo-tagged photo using satellite imagery, street maps of the area, and building websites to eliminate redundancies and unusable data. This process generally produced 2 to 3 viable images per gridded square. Banks of photos amassed from each grid were used in conjunction with satellite imagery to create the summary Lynchian images (Figure 4). Paths, edges, districts, nodes, and landmarks (major and minor) were identified for each sample area using Lynch’s method for analysis and comparison (Image 46).

Esentepe, grain

As suggested by the preliminary analysis of the area, the formal settlement of Esentepe contained two dominant edges, the inter-district highway in the north and the industrial pond on the east side of the sample area. The settlement’s users heavily documented both edges and their images confirm the high visibility of each in the landscape. Five large districts were identified from user photos as well, one above the highway and four below it. The most expansive of these districts are industrial in nature, with urban fabric of a considerably larger scale than the surrounding residential areas.

Gumuspinar, grain

The image of the informal settlement of Gumuspinar contained two distinct edges. The first edge separates the open farmland in the northwest corner of the sample area from the residential area below the D-100 highway, while the second edge splits the sample area into east and west sectors, forming a natural barrier between residential districts in the north. A larger number of paths were documented in this mahalle as well. Unlike the previous settlement, these paths intersect the edges present in the sample area, diminishing their potential as barriers to movement. Four districts were documented in Gumuspinar, all exhibiting predominantly residential fabric. Of these districts, only one occurs on the east side of the central highway. It appears to have been formally planned, as it is considerably larger than the others, with radial streets and repeated building footprints. The other districts are smaller and irregular in nature.
Cevizli, grain

The image of the blended settlement of Cevizli also contained four districts, though these were of significantly larger scale than those observed in the other mahalleler. The inter-district highway in the north formed the most distinct edge in the settlement. Its position on the grid not only partitioned the formal housing districts from the rest of the mahalle, but severely limited movement from north to south overall. Only one path was documented above the highway edge, which seems to echo the trajectory of the highway at a scale more conducive to residential movement. Paths between the two lower districts help to solidify the change in fabric from industrial to residential, though a larger number of paths cross between districts in this settlement. More than half of the nodes observed in the mahalle occur at path intersections. Two of these nodes seem to signal a particularly walk-able segment of the blended settlement not seen in other mahalleler.

3.3 Access Analysis

Path network diagrams were constructed for each settlement in two stages. The first diagram aggregated the transit network and vehicular path data collected from Google Earth, the İstanbul Electric, Tramway and Tunnel Department, and the İstanbul Büyükşehir Belediyesi (IBB) into a single square mile map, in accordance with the Allen Jacobs’ method. Though several transit systems exist in Istanbul, public buses called otobüsler had a significant presence in the sample mahalleler. A distinction was made between the transit, highway, street, and road systems, with bus routes receiving the highest position in the visual hierarchy, shown in red (Figure 6). Transit and highway networks are given preference as “higher order” path systems in the remainder of this analysis, with street and road networks designated as “lower order” systems by comparison.

When pedestrian path networks proved too difficult to assess at the scale of the square mile diagram, a second round of stratified sampling was applied to each mahalle. Purposive techniques were again employed, and three grid units were selected from each mahalle to ensure an analogous sampling. The figure ground and higher order paths were reassessed in these samples and combined with pavement data available at the scale of the grid unit. Pedestrian-accessible areas are thus defined in black and gray tones, limited to areas not covered by industrial buildings, residences or higher order paths (i.e. highways or streets). As such, all of the following inferences about pedestrian access within the sampled mahalleler are based on the author’s direct observations of satellite imagery, user-generated photos, and previous diagrams.

4. Translating Urban Performance

Once individual analyses of the sample areas for each construct were complete, a cross-case synthesis was conducted using the criteria defined by Kevin Lynch and Jane Jacobs so that more formal conclusions could be drawn about the performance of urban fabric in each settlement. The qualities of formal, informal, and blended fabric have been coarsely divided in previous discussions, though such divisions are much more difficult to identify at the pedestrian level. In light of this, the following narrative attempts to aggregate and synthesize the data collected from each sample area into more comprehensive statements about the variable levels of density, grain, and access found in the district of Kartal.

Density

Lynch strongly suggests that there is no general optimum density that can be applied to urban fabric, for desired concentrations vary widely by use, age, societal expectations, and countless other considerations (A Theory 262). Despite this tangible ambiguity, density is one of the most primary and influential factors in urban design of any type and its presence or absence
in the district of Kartal merits further discussion. Jacobs writes extensively about the myths of
density and over-crowding. She concludes that despite the warnings of planning literature against
high-density settlements, low-density areas are often dull and lifeless, unable to sustain a diverse
population or amenities (J. Jacobs 204). Jacobs draws sharp lines between dwelling and use
density however, claiming that a dense concentration of people, not structures necessarily, are
required for vibrant urbanism.

Of the mahalleler sampled in this analysis, the informal and blended settlements had the
highest concentrations of housing, though the composition, or grain, of the two sectors were very
different. Gumuspinar was completely dominated by residential use, with nearly an equal amount
of low density and combination housing observed. In contrast, Cevilzi had a much higher
proportion of low-density housing, with less than 50 percent of the settlement devoted to
residential use. The formal settlement of Esentepe had a much lower concentration of housing,
due in part to its even use distribution. This sample area had the largest percentage of highly
developed units however, leaving little undeveloped land for green space or expansion.

The need for open, green, and derelict spaces is addressed more directly in Lynch’s text. He writes that open space is the “negative, extensive, loose, and uncommitted complement to the
[built] system” that is necessary for the “freely chosen and spontaneous activity” required to
sustain city residents (A Theory 396). The informal settlement of Gumuspinar had the highest
percentage of low-development and undeveloped land, units most likely to exhibit spaces of this
type. This percentage was not significantly higher than that of the other two samples, however.
Though dedicated parks were much easier to identify from photos and satellite imagery of each mahalle than derelict and waste areas, the blended settlement of Cevizli seemed to contain the
highest proportion of derelict space.

The farmland observed in the northwest corner of Gumuspinar was the highest single
concentration of green space in the sampled areas. This settlement also contained the largest
number of residential parks and fields overall. The formal settlement of Esentepe exhibited a
much smaller percentage of open and derelict space by comparison, though the industrial pond
on the east side of the sample area created a significant figure in the landscape. The open spaces
that did exist were distributed more evenly throughout the mahalle, however, and may provide
greater benefits to the population than more dense concentrations.

Grain

In Jane Jacobs’ defense of urban diversity, she preys on the master plan mentality,
commenting that “large swatches of construction built at one time are inherently inefficient for
sheltering wide ranges of cultural, population, and business diversity” (191). Both Jacobs and
Lynch describe the ideal urban situation as one containing multiple functions, textures, and
amenities, though their terms for this quality vary. The types of grain defined by Lynch are not
mutually exclusive. The majority of planning professionals agree, for example, that the ideal
grain composition for residential areas would be fine and blurred, so that “each small area [of
urban fabric] should be a microcosm of the whole” (A Theory 267).

Of the mahalleler examined in this analysis, the informal settlement of Gumuspinar
exhibited the highest proportion of finely grained fabric, though relatively little industrial fabric
was observed. The formal housing district in the northern portion of the sample area presents the
 sharpest departure from the mixed residential fabric dominant in the rest of the settlement,
though the presence of the inter-district highway negates the shift. Regardless, the formal
buildings are only slightly larger and more widely spaced than the informal buildings and do not
create an obvious shift in texture.
Transitions between industrial and residential fabric in the blended settlement of Cevizli are much more pronounced, however, and constitute a coarser grain in the settlement overall. Large industrial buildings and waste areas were often separated from low-density residential areas by a single street or path, eliminating the physical and visual transition time Lynch suggests is most desirable in urban fabric. The residential grain in Cevizli was much more fragmented as well. Mass and formal housing was relegated to the north and west areas of the settlement, with few units exhibiting a mixture of high and low-density housing overall.

Though the highest proportion of coarse industrial fabric was observed in the formal settlement of Esentepe, a high number of finely grained residential units were also present in the mahalle. Mass and low-density housing were mixed throughout the residential area with a surprising number of small-scale industrial buildings as well. The presence of multiple grains within a single settlement seemed to suggest a discontinuity between districts, but transitions between the industrial and residential areas of Esentepe were sufficiently gradual to be considered blurred.

Access

Jacobs’ notion of diversity surfaces again in Lynch’s discussion of access, which he calls for alongside equity and control to create a pleasing system. Both theorists agree that access cannot be measured by the sheer number of options or amenities available in a place, for “to have everything instantly available is no more desirable than to live in an infinitely adaptable world” (A Theory 191). Lynch concedes, however, that a good environment is one which “affords obvious and easy access to a moderate variety of people, goods, settings,” with the potential that these choices could be expanded at will (192). To this end, various types of paths and access networks are necessary, “in sufficient number,” to ensure vital urban space (272). Jane Jacobs elaborates on this by listing short blocks and numerous intersections as additional requirements.

A full range of paths was observed in each of the sampled mahalleler, though the types (transit, highway, street, road, and pedestrian) presented at varying strengths in each. The transit network of Cevizli was the most expansive of the three samples, able to serve residential areas (formal and informal) and industrial areas equally. The settlement’s lower order paths were not so evenly distributed. Though larger streets dominated in the formal housing districts of the north, fewer numbers of them were present in the informal housing districts of the south. Small walk-able roads were most prevalent in the central residential area of the blended settlement where pedestrians were more likely to travel and congregate. Unpaved paths were less likely in the industrial areas of the settlement at the pedestrian level, though both paved and unpaved paths were observed in the formal and informal housing districts.

Higher order paths (transit and highways) were more widely distributed in the informal settlement of Gumuspinar, despite having relatively little industrial fabric. Small roads were observed throughout the settlement, facilitating vehicular access even in the most densely concentrated residential areas. A greater number of street and road intersections were observed in this mahalle than the other sample areas combined, suggesting that more path choices exist for the user at this level. Paved and unpaved pedestrian paths were identified throughout the informal settlement as well, due in part to the fine residential grain discussed previously. Walk-able streets and roads were identified in every pedestrian unit, with no visibly restricted areas.

The formal settlement of Esentepe exhibited the least expansive path network of the sampled areas. Street and road systems were extremely limited in the industrial areas above and below the inter-district highway, though large paved areas may provide additional access. The road system is much more dense and complex in the residential districts, however, allowing
greater vehicular access to these areas than the rest of the settlement. Though massive industrial structures dominate the settlement visually, the concentration of lower order paths in the south suggests more consistent movement and use in the residential areas. A similar conclusion has been drawn from the shape and positioning of transit networks in the mahalle. Of the two routes observed, the winding nature of the southern route seems to provide more flexibility to its users than the highway-centric routes in the north, signaling a greater need for transit access in this area. Pedestrian paths were much more consistent. Paved and unpaved paths existed in each of the sampled units, with fewer paved areas observed in the informal residential fabric.

4.1 Concluding Cross-case Synthesis

Each of the initial research questions was addressed in the course of this study, beginning with the identification of the informally and formally designed regions of Kartal seen in Figure 2. From this map, the blended regions of the district were also observed and incorporated into the remainder of the analysis as a third embedded unit. The spatial typologies defined by Kevin Lynch were addressed in the second stage of analysis. Though certain characteristics were identified in each of the sampled mahalleler, more typologies were exhibited by the blended settlement of Gumuspinar than either of the other sample areas.

In the next stage of analysis, three types of structural and spatial densities were observed in the sample regions: built, dwelling, and use density. The built densities of the sampled mahalleler were generally consistent, though the dwelling and use densities varied widely by development pattern. Gumuspinar had an extremely high proportion of residential fabric in comparison to the formal and blended settlements, with other uses each accounting for less than 10 percent of the total land area. In a similar manner, low-density housing existed in much higher quantities in the blended settlement of Cevizli, though the distribution of industrial and blended areas seemed to suggest a more even mixture of uses.

The spatial and architectural grain types defined by Lynch were observed in the fourth stage of analysis and have been analyzed in great detail already. Though none of the sampled mahalleler conformed to the planning ideal of a fine and blurred grain, the blended settlement of Cevizli seemed least amenable. The informal settlement of Gumuspinar exhibited the finest grain overall, but did not contain a sufficient number of large industrial buildings to justify comparison with the other mahalleler on this point. The formal settlement of Esentepe exhibited a fine and blurred grain in certain instances, but as previously mentioned, this was not a widespread phenomenon.

In the final stage, variable levels of pedestrian, vehicular, and transit-oriented access were observed in the sampled mahalleler. Of the three settlements, the most comprehensive transit network was observed in the blended mahalle of Cevizli, though the formal mahalle of Esentepe was moderately well served. The informal mahalle of Gumuspinar exhibited the most comprehensive vehicular network, and seems to have the highest level of pedestrian access as well, though no conclusive determination on this point could be made from the available data.

5. Discussion

The performance construct, though only one of many measurements considered in the field of urban design, remains critical to the analysis and refinement of cities. This study, from its inception, was an attempt to address the complex phenomena involved in the formation of vital and diverse urban fabric through a stratified, descriptive method of analysis, and to determine, to a lesser degree, the benefits and limitations of remote form analysis. The study stands as a substantive-descriptive analysis of selected mahalleler within the district of Kartal, a single vilayet, or province, within the greater city of Istanbul.
5.1 Implications

The implications of this study for the district of Kartal are numerous, as it provides a comprehensive picture of three different types of settlements within its administrative boundaries. The analytical maps produced by the study revealed issues with the access networks in each settlement, as well as certain inconsistencies in the grain and density measurements of each. This study additionally expands the body of data publicly available for analysis in Kartal, a district of Istanbul of considerable interest to the country of Turkey, and the world, as globalization brings increased social and economic attention to the area. This information, and the cursory maps, could prove extremely useful as development and revitalization schemes are formed and implemented in the district in the future.

The study has the potential to serve as an analytical base for further research in the city of Istanbul as questions about the risks and benefits of mixed-use and development arise. By synthesizing the theories of Jane Jacobs and Kevin Lynch into a single, stratified method, this study has not only produced a more rigorous tool for remote analysis, but one that can be applied to numerous other urban settlements in the future. As was previously mentioned, the method relies on western planning theory, not Turkish development patterns and culture. Though certain biases exist, the use of dominant theories in this study ensure that the method is not confined to the borders of Turkish development and culture and can thus be applied to a much larger range of blended settlements.

The substantive-descriptive data produced for the sampled mahalleler of Kartal, though quite rigorous in its own regard, has the potential to provide an even richer base for additional research and planning endeavors. The stratified nature of the method and hierarchical rating system it produces is well suited to Geographic Information System (GIS) analysis, and could be used to create a modeling base for such programs in the future. A model of this type would allow intricate maps of performance criteria to be constructed and interpreted by researchers quickly and accurately, to the benefit of planning professionals across the globe.

6. Conclusion

For any urban analysis to be considered viable it must produce an intimate understanding of the multiple and complex phenomena at work in a city’s growth and development. As demonstrated by this study, no aspect of urbanism can be evaluated on its own, but must be measured in concert with a host of other variables. Repeated analysis ensures that the elements of “good” urbanism are nourished and cultivated over time, while detriments are identified and removed, like weeds in a garden. Performance ratings are crucial in this sense because they beg for iteration; a city’s performance, on any level, must be assessed again and again as development continues.

Descriptive-substantive measurements facilitate the method by providing the most comprehensive knowledge of an urban condition without defaulting to prescriptive generalizations, as demonstrated by this study. Though user-generated settlements have been romanticized and condemned at regular intervals by users and professional planners alike, a critical evaluation of performance avoids such biases altogether. This study not only illuminates the successes and failures of the traditional formal and informal fabric found in the district of Kartal, but also begins to identify what has been gained and lost when the two are merged in blended areas. Such determinations are particularly poignant for the city of Istanbul, and other cities, where cultural heritage is closely tied to the simultaneous chaos and wonder of informal settlements.
Works Cited


COMMUNICATING PRODUCT SUSTAINABILITY: CONSUMER RESPONSES TO SUSTAINABILITY LABELING IN A RETAIL LABORATORY ENVIRONMENT

By Siqi Feng
Sam M. Walton College of Business, Department of Marketing

Faculty Mentor: Dr. Scot Burton
Sam M. Walton College of Business, Department of Marketing

Abstract

In effort to enhance sustainable development, manufacturers and retailers have collaborated to develop a standardized sustainability index based on supply chain life cycle information. However, it is unclear whether this index will help consumers make more sustainable purchases. Research conducted in a retail laboratory addresses consumer attitudes, purchase intentions, and product choices with and without a credible standardized sustainability index, and with or without provision of background sustainability information. Results from a pilot study and two mixed design experiments indicate that, on average, consumers focus more on brand equity than on sustainability levels when they make brand choices. While the disclosure of credible information for brands within a product category affects brand-level sustainability perceptions, there are limited effects on brand purchase intentions and choices. Results also reveal a consumer misconception that nationally recognized brands are more sustainable. Implications of results are offered for producers, retailers, and public policy makers.
Over the past few decades, sustainability has become a mainstream topic worldwide. Although there is increasing consumer awareness of environmental factors, and more people claim to be concerned about the environment, there is still a relatively small number of consumers who are acting on that claim. When making daily purchases, consumers do not seem to be equipped or motivated enough (Young, Hwang, McDonald, & Oates, 2010) to choose sustainable products.

Many companies have established sustainability departments to incorporate sustainability throughout their business. The Sustainability Consortium (2013) is working with many corporate leaders in developing life-cycle assessments to better gauge the sustainability levels of their consumer-packaged goods. Despite these substantial efforts taken by companies, if consumers do not buy more sustainable products, it is likely that companies ultimately will lose their motivation to produce sustainable products. The current state of consumer consumption is not sustainable. Green marketing is needed to increase consumer knowledge and change consumer attitudes and behaviors toward the use of more sustainable products (Cherian & Jacob, 2012); this will be very important to the success of sustainable development.

This study was designed to determine whether consumer product attitudes, purchase intention, willingness to pay a price premium, and quality perception of products are affected by providing consumers with labels and information regarding the sustainability of the products offered at the retail store shelf. The answers to these questions may provide beneficial information to corporations, nongovernmental organizations, and consumers. In other words, all stakeholders will be able to gauge the potential results on consumer behavior by increasing the availability of sustainable information using labels associated with options in product categories.

**Literature Review**

Ajzen’s (1991) theory of planned behavior and Fisk’s (1973) theory of responsible consumption provide insights on the potential effects of labeling on consumer sustainability evaluation, purchase intentions, willingness to pay price premium, and quality perception of products with different sustainability ratings. Ajzen (1991) found that consumer choice behaviors are related to various motivational factors or incentives. Many factors, such as behavioral control, attitude, and social norms, influence purchase intentions, and ultimately, consumer behavior. Thus, he proposed the theory of planned behavior to explain consumer choice. According to this theory, product attributes, such as favorable (unfavorable) information about sustainability, can positively (negatively) affect brand attitudes and purchase intentions and ultimately the actual choice behavior for packaged good products. However, there are many diverse product attributes (e.g., brand awareness, perceived quality, effectiveness, price) that affect brand attitudes, and for many consumers sustainability may be a secondary consideration that has a minimal impact.

Additionally, the theory of responsible consumption suggests that consumers will use limited resources on earth logically and efficiently to support the world’s growing population (Fisk, 1973). Because there are scarce resources on earth, consumers need to be responsible in their consumption behavior so that they will not totally deplete available resources. This theory suggests that at least for some consumers, when provided with information about the importance of sustainability and the details of the sustainability levels communicated by labels, they will prefer sustainable products in their effort to live sustainable lives.

Accurate and reliable eco-labeling is at least potentially important in helping consumers make sustainable decisions by promoting more sustainable consumption. Labels are communication tools to inform buyers of the claims about the product made by the sellers.
Sustainability labeling has multiple functions that vary with the different stakeholders. For consumers, eco-labels provide information about the sustainability levels of products (de Boer, 2003). In regards to sustainability, accurate and objective labeling can be used to provide a type of quality assurance. However, research regarding eco-labels and how they have affected product choices and performance in the marketplace varies.

For example, previous research demonstrates that even though 87% of consumers claim they are concerned about the environment, only one third actually engage in environmental purchases (Oppenheim & Bonini, 2008). One of the main reasons for these behaviors is based on their lack of trust in eco-labels and environmental claims made by companies. In a study conducted by Pricewaterhouse Cooper in 2009, only 16% of consumers reported trusting environmental claims. The remainder did not trust the intent of the companies when making environmental claims, and did not view the information offered by these companies to be credible or convincing (Bybee, 2010).

Additionally, the large number of eco-labels available in the market can be extremely confusing for consumers. Currently, there are over 430 eco-labels ("Ecolabel index," 2012). This contributes to consumer confusion and questions regarding which products are truly sustainable (Seifert & Comas, 2012). Research conducted by the Natural Marketing Institute (2012) revealed that 51% of American consumers believe there are too many green certifications, while 75% believe it is difficult to assess the credibility of the labels. Additionally, 59% of American consumers want just one over-arching label across industries. This will provide them with a simple solution for determining the sustainability level of alternative products on the market, instead of attempting to learn about all of the eco-labels available, understanding their significance, and researching the credibility of the labels (Watanatada & Mak, 2011).

Additional research has considered consumer attitudes and behaviors when eco-labels are present. One study surveyed Swiss consumers to compare existing product attributes, such as the brand and price, to eco-labels in their importance in consumer buying decisions. This study showed differences in responses to products in the lighting and appliance sectors due to the varying degree of involvement for the purchases in these two product categories (Sammer & Wüstenhagen, 2006). Even though there is high consumer awareness for eco-labels, they may not be important enough to influence many purchasing decisions. That is, the labels may only affect behavior for the small segment of consumers who are the most environmentally aware. Their research has found that brand names and equity are important, especially in high involvement categories. In low involvement product categories, consumers seemed more willing to pay a price premium (Sammer & Wüstenhagen, 2006). As shown in this study, the importance of sustainability seems to differ across types of product categories.

In a Regeneration Consumer Study conducted in September and October 2012 by BBMG, GlobeScan, and SustainAbility, in which over 6,000 consumers were surveyed in six major international markets, results indicated that consumers in developing countries are more likely to agree that they need to consume in a sustainable fashion to contribute to a better environment for future generations than consumers in developed countries (GlobeScan, 2012). Likewise, consumers in less developed countries appeared more likely to adopt sustainable behaviors, as compared to those in more developed countries (GlobeScan, 2012). Additionally, research conducted in the United Kingdom and Greece, using Ajzen’s theory of planned behavior, suggests that the theory of planned behavior model appears to be more strongly supported in the United Kingdom than in Greece. The authors suggest that this theory might be more appropriate in more established markets (Kalafatis & Pollard, 1999). Studies conducted in
developing countries, such as Egypt, have illustrated that consumers are becoming increasingly aware of environmental issues. Although sustainability is not their priority, they do show positive attitudes towards the environment (Mostafa, 2007).

**Research Objectives**

Despite prior research, there are only a limited number of studies conducted in the United States that have examined consumer evaluations and choices for actual brands following the disclosure of objective, accurate sustainability ratings for different products in a category. Also, previous researchers have not explored the ability of consumers to accurately estimate the sustainability level of a product without a label when brand competitors on the shelf have sustainability labels present. To address these issues, this research project was conducted in a retail laboratory and examines the following specific questions regarding the effects of sustainability labeling:

1. For brands in a product category, how does the disclosure (versus the absence) of the sustainability levels affect consumer evaluations (sustainability perceptions, product quality), purchase intentions, and choices among the brands in the category?
2. How does the presence of a lesser or a greater sustainability level affect brands with lower versus higher levels of consumer familiarity and equity?
3. Does the presentation of information (disclosed via a newspaper article) emphasizing the importance of consumer sustainability moderate the influence of the presence of sustainability labeling?
4. What inferences do consumers draw about sustainability when brand level information is not disclosed? Are sustainability inferences related to the levels of brand familiarity and brand attitudes?

In contrast to most prior research examining the disclosure of brand level sustainability information, the two mixed design experimental studies are conducted in a retail store laboratory environment, with multiple brands offered on store shelves.

**Pilot Study**

A pilot survey was distributed to determine consumer present perceptions about specific brands used in these two experiments. There were 39 participants in this pilot study. Participants were asked about their overall brand attitude, familiarity, and sustainability perception of products for each of the three brands in two product categories (laundry detergent and dish soaps) that were used in these experiments. Each measure was assessed using a seven-point scale ranging from 1 (*not at all*) to 7 (*very*). For example, for the brand attitude for laundry detergent, participants were asked, “What is your overall attitude of each of the following brands of laundry detergent?”; the categories included Wisk, Era®, and Arm & Hammer® (with endpoints of *very unfavorable* and *very favorable*).

As shown in Table 1, for the category of laundry detergents, the differences in brand attitude, familiarity, and sustainability level for the more recognized brand and other two brands are statistically significant (*F*-values from 5.56 to 19.14, *p* < 0.01 for all). Brand attitude, familiarity, and sustainability level perceptions were significantly higher for Arm & Hammer in the laundry detergent category. Also, the differences between the sustainability level perceptions for all of the laundry detergents are statistically significant. The results were similar to that for dish soaps, where there is a statistically significant difference between the more recognized brand (Dawn) and other two brands, for which evaluations are similar. While the perceptions of the sustainability level of Dawn (*M*=5.03) are stronger than the other brands (*F*=13.6, *p* < .01),
its objective level is lower than Method or Mrs. Meyers, based on GoodGuide (2013) sustainability ratings.

Table 1
Pilot Study Results: Mean Levels for Attitude and Familiarity of Target Brands Used in Experiments 1 and 2

<table>
<thead>
<tr>
<th>Laundry Detergents:</th>
<th>Wisk</th>
<th>Era</th>
<th>Arm &amp; Hammer</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brand Attitude</strong></td>
<td>3.46</td>
<td>3.38</td>
<td>4.41</td>
<td>19.14***</td>
</tr>
<tr>
<td><strong>Familiarity</strong></td>
<td>2.11</td>
<td>2</td>
<td>4.87</td>
<td>56.62***</td>
</tr>
<tr>
<td><strong>Sustainability level</strong></td>
<td>3.56</td>
<td>3.26</td>
<td>4.26</td>
<td>5.56***</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Dish Soaps:</th>
<th>Dawn</th>
<th>Mrs. Meyer's</th>
<th>Method</th>
<th>F-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Brand Attitude</strong></td>
<td>6.03</td>
<td>3.38</td>
<td>3.64</td>
<td>65.12***</td>
</tr>
<tr>
<td><strong>Familiarity</strong></td>
<td>6.36</td>
<td>2.05</td>
<td>2.28</td>
<td>109.23***</td>
</tr>
<tr>
<td><strong>Sustainability level</strong></td>
<td>5.03</td>
<td>3.64</td>
<td>3.61</td>
<td>13.6***</td>
</tr>
</tbody>
</table>

*Note. Means in the table are based on a 7-point scale.*

***p<.01

The correlations between attitude, familiarity, and sustainability level for consumers are positive and statistically significant for all the brands, with the exception of Era, for relationships between attitude and familiarity and familiarity and sustainability level. As shown in Table 1, given no information, brands with a more favorable attitude and greater familiarity are perceived to have higher sustainability levels.

Experiment 1: Laundry Detergent

The purpose of experiment 1 is to examine the influence of the sustainability disclosure on consumer attitudes, purchase intentions, and choices for laundry detergents. For the three brands examined in this experiment, the brand with the highest level of awareness (Arm & Hammer) also had the most favorable level of sustainability based on GoodGuide (2013) ratings.

Methodology

Study design. This research project was part of a larger study conducted in the Walton College retail lab. Participants were asked to examine products and to make evaluations and choices from laundry detergents on the retail shelves. The study was a 3 (brand-level sustainability rating: low vs. moderate vs. high) x 2 [brand-level disclosure: absent (control vs. present)] x 2 [sustainability information provision: absent (control vs. present)] mixed experimental design. The sustainability indices used were based on GoodGuide ratings for the selected brands. Participants in the study were shown laundry detergents designed to match a real retail environment using actual brands (e.g., Wisk, Era, Arm & Hammer). In the brand-level sustainability label present condition, each brand of laundry detergents has a different sustainability rating (e.g., Wisk=3.7, Era=5.4, and Arm & Hammer=7.0), as found in the GoodGuide. The participants in the disclosure present condition saw sustainability labels for each of the laundry detergent brands, while participants in the absent condition did not see
sustainability labels for any of the products. Furthermore, for the second between subjects factor, the sustainability information manipulation, participants read an article about the standardized sustainability index and the importance of sustainability in the information present condition (see Appendix A). In the sustainability information control condition, they read an article about identity theft that was unrelated to sustainability.

**Procedures and participants.** The population from which participants were recruited was college students enrolled in Walton College of Business courses at the University of Arkansas. The participants were recruited by professors who, in return, provided course credit for participation in marketing research studies. For this part of the study, there were a total of 213 participants ($M_{age} = 20.83$, 114 females and 109 males). Due to the increasing focus in higher education on the subject of sustainability, it was predicted that the awareness of participants would be relatively high. This generation potentially can influence the generation before them and the generation to follow. That is, if this group can engage in sustainable behaviors, there is likely to be a ripple effect on the creation of a more sustainable society. The participants in the study were randomly assigned to experimental conditions in the study. The University of Arkansas Institutional Review Board (IRB) approval was obtained prior to collection of the data.

Before entering the retail lab, study participants read a “USA Today” article constructed as part of the study and embedded between two newspaper advertisements not related to the study. The control group read an article regarding identity theft, while the experimental group read information regarding the importance of sustainable consumption and the standardized sustainability index. Copies of both of the articles are provided in Appendix A. Then, participants were escorted to the retail store laboratory where they spent approximately 15 minutes examining laundry detergents and answering questions. During this portion of data collection, participants were asked to go through the product evaluation and selection process and answer survey questions regarding their choices and evaluations. In the control group, the products were presented as currently found on the market (i.e., without Sustainability Index scores). For the sustainability disclosure treatment group, the Sustainability Index score was provided for all of the brands. The Sustainability Index scores of the products were presented as ‘shelf talkers’ next to the products. (The lab set-up for the experiment is presented in Appendix B.) After the participants identified the products they were going to purchase, they were queried about their evaluation of product sustainability, purchase intentions, willingness to pay price premium and quality perception for all of the products in the laundry detergent category. When the participants finished examining the products, they were escorted to a computer lab, where they engaged in a follow-up computer-based survey that took approximately five minutes to complete.

**Measures.** All of the dependent variables were measured when participants were in the retail lab examining the products on the shelves. Participants were first asked which laundry detergent they would choose ("Which laundry detergent would YOU be most likely to purchase?"). Then, a seven-point Likert scale was used to assess purchase intentions ("Assuming you were going to buy a laundry detergent, would you be more likely or less likely to purchase this product?") with the endpoints of 1 (less likely) to 7 (more likely) (Kozup, Creyer, & Burton, 2003). Participant willingness to pay a price premium was measured by responses to the statement, “The price of this laundry detergent would have to go up quite a bit before I would switch to another laundry detergent” [endpoints of 1 (strongly disagree) to 7 (strongly agree)]. The perception of sustainability level was gauged by having participants rate the sustainability...
for each laundry detergent using endpoints of 1 (not sustainable at all) to 7 (very sustainable). Multiple items were used to assess participant perceived quality for the brands [“Compared to other laundry detergents, what is the quality of this detergent in terms of getting your clothes clean?” with endpoints of 1 (much lower than average quality) to 7 (much higher than average quality); 1 (not effective at all) to 7 (very effective); and 1 (poor performance) to 7 (excellent performance)] (Boulding & Kirmani, 1993). These measures were only assessed for the products with the lowest and highest brand-level sustainability labels.

The quality items were combined into a single measure score with an acceptable level of reliability (coefficient α’s > .90). Thus, measures in the study allowed an assessment of participant product choices, as well as repeated measures of product evaluations for two brands offered in the same product category. Following responses to questions completed in the retail lab, participants were escorted to a nearby computer lab where they completed a web-based survey that included questions for manipulation checks, demographic questions, and questions used to identify any possible demand artifacts. Data collected in the retail lab and the computer lab were merged, and subsequently analyzed using SPSS 20.0 software.

**Results**

The manipulation checks were successful and there were no major problems with demand effects. The choice results are shown in Figure 1. The difference between consumers’ purchasing choices in the brand-level sustainability label absent vs. present condition was not statistically significant (Pearson $\chi^2 = 1.83, p = 0.40$). However, for the product with a lower brand-level sustainability label, the purchase choice decreased by 5.8 percent when the label was present. For the brand with the high brand-level sustainability label, choice increased by 3.2 percent.

For the effects on product evaluation and purchase intentions, a mixed design analysis of variance was performed. Results are shown in Table 2. All of the main effects for the experimental independent variables for the dependent measure of sustainability evaluation are statistically significant. Also, all of the dependent measures are significant for the brand-level sustainability label experimental condition. For sustainability evaluation, the two-way interaction

---

**Figure 1.** Purchase Choice in Brand-Level Sustainability Disclosure Present And Absent Conditions

![Figure 1](image-url)
for brand sustainability level and the presence of the sustainability label is significant \((F = 42.96, p < 0.01)\). The plot of means is shown in Figure 2. When the objectively low sustainability level for the Wisk brand is disclosed, its evaluation is reduced, relative to the no label control. For the more well known brand, which also had a more favorable sustainability level (Arm & Hammer), adding the disclosure modestly increased perception.

**Table 2**

*Effects of Brand-Level Sustainability Label and Sustainability Knowledge on Dependent Measures*

<table>
<thead>
<tr>
<th></th>
<th>Sustainability Evaluation</th>
<th>Product Quality</th>
<th>Purchase Intention</th>
<th>Price Premium</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects:</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustainability Level (SLEV)</td>
<td>419.1***</td>
<td>191.6***</td>
<td>271.0***</td>
<td>90.8***</td>
</tr>
<tr>
<td>Sustainability Knowledge (SK)</td>
<td>4.02**</td>
<td>0.01</td>
<td>1.19</td>
<td>1.78</td>
</tr>
<tr>
<td>Sustainability Label (SLAB)</td>
<td>14.04***</td>
<td>2.82*</td>
<td>0</td>
<td>0.36</td>
</tr>
</tbody>
</table>

| **Interaction Effects:** |                           |                 |                    |               |
| SLEV x SK                | 1.04                      | 0.01            | 4.49**             | 3.13*         |
| SLEV x SLAB              | 42.96***                  | 1.03            | 1.96               | 0.51          |
| SK x SLAB                | 7.61***                   | 0.5             | 1.65               | 0.62          |

***p<.01, **p<.05, *p<.10.

*Figure 2.* Brand Sustainability Level and Presence of the Sustainability Label Interaction Effects on Sustainability Perception
The interaction for sustainability knowledge and sustainability label on the sustainability evaluation also is significant ($F=7.61, p < 0.01$). In addition, for purchase intention, the two-way interaction of brand sustainability level and sustainability knowledge was statistically significant ($F=4.49, p < .05$). The plot of means is shown in Figure 3. When the information on sustainability is provided in the article, the purchase intentions of the low sustainability brand (Wisk) decreases, but there is little effect for the more sustainable brand (Arm & Hammer). The interaction of sustainability level and sustainability knowledge did not reach statistical significance ($F=3.13, p < 0.10$), and all of the three-way interactions for this experiment were not significant.

![Purchase Intention](image)

*Figure 3. Brand Sustainability Level and Presence of the Sustainability Label Interaction Effects on Purchase Intention*

The laundry detergents used in Experiment 1 showed interesting results related to brand level sustainability. Arm & Hammer has the highest brand equity as well as the highest sustainability level of the brands selected. The disclosure of actual brand level sustainability has positive influence on consumer evaluation of product sustainability, but a lesser effect on purchase intention and choices of the brands of laundry detergents. As shown in Figures 2 and 3, the sustainability disclosure appeared to have a more negative influence on low sustainability brand (Wisk) than on the higher sustainability brand (Arm & Hammer).

Not surprisingly, the presence of brand level sustainability had the strongest influence on consumer sustainability perceptions. The pattern of results is consistent with the larger literature showing the stronger influence of negative (than positive) information (Baumeister, Bratslavsky, Finkenauer, & Vohs, 2001). The presence of a less favorable sustainability level seems to hurt brands with a lower level of consumer familiarity and equity, while a more favorable
sustainability level only marginally helped the brands with a higher level of consumer familiarity and equity.

**Experiment 2: Dish Soap**

In a voluntary labeling environment, some brands may not include a sustainability level disclosure on the label. When some brands include sustainability information but others do not, what will consumers infer about the brand that does not include a sustainability level? The purpose of Experiment 2 is to extend findings from Experiment 1 by examining consumer inferences and responses when one product label in the dish soap category does not include a sustainability level. Thus, a main component of this study is to assess what consumers infer about the sustainability of a product when there is no sustainability information present for only that one product.

**Methodology**

**Study design.** This experimental method is very similar to Experiment 1. In addition, the research also took place in the retail lab. This research project was part of a larger retail shopping study that asked participants to examine products and to make evaluations and choices from the retail shelves. The study was a 3 (brand-level sustainability rating: low vs. moderate vs. high) x 3 [brand-level disclosure: absent (control), partially present, fully present] x 2 [sustainability knowledge: absent (control) vs. present] mixed experimental design. The sustainability indices used were actual GoodGuide ratings that matched the selected brands. In this experiment, participants in the study were shown dishwashing soaps using the real brands of Dawn, Mrs. Meyer’s, and Method. In the brand-level sustainability label present condition, each brand of dishwashing soap has a different sustainability rating (Dawn=5.0; Mrs. Meyer’s=7.6; Method=8.5; see Appendix B). Please note that in this study, the most familiar brand with the most favorable consumer attitude has the lowest sustainability rating (shown in GoodGuide). The participants in the full disclosure present examined sustainability labels for all of the dishwashing soaps. In the partial disclosure condition, the brand Mrs. Meyer’s was the only brand presented without any sustainability levels. The participants in the absent condition did not examine sustainability labels for any of the brands. As in the first experiment, for sustainability information provision, in the present condition participants read an article about the standardized sustainability index, while in the absent condition they read an article about identity theft. Both articles are provided in Appendix A.

**Procedures and participants.** The participants in this experiment were drawn from the same sample population as Experiment 1. For this study, the mean age was 21 years and 51% of the 213 participants were female. All participants were randomly assigned to experimental conditions in the study. The procedures used for examining the dish soaps on the retail shelf were the same as those used in Experiment 1. In this study, the sustainability information manipulation did not affect the dependent variables and is therefore dropped from further discussion.

The primary difference from Experiment 1 occurred when participants were only shown partial information, with the sustainability index number missing for one brand (as shown in Appendix B). As shown in the pilot test, this brand (Mrs. Meyers) had a relatively low level of familiarity and a weak brand attitude, but its actual sustainability level, based on the GoodGuide, was favorable. After the participants had identified the products they were going to purchase, they were queried about their evaluation of product sustainability, purchase intentions, willingness to pay a price premium and quality perception for all of the products in the dish soap...
category. When participants finished examining the products, they were escorted to a computer lab, where they completed a follow-up computer-based survey similar to Experiment 1.

**Measures.** As in Experiment 1, all of the dependent variables for this experiment were measured when participants were in the retail lab examining the products on the retail shelves. They were first asked which dish soap they would choose (“Which one dish soap would you be most likely to purchase?”). Then, the same seven-point Likert scale was used to assess purchase intentions (Kozup, et al. 2003), participant willingness to pay a price premium, and perception of sustainability level. Two questions were asked regarding the perceived quality of the dish soap (1 much lower than average quality to 7 much higher than average quality) (Boulding & Kirmani, 1993) and (1 not effective at all to 7 very effective). These measures were assessed for all of the brands of dish soap. Participants were also provided with a text box to enter a sustainability rating on a scale of 1-10 for the middle brand-level sustainability label product, Mrs. Meyers (for which no information was provided for two of the three label conditions). This allowed researchers to gather data on participant estimates for this product for each of the absent, partial, and fully present disclosure conditions. After this phase, participants were also asked to complete a web-based survey that included questions that expose possible demand artifacts, assess the manipulations, and provide demographic information.

**Results**

Data from Experiment 2 were analyzed using 2 x 3 mixed analyses of variance; results are shown in Table 3. The main effects of the brand sustainability level on sustainability evaluation and purchase intention were statistically significant. For the two-way interactions, the interactions between brand sustainability level and the sustainability level were statistically significant for sustainability evaluation ($F=7.74, p < 0.01$), while they were not significant for purchase intention. Plots are shown for each of the dependent variables in Figures 4 and 5; as you will see, the pattern is intriguing. For example, as shown in Figure 4, the addition of the label has clear effects on the sustainability evaluation with the familiar brands with relatively lower objective sustainability level (Dawn) decreasing, and the less familiar brand (with stronger objective sustainability levels) benefitting from the information. However, the purchase intention plot in Figure 5 indicates a totally different pattern. The relatively low level of objective sustainability does not diminish the purchase intention for the higher familiarity and higher equity brand (Dawn), while demonstrating little positive effect for the objectively higher sustainability brand.

**Table 3**

*Effects of Brand-level Sustainability Label on Dependent Measures*

<table>
<thead>
<tr>
<th></th>
<th>Sustainability Evaluation</th>
<th>Purchase Intention</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Effects:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sustainability Level (SLEV)</td>
<td>20.56***</td>
<td>104.83***</td>
</tr>
<tr>
<td>Sustainability Label (SLAB)</td>
<td>2.26</td>
<td>0.08</td>
</tr>
<tr>
<td><strong>Interaction Effects:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SLEV x SLAB</td>
<td>7.74***</td>
<td>1.87</td>
</tr>
</tbody>
</table>

***$p<.01$, **$p<.05$, *$p<.10$.**

INQUIRY, Volume 15
In the partial disclosure condition, Mrs. Meyer’s is the only product without brand sustainability label information. When participants were asked to evaluate the sustainability level of Mrs. Meyer’s, its sustainability evaluation score decreased significantly, as shown in the partial condition ($M=3.84$) in Table 4, as compared to the evaluations when the all of the labels were present or absent for the category. Thus, a negative inference is made in the partial...
disclosure condition. However, when all the labels were present for the product category, the score for Mrs. Meyers increased significantly ($p < .05$).

**Table 4**

*Sustainability Inferences when Information is Not Disclosed for a Brand*

<table>
<thead>
<tr>
<th>Sustainability Label Condition</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>4.44$^a$</td>
</tr>
<tr>
<td>Partial</td>
<td>3.84$^b$</td>
</tr>
<tr>
<td>Full</td>
<td>4.91$^a$</td>
</tr>
</tbody>
</table>

*Note.* In the control condition, no sustainability information was disclosed, in the partial condition it was disclosed for two of the three brands, and in the full condition information was offered for all brands on the retail shelf. The sustainability information for the non-disclosed brand (Mrs. Meyers) in the partial condition is favorable (rating=7.5 out of 10). Means are based on 7-point scales; different letters indicate significant differences at $p < .05$.

**Discussion**

Experiment 2 differs from Experiment 1 in that the product with the *high* brand familiarity and equity (Dawn) had a relatively *low* level of sustainability. The disclosure in this case only influenced consumer evaluations of the product sustainability level; purchase intentions and choice for this brand were unaffected by the disclosure of sustainability level. Thus, for this brand with a higher level of consumer familiarity and equity, the presence of a relatively less favorable sustainability level did not lower purchase intention. This result shows that while the disclosure of brand level sustainability information can affect consumer evaluations of the products, it does not necessarily impact consumer purchase intentions or choices. In other words, brand awareness and equity has the dominant influence on consumer purchase intentions and choice.

This experiment also provides some insight into consumer inferences regarding brands that do not offer sustainability information in a voluntary disclosure environment for the dish soaps. In the partial disclosure condition, consumers infer a significantly lower sustainability level for Mrs. Meyer’s, which is the only brand not offering sustainability information. This lower sustainability evaluation may be because consumers believe this brand must not be sustainable, since the producer chose not to share its sustainability information. However, effects may also be related to the placement of the products. Mrs. Meyer’s was placed to the left of Dawn and Method, which had progressively higher brand sustainability levels. Consumers could have inferred that because the Mrs. Meyer’s product was placed to the left of those two brands, its sustainability score was lower. The evaluation for Mrs. Meyer’s was much higher in the full disclosure condition. This suggests in a voluntary labeling environment, consumers will have lower sustainability evaluations for the producers that choose not to disclose their sustainability level.

In addition, although Mrs. Meyer’s had lower sustainability evaluations in the partial condition, its sustainability ratings in the partial and full disclosure condition did not seem to affect consumer purchase intentions. This concurs with the findings in both experiments that a lower/higher sustainability level often may have little influence on consumer purchase intentions and choices.
General Discussion

Even though in the 2012 survey, 57% of consumers in the developed nations who were polled agreed that consumers should consume sustainably in order to improve the environment (GlobeScan, 2012), the results of my two experiments do not offer consistent support that this reported attitude extends to purchase intentions and choices at the retailers’ shelf.

The pilot study results appear to closely correspond with the experimental results in explaining the major effects that initial brand familiarity and attitudes have on consumer purchasing intentions. While there are some effects of the disclosure on intentions in the first experiment, in general, across the brands in the laundry detergent and dish soap product categories, the disclosure of actual sustainability levels seemed to have limited effects on purchase intention and choice. However, in contrast, the disclosure did influence consumer sustainability perceptions. This result suggests that, at least initially, there may be limited effects of retailers or producers providing brand level sustainability to consumers.

For the sustainability evaluation of brands with high brand equity, the presence of a relatively lower sustainability level did not have a substantial negative effect on purchase intention. However, the presence of a higher sustainability did marginally help the higher equity brand. The most substantial effect on purchase intentions was for the brand with a lower level of consumer familiarity and equity; results in Experiment 1 show that a poor sustainability level tended to lower intentions and thus have a negative effect on the brand. This pattern of findings is consistent with much of the literature in marketing and psychology on the asymmetry of information, indicating that negative information has stronger effects than does positive information (Baumeister et al., 2001).

Although there are a number of limitations to these experiments, results provide several potential implications for producers and marketers of sustainable products, particularly the brands with low equity. These producers should focus more on advertising the brand in conjunction with the product sustainability level, because sustainability alone often will not affect consumer attitude and purchase intention for the product.

The presentation of information related to the sustainability index and sustainability only showed an effect in the first experiment. While increasing consumer awareness for the importance of sustainability is very important, one news article is probably not enough to sway an average consumer’s attention and use of the information in forming evaluations and in making choices. In building sustainability awareness, related stakeholders need to find the appropriate communication vehicle and level of information provided to consumers.

In an environment with voluntary sustainability disclosures, the choice not to disclose the sustainability information can harm consumer evaluations of product sustainability. Certainly, brands with favorable levels should consider providing the information to improve the sustainability evaluation of their products. However, product manufacturers and marketers must keep in mind that even when the consumer had an unfavorable initial perception of sustainability that was later affected by a favorable sustainability disclosure, in Experiment 2 consumer sustainability perception had a limited influence on their purchase intentions and choices for a brand with a low level of consumer familiarity. In other words, a favorable sustainability level will not overcome low awareness or a weak initial brand attitude.

While some procedures in this research study attempt to address a number of threats to external validity by using actual brands and a retail store lab environment, there are other limitations to generalizability. Specifically, the retail lab still differs from actual retail
environments in which there are many product options in a category, various types of promotions, and other situational and market variables that may impact evaluations and choices. Also, price information was not provided for the product options used in these experiments. Thus, while these limitations restrict our ability to generalize the findings, they do offer opportunities for additional research.

There are a number of other areas of opportunity for future research. While not discussed in this thesis, the qualitative results regarding participant stated reasons for their choices show that some participants have different perceptions of what is an acceptable sustainability level for brands. Thus, a future research topic could focus on determining acceptable sustainability levels for national brands and how and why these perceptions of acceptability differ across consumers (Cho, Burton, & Soster, 2012). Also, the information manipulation in my experiments demonstrated mixed results. Future research may address the most effective types and levels of public service information or promotion that may have the strongest influence on consumer behavior regarding the sustainability of their choices in a retail environment.

**Conclusion**

In these studies with actual brands that vary in levels of familiarity and consumer equity, sustainability labeling had a limited influence on consumer purchase intentions. It seems likely that typical consumers will often be more likely to make their purchase decisions based on their attitude or familiarity with the brand rather than sustainability of the product. While consumers may be becoming increasingly concerned about the environment, the effects of brand level sustainability ratings are not necessarily reflected in consumer purchase intentions and choices when familiarity and brand equity constructs are considered in evaluations.
References


INQUIRY, Volume 15


Appendix A

Sustainability Information Manipulation

USA Today article used for control condition

Your Identity: Protect It

12:58 p.m. EST January 28, 2013

The FTC estimates that as many as 9 million Americans have their identities stolen each year.

Identity theft occurs when someone uses your personally identifying information, like your name, Social Security number, or credit card number without your permission, to commit fraud or other crimes. The FTC estimates that as many as 9 million Americans have their identities stolen each year. In fact, you or someone you know may have experienced some form of identity theft.

While some identity theft victims can resolve their problems quickly, others spend hundreds of dollars and many days repairing damage to their good name and credit record. You may not find out about the theft until you review your credit report or a credit card statement and notice charges you didn’t make—or until you’re contacted by a debt collector.

Thieves may open new credit card accounts in your name. When they use the cards and don’t pay the bills, the delinquent accounts appear on your credit report. They may change the billing address on your credit card so that you no longer receive bills, and then run up charges on your account. Because your bills are now sent to a different address, it may be some time before you realize there’s a problem.

The best way to find out if your identity was stolen is to monitor your accounts and bank statements each month, and check your credit report on a regular basis.
USA Today article about sustainability index

Standardized Sustainability Index

12:39 p.m. EST January 26, 2013

Consumers can play an essential role in sustainable development by purchasing sustainable products.

In the first half of the 21st century, the human population and power to influence Earth's climate and biology are expected to surge. Increasing consumption levels are thought to be associated with many direct and indirect consequences, including issues related to consumer health, food supplies, natural disasters (e.g., drought, wildfires, storm severity, flooding), water availability, and changes to various biological systems. For several decades, scientists, economists, and other experts have been trying to design strategies for meshing human activities with the limits of the planet and the needs of future generations. This concept, called sustainable development or sustainability, has become a goal of the United Nations, World Bank and a growing list of corporations.

Consumers can help drive sustainable production and play an essential role in sustainable development by purchasing sustainable products. There are a growing number of "eco-labels" providing information regarding the environmental impacts of products and services to help consumers make decisions. However, with over 400 eco-labels, consumers are often confused about what the environmental standards or certifications mean, therefore complicating their decision in making sustainable purchases. Recently, manufacturers and retailers are working together on a standardized index using a common language for product evaluation to provide consumers with transparency of the products they purchase. This standardized index, ranging in values from 1 to 10 to indicate overall sustainability performance, makes it easier for consumers to evaluate products' environmental impacts, and ultimately promote sustainable development.
Appendix B

Retail Lab Examples of Sustainability Disclosure Conditions

Control condition, no sustainability labels
Full sustainability label disclosure

Experiment 2: Partial sustainability label disclosure
While billions of dollars flow into low-income countries each year to ease poverty, assessing their effectiveness beyond the life of a development program proves challenging. In 2009, the World Cocoa Foundation (WCF) undertook the Cocoa Livelihoods Program (CLP) to strengthen cocoa-growing communities with training and credit programs. Using primary 2010-2011 production data collected in Ghana, this study estimates the net present value (NPV) change associated with CLP training over 50 years (two cocoa tree life cycles), assuming producers can utilize what they have learned well after the courses are completed. Applying regression analyses to determine the effect of CLP on yield, findings suggest that average cocoa yield increased 75.24%, resulting in a NPV increase of $401.00 per hectare annually. Given that it costs WCF $252 per participant, this equates to a benefit-cost ratio (BCR) of 79.56:1, or $79.56 dollars in additional cocoa producer profit for every dollar WCF invests into poverty reduction.
Introduction

While billions of dollars flow into low-income countries each year to help alleviate poverty, assessing the effectiveness of these dollars is a challenging task. Because of poor infrastructure and communication networks, as well as a lack of transparency in the sources of information, collecting and evaluating data to measure the impact of development projects in low-income countries is difficult. Meanwhile, the global economic recession, coupled with budget cuts across high-income countries, has resulted in fewer unrestricted funding sources for large-scale development projects (Ozgediz, 2012). Donors for poverty alleviation projects are increasingly asking for higher resolution impact and evaluation data for their projects. Thus, to adequately measure the impacts of a poverty alleviation project, monitoring and evaluation teams must be inherently results-oriented with the data to support claims (United Nations Development Program, 2009).

The literature is rich in studies that measure the benefits of rural development programs. However, many of these studies lack a temporal dimension because they measure costs and benefits for only capital investments and for only a single, static year, while not accounting for skill enhancement dividends paid over a longer horizon. Extending the time horizon beyond a single year requires both knowledge of future yield curves and a quantifiable benefit to the development program, as this study includes. Farmer training programs can involve human capital acquisition and the benefits can persist long after the training program has officially ended. As a result, farmers develop skill sets that can extend well past the single year (or few years) of the training program. By accounting only for net producer benefits during the life of the development program, the cost-benefit analyses (CBA) may not truly capture the full net benefits of a given program.

Therefore, a more comprehensive approach of cost-benefit analysis must be utilized when evaluating projects that invest in human capital. Such analyses should give future donors a more complete portrait of potential investment returns. With that in mind, this study undertakes a cost-benefit analysis of a 2009-2014 Bill and Melinda Gates/World Cocoa Foundation (WCF) training program for Ghanaian cocoa producers. The goal of the training program is to teach cocoa producers in five West African countries agricultural practices such as proper pruning, drying techniques, and harvesting methods to improve their agricultural production and thus their livelihoods. To more comprehensively measure the costs and benefits of such a program, the values should be calculated over an extended horizon, rather than simply accruing the five-year benefits that correspond with the life of the program itself. Net present value (NPV) is a standard measure of intertemporal, net benefits resulting from an investment. By calculating the NPV change due to the human capital obtained, the complete net benefits of the grant and training program(s) can be measured. This type of intertemporal accounting of net benefits makes the full return to grant programs more clear.

In Ghana, approximately 52% of the population lives on USD $2 a day or less and 27% live on $1.25 or less per day. Nineteen percent of rural households produce cocoa; thus, measuring the full impact of agricultural development programs can generate information needed to more efficiently invest scarce resources (Breisinger, Diao, Kolavalli, & Thurlow, 2008; World Bank, 2013).

With the introduction of structural adjustment programs (SAPs) in the 1980s, there was an overall decline in agricultural research, farm extension, and rural banking services that play an integral role in tree crop production enterprises like cocoa in Ghana. To fill this void for cocoa, in 2009 WCF undertook the Cocoa Livelihoods Program (CLP) in conjunction with the Bill and
Melinda Gates Foundation and sixteen member companies involved in the chocolate, cocoa, and coffee industries. The goal of CLP is to increase cocoa production and thereby strengthen the economies of cocoa-growing communities. CLP operates production and management training and credit programs to help accomplish its goals. To estimate the benefits of this program, this study uses primary data collected from the 2010-2011 growing season in Ghana to estimate the impact that the training program has had on producer output and, subsequently, returns. The primary data allowed a comparison between yields and costs for farmers who attended the farmer training and for those that did not. From this comparison, the study implements an NPV model using the 25-year parabola shaped lifecycle yield curve (average productive life) of a cocoa tree in Ghana based on research conducted by the International Institute of Tropical Agriculture (IITA) and Mahrizal, Nalley, Dixon, and Popp (2013). The NPV model estimates the value of CLP training over two production cycles, or a 50-year period, assuming that one hectare is going to be planted after a producer completes CLP training. The hypothesis of the study is that CLP farmers will experience an increase in livelihood quality due to increased cocoa yields associated with farmer training.

**Literature Review**

**Poverty in Ghana**

Real Ghanaian gross domestic product (GDP) has increased 4% annually since 1986, helping real per capita income grow by over 30% for the period 1986 to 2004 (Brooks, Croppenstedt, & Aggrey-Fynn, 2007). Between 2007 and 2011, the annual GDP growth rate was 8.3% (World Bank, 2013). In 2011, the country’s per capita income reached $1,410 and it attained lower middle-income status according to World Bank classifications. However, this increase is more than likely linked to oil discovery and high gold prices, which can lead to unevenly distributed growth and development (World Bank, 2013).

As illustrated in Table 1, food poverty (the estimated food expenditure per person per year needed to meet minimum nutritional requirements hence “extreme poverty”) as well as overall poverty (measured at an income of $1.25 per day) has consistently fallen since 1991 (Breisinger et al., 2008; Ghana Statistical Service, 2000; National Development Planning Commission, 2012). Ironically, farm households experienced a higher incidence of food poverty compared to the national average, ranging from 52% to 45% between 1991 and 1998, respectively. In the past 30 years, the percentage of the poor that produce food crops has increased while the share attributed to export crop producers has decreased (National Development Planning Commission, 2012). Thus, in Ghana, like many low-income countries, those who are the poorest and the most food insecure generally comprise smallholder agricultural producers. Considering the divergent results between food crop producers and export crop producers in recent decades, cocoa production exists as a potential pathway to economic development for Ghana’s rural poor.
Table 1

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>% of Population in Food Poverty</td>
<td>36.5</td>
<td>27.0</td>
<td>18.2</td>
</tr>
<tr>
<td>% of Population in Poverty</td>
<td>51.7</td>
<td>29.5</td>
<td>28.5</td>
</tr>
<tr>
<td>% of Poor Food Crop Producers</td>
<td>57.3</td>
<td>58.1</td>
<td>68.5</td>
</tr>
<tr>
<td>% of Poor Export Crop Producers</td>
<td>7.8</td>
<td>6.9</td>
<td>6.2</td>
</tr>
</tbody>
</table>

In Ghana, 60.1% of cocoa farmers had incomes that were below the poverty line in 1991. By 2007, that figure had dropped to 23.9% (Coulombe & Wodon, 2007). Economic growth has also positively affected poor cocoa farmers more than the poor in other sectors of the economy (Breisinger et al., 2008). Much of this growth can be attributed to improved cocoa varieties. However, hybrids of cocoa may cause greater soil damage than conventional varieties if used without fertilizers, thus necessitating the need for production skill development to properly manage crops and credit access to purchase inputs. In recent years, poverty has actually increased for the more arid, northern regions of Ghana where farmers are less involved in cocoa production, largely due to a decrease in agricultural and non-farm income (Brooks et al., 2007). Many cocoa-growing regions have poverty rates below the national average (Breisinger et al., 2008). Nevertheless, Victor, Gockowski, Agyeman, and Dziwornu (2010) estimate that the average annual per capita income among cocoa-producing households is $153.30, indicating there is still ample room for income enhancement.

**Impact of Structural Adjustment Programs on Cocoa**

In the early 1980s, the World Bank and International Monetary Fund began instituting structural adjustment programs (SAPs) that led to a reduction of government initiatives to “[open] up economic activities to the free play of market force”; the lack of government support led to a decline in the agricultural research, farm extension, and rural banking initiatives that play an integral role in tree crop production enterprises like cocoa (Nyemeck, Gockowski, & Nkamleu, 2007; Wilcox & Abbot, 2006). This decline in public funding was coupled with a decline in official development assistance, decreasing by almost half between 1980 and 2005 when adjusted for inflation. The result included fewer funds to implement agricultural development projects in West Africa and across the globe (Cabral, 2007).

Before the SAPs, many West African cocoa producers received free or subsidized fungicides, herbicides, fertilizers, and technical training. The absence of this support has led to declining yields and increasing yield and income volatility for cocoa producers, particularly for rural poor farmers who live on marginal land susceptible to weather and yield variability (Nyemeck et al., 2007). This deficiency of support can lead to lower output, sale of productive assets, reduced consumption, and/or reduced investments in education if problems persist (Hill & Torero, 2009). Current agricultural loans to Ghanaian cocoa farmers come in the form of input packages such as fertilizer, fungicide, and/or insecticide, primarily through farmer associations.
or non-governmental organizations (NGOs). A larger banking (lending) system that provides credited inputs (providing inputs on loans at the beginning of the season) to more producers has the potential to: (a) ease the capital constraints currently imposed on farmers by smoothing seasonal cash flow deficits that are currently solved by discretionary use of limited resources by households, and (b) improve the ability of cocoa producers to obtain and utilize agricultural inputs (Nyemeck et al., 2007).

**Cocoa Production in Ghana**

Agriculture represented 32.3% of GDP in 2010, the second highest export behind gold (Mhango, 2010; World Bank, 2012). In 2005, cocoa production comprised 18.9% of the agricultural GDP and 7.3% of overall Ghanaian GDP (Breisinger et al., 2008). By 2015, cocoa is projected to account for 16.5% of agricultural GDP and 6.5% of overall GDP (Breisinger et al., 2008). During the 2010 growing season, Cameroon, Côte d’Ivoire, Ghana, and Nigeria together accounted for 71.4% of world cocoa production (World Cocoa Foundation [WCF], 2012). Ghana alone represented 20.5% of global cocoa production in 2010 and was (and remains) the second largest exporter behind Côte d’Ivoire (WCF, 2012). Yet, it should be noted that the number of beans harvested per hectare in Ghana is “among the lowest in the world” (Opoko, Dzene, Caria, Teal, & Zeitlin, 2009).

The Ghana Cocoa Board (COCOBOD) is the sole exporter of Ghanaian cocoa, guaranteeing farmers a minimum price at 70% of the net free on board (FOB) price (Kolavilli, Vigneri, Maamah, & Poku, 2012). In the 1998 growing season, the actual Ghanaian farm gate price as a percent of the FOB price increased to nearly 80% (Kolavalli & Vigneri, 2011). For the 2012 growing season, farmers received 76.04% of the FOB price (Government of Ghana, 2012). Still, net FOB prices in Ghana are lower than its more liberalized neighbors such as Côte d’Ivoire, Togo, Nigeria, and Cameroon (Kolavalli & Vigneri, 2011; Mohammed, Asamoah, & Asiedu-Appiah, 2012). Ghanaian cocoa production is partially liberalized, allowing private licensed buying companies (LBCs) to buy, sell, and transport cocoa. However, COCOBOD sets a minimum price and is currently the only exporter. COCOBOD’s primary LBC competitors are Kuapa Kokoo, Olam, Armajaro, and Global Haulage (Kolavalli & Vigneri, 2011). While LBCs are allowed to export, none have reached the minimum quantity of beans to be eligible to export (Kolavalli & Vigneri, 2011). Given COCOBOD’s predetermined minimum pricing system, the LBCs’ sole option for competing with each other on price is through price bonuses for higher quality cocoa (often linked to a certification program). They can also differentiate themselves through gifts such as exercise books, cakes of soap, salt, subsidized inputs, or credit programs largely implemented through farmer-based organizations (FBOs) like Cocoa Abrabopa (Kolavalli & Vigneri, 2011; Laven, 2007; Opoko et al., 2009). Cocoa Abrabopa is a not-for-profit partner of the Dutch/Ghanaian agricultural company Wienco and provides credit for farmers to buy Wienco agricultural inputs before the season begins. LBCs rarely pay above the minimum COCOBOD price due to the cost associated with doing so (Kolavalli et al., 2012; Seini, 2002).

**The World Cocoa Foundation and the Cocoa Livelihoods Program**

The World Cocoa Foundation is a Washington, D.C.-based NGO with programs in Central and Latin America, Southeast Asia, and West Africa. The Foundation promotes sustainable cocoa production, both economically and environmentally, while also working to improve the livelihoods of cocoa growers and cocoa-growing communities. The Cocoa Livelihoods Program (CLP) is supported by $17 million from sixteen member companies involved in the chocolate, cocoa, and coffee industries such as Hershey’s, Nestle, Kraft, Mars,
Starbucks, and Godiva. Additionally, it has received financial support of $23 million from the Bill and Melinda Gates Foundation, as well as technical support from the German government’s Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ), Canada-based NGO Société de Coopération pour le Développement International (SOCODEVI), U.S.-based NGO TechnoServe, IITA’s Nigeria-based Sustainable Tree Crops Program (STCP), and the governments of Ghana, Liberia, Cameroon, Nigeria, and Côte d’Ivoire.

In Ghana, CLP operates three cocoa farming training programs and one credit operation. The cocoa training programs relate to the following three areas (in this order): production management, business management, and input management. The credit operation provides input loans via TechnoServe, which allow farmers to receive inputs on loans before the season begins. The three training programs are respectively labeled farmer field school (FFS), farmer business school (FBS) and input promoter (IP). When the funding expires in 2013, CLP will have granted credit access to and trained 44,200 Ghanaian cocoa farmers between 2009 and 2013. The number of farmers trained per country was proportional to the share of cocoa production within the five West African countries and multiplied by the 200,000 total farmers to be trained in West Africa. Farmers wishing to participate in CLP are asked to form groups of 15-30 individuals. Further selection criteria are: farmer age not greater than 60 years old, farms with at least 2.5 acres planted with hybrid cocoa with a maximum age of ten years, and farmer has access to at least one hectare of land to establish a new cocoa farm planted with hybrid cocoa.

COCOBOD instructs the FFS curriculum. The immediate impact of FFS should be improved agronomic production skills to better manage the agronomic health of cocoa trees through fertilizer use and prevention of disease and pests. Specifically, FFS provides training in safety practices, fermentation methods, replanting, farming techniques, estimating farm size, pruning, and managing persistent pests like mealy bugs and aphids. FFS also educates farmers on broader social goals such as HIV awareness and children’s education. FFS in Ghana is not a traditional FFS; for example, the curriculum is customized based on preliminary questions to ascertain and address specific farm and farmer needs and deficiencies.

The second phase of CLP is FBS, which is taught by GIZ. FBS provides farmers with the financial tools to balance a budget, work within FBOs, and function as a farmer entrepreneur. The program is primarily concerned with shifting farmer perceptions from farming as a lifestyle to farming as a business. The curriculum accomplishes this by reviewing the farming measurements (hectare, kilometer, kilogram, liters), observing caloric values to ensure families receive the required nutrition, stressing the importance of a balanced budget, practicing balancing a budget, and demonstrating the benefits of crop diversification. The course also evaluates financial services, methods to increase cocoa quality, FBO membership, and the advantages of replanting cocoa. The central message of FBS is that farming is an entrepreneurial activity.

TechnoServe, which also facilitates the credit program, teaches the final phase of CLP: input promoter. The course involves using inputs and, upon graduation, farmers are able to receive input loan packages via TechnoServe at a 10% down payment, underwritten by Opportunity International. The curriculum specifically assesses ways in which the farmer can expand production through the use of inputs, such as chemical fertilizer, fungicides, and insecticides. Safety precautions when spraying and mixing chemicals are also included in the program. By the final phase of CLP, farmers should know proper crop management techniques, how to budget and coordinate financial resources, and how to safely use chemical inputs.
Previous Cost-Benefit Analyses in Development Programs

Several past cost-benefit analyses of tropical agriculture are used for comparisons of the results of this study. Wience’s FBO Cocoa Abrabopa, in conjunction with the Center for the Study of African Economies (CSAE), conducted a study in 2007 to assess the impact of Cocoa Abrabopa’s field representative training and farmer loan program in Ghana (Opoko et al., 2009). The program differs from CLP in that farmers are not provided with training. Instead, Cocoa Abrabopa representatives are trained in production practices like FFS and then go into the field to advise the 11,000 member-farmers. These representatives do not directly sell inputs to farmers, but do provide group-based input loans. Cocoa Abrabopa also gathered information from non-participating farmers to directly compare participating farmers to non-participating farmers. While their study included 239 farmers in the sample, the methods used to collect the data are not clear. The notable results of the study were a recognizable 40% average increase in yield for the 2007/2008 growing season and an economic return of over 250% (benefit cost ratio (BCR) of 2.5) after accounting for the cost of the input loan but excluding operational costs of the program (Opoko et al., 2009). The study found increased labor use was not substantial enough to alter the cost-benefit ratio. More importantly, the study found that incorrect use of fertilizer and other inputs was still a common problem, signifying that credit accessibility is only part of the solution, while training on proper input usage can be as pivotal as the availability of inputs themselves.

Victor et al. (2010) conducted another CBA for cocoa production, estimating the costs, benefits, and NPV of Rainforest Alliance-certified cocoa production in Ghana. Certification requires farmers to adopt medium shade density (70 trees per hectare with a minimum of 12 compatible indigenous species) to “increase biodiversity and other environmental services” (Victor et al., 2010, p. 5). The other major burden of certification is purchasing protective equipment for pesticide mixing and application. The core benefit was the 144 Ghana cedi (GHC) per ton price premium for certified cocoa, a value assumed by Victor et al. The NPV of certification calculated over 15 years for high input, medium shade Amazon-certified cocoa was positive for an 85% FOB price share with a 1.075 BCR and again positive for a hypothetical 25% training yield increase with a 1.087 BCR (Victor et al., 2010). These estimates included a training yield gain and accounted for human development capital that remained unaddressed in prior studies. The study notes its limitations in not incorporating all certification costs and not accounting for future price or cost volatility.

Another cost-benefit analysis conducted by Alam, Furukawa, and Harada (2009) examined a participatory agroforestry program in Bangladesh, intended to combat unregulated, unsustainable deforestation. The study observed financial viability, environmental sustainability, and management issues of a forestry program created to manage farmer needs within forest ecosystems. Participating farmers were allotted one hectare each. Costs were calculated for land preparation, maintenance, pesticides, fertilizer, seeds, and labor. Benefits included income attained from pineapple, ginger, and mustard production, among others. The study found a BCR of 4.12 and an NPV of $17,710 over a 10-year rotation. The researchers illustrated the financial viability of sustainable agroforestry programs (Alam et al., 2009).

A CBA was also conducted for an agricultural development project in Senegal. Weiler and Tyner (1981) analyzed the Nianga area irrigation project that included: 35 mixed agriculture farms, a seed farm, an experiment station, a training program, housing for government workers, and a sheep feed lot, covering 632 hectares of irrigable land in total. They calculated the BCR to be between 1.365 and 0.851 depending on the discount rate (5% or 15%). The authors cited operating expenses and preliminary costs as the primary factors affecting the BCR.
Mahrizal et al. (2013) utilized cocoa production data collected by STCP and IITA to estimate an optimal replacement rate (ORR) and initial replacement year (IRY) to maximize a 50-year NPV for a hectare of cocoa production in Ghana by employing a phased replanting approach. The authors found that the annual ORR is 5% to 7% across the three different production systems studied: Low Input, Landrace Cocoa (LILC); High Input, No Shade Amazon Cocoa (HINSC); and High Input, Medium Shade Cocoa (HIMSC). They also estimated that the optimal IRY ranges from year five to year nine as a function of cocoa prices, fertilizer prices, labor prices, and percentage yield loss due to disease outbreaks. From the ORR and IRY values, the authors estimated economic gains that exceed currently practiced replacement approaches by 5.57% to 14.67% across production systems with reduced, annual income volatility. They concluded their method could be used to increase cocoa yields and stabilize income over time, thus facilitating substantial quality of life improvements for many subsistence cocoa farmers in Ghana and around the world.

Methodology

Data

The primary researcher and two WCF staff members conducted a survey in ten WCF CLP communities in July 2011 in the cocoa growing regions of Ghana; the communities were selected using cluster sampling of three production regions. The villages (district in parentheses) were: Adankwame (Atwima Nwabiagya), Afere (Juaboso), Datano (Juaboso), Bonzain (Juaboso), Ntertreso (Sefwi Wiawso), Domeabra (Sefwi Wiawso), Akim-Aprade (Birim South), Oforikrom (Birim South), Anyinam-Kotoku (Birim South), and Djanikrom (Birim South). Figure 1 illustrates geographic positions.

Figure 1. Location of Cocoa Livelihood Program (CLP) Villages used in the Study. Map Source: ArcGIS (2013).
The University of Arkansas Institutional Review Board (IRB) approved the research. Consent was obtained using modified informed consent and the provision of information was completely voluntary; participants could choose to not participate in the survey at any time. Farmers were compensated with snack food items to incentivize their involvement. All CLP communities were grouped according to training received (FFS, FFS/FBS, or FFS/FBS/Input) and selections for the survey were randomly made within the respective groups. Once the 10 communities were chosen, purposive sampling was employed to select both male and female cocoa producers. Women were intentionally overrepresented in the sample to provide reporting data to donor agencies on female farmers’ practices and yields.

The targeted and attained sample size was comprised of 183 farmers (126 men and 57 women). The sample size was calculated to have approximately 18 farmers from each of the ten communities. The sampling frame was obtained from Fortson, Murray, and Velyvis (2011), who conducted a study by Mathematica Policy Research Inc. during the 2009/2010 cocoa growing season on behalf of WCF to measure yields of farmers “most likely to benefit from the program.” Thus, the sample identified by Mathematica should be representative of cocoa producers in Ghana who are likely to participate in the training program. It should be noted that each community had received some form of CLP training by the time this study was implemented. Of the 549 training units (one farmer graduating from any one of the three programs) experienced by the 183 farmers in our survey, 256 (46.6%) occurred after the 2010-2011 harvest. Because the yields from this group’s farmers were not affected by the training at the point of data collection in summer 2011, they comprise the control group for measuring the impact of the training programs.

The CLP survey was implemented to collect qualitative and quantitative information about the producers and their production behavior. Data collected included: (a) name, (b) gender, (c) age, (d) district, (e) village, (f) total area planted in hectares, (g) FBO membership, (h) total farm yield (measured in 64kg bags), (i) WCF training received including the year, (j) source of planting material for their farm both pre- and post-training, and (k) implementation of different farm management practices. Farm size was based on farmer estimations since many farms were non-contiguous and GPS mapping was not common. Since FFS incorporates a module on the proper measurement of farm size, producer-reported farm size should be a relatively accurate approximation. For observations where multiple family members co-managed a farm, only one manager was interviewed. For farms with both a farm manager and a farm owner in which only one received training, the two were interviewed together. If language barriers existed between farmers and interviewers, a translator was utilized. The questionnaire was administered with the assistance of local technical partners under supervision of the WCF Monitor and Evaluation team.

**Methods and Data**

To estimate the yield enhancement attributable to the various levels of CLP farmer training, a semi-log linear regression model was specified and estimated by ordinary least squares. The dependent variable was yield measured in kilograms of cocoa beans per hectare. The independent variables were FFS training, FBS training, input promoter (IP) training, gender, farm size, FBO membership, fertilizer use, fungicide use, insecticide use, herbicide use, improved cocoa varieties, seed source, and location.

The model can be written as:

\[
\log Y_i = \alpha + \beta_{1i}FFS + \beta_{2i}FBS + \beta_{3i}IP + \beta_{4i}Gender + \beta_{5i}FarmSize + \beta_{6i}FBO + \beta_{7i}Fert + \beta_{8i}Fung + \beta_{9i}Insect + \beta_{10i}Herb + \beta_{11i}ImprVar + \varphi_{1i}SeedSource + \varphi_{2i}Location + e.
\] (1)
The dependent variable $Y_i$ represents yield of dried cocoa beans for individual farm $i$ in kilograms per hectare. A natural log was used because a semi-log regression model calculates the percentage yield increase associated with training (rather than in kilograms per hectare), resulting in a more accurate NPV model. $FFS$, $FBS$, and $IP$ are binary variables taking on a value of one if the $i^{th}$ participant had completed the CLP farmer field school (FFS), farmer business school (FBS) and input promoter (IP), respectively. The control producer group consists of those farmers who had no CLP training. $Gender$ is a binary variable taking on the value of one if the $i^{th}$ participant is male. $FarmSize$ is the natural log of participant $i$’s cocoa farm size in hectares. $Fert$, $Fung$, $Insect$, $Herb$, $ImprVar$, and $FBO$ are binary variables taking on the value of one if the $i^{th}$ participant used inorganic fertilizer, fungicide, insecticide, herbicide, improved cocoa varieties, or was a member of an FBO, respectively. Ideally, the amounts of fertilizer, fungicide, herbicide, pesticide, and insecticide would have been collected. However, given the non-contiguous nature of most producers’ farms, the two growing seasons for cocoa, and that fertilizer may not be applied every year, these data sets were not obtained. The coefficient vector $\varphi_1$ contains coefficients for the origin of seed stock binary variables (own farm and friend’s farm, with government certified seed acting as the control origin) and $\varphi_2$ contains coefficient binary variables indicating the location of the farm (the districts Atwima Nwabiagya, Juaboso, and Sefwi Wiawso, with Birim South acting as the control district). Because of the cross sectional nature of the sample, the standard errors of the estimated coefficients are heteroscedasticity consistent standard errors as provided by White (1980). As a result, the ratio of the estimated coefficients to their estimated standard errors is distributed asymptotically as standard normal under the null hypothesis.

**Net Present Value**

Given the estimated yield increases from the various CLP training programs from equation (1), a net present value (NPV) of total benefits can be calculated using the methods implemented in Mahrizal et al. (2013). Like Mahrizal et al., this study solves for the optimal IRY and ORR. Given this solution, the net future value (NFV) in each year is computed as a function of returns, the replacement rate, year of replacement, and inflation rate. Then, the NPV is computed as the sum of the annual, discounted NFV in each year. This study considers the importance of both the inflation rate (because it is often high in low-income countries), since it increases the nominal price level over time and strongly affects the future value of money, and the importance of the discount rate, since it determines the present value of net returns from future periods.

A baseline NPV was computed using the results of the Mahrizal et al. (2013) study that used the same production data set as this study. A two-dimensional matrix was constructed in Excel with varying annual replacement rates (the ORR) along the columns and an initial year for beginning replacement along the rows (the IRY). Each element in this matrix is the NPV for a given replacement rate and the associated initial replacement year. The ORR ranges from 4% to 10% and the IRY ranges from year 5 to year 20. As Mahrizal illustrates:

Replacing cocoa trees by less than 4% or over 10% indicates that the completion of replacement of an entire farm for one production cycle would take 33.3 to 100 years or 9 years or less, respectively. Setting the IRY at less than 5 years of age or over 20 years of age is not necessary since the cocoa trees bear fruit starting at age three and decreasing yields begin after year 20 (p. 17).

The combination of percentage of replacement rates and IRY, which gives the highest NPV, is the optimal solution. Furthermore, “for all scenarios solved, all optimal solutions were in the
interior of the matrix (i.e., no corner solutions). This justifies having $4\% \leq \text{ORR}, \leq 10\%$ and $5 \leq \text{IRY} \leq 20$ in the search procedure for the ORR and optimal IRY” (Mahrizal, 2013, p. 17).

From the optimal ORR and IRY that maximizes NPV solved for in the Mahrizal et al. (2013) study, a baseline scenario can be computed to estimate the NPV for participants who are maximizing NPV without the benefit of CLP training. A baseline was established using ORR and IRY to highlight the maximum potential profit that could be achieved for producers given current production practices without CLP training. Given the biological life cycle of a cocoa tree, which has a production peak with a decreasing yield over time (see Figure 2), an alternative baseline, not addressed in this study, would be to simply not replace trees, letting the entire orchard reach zero yield, and subsequently replacing all of the trees at once. Following Mahrizal et al. (2013), who concluded that cocoa yield decreases at an increasing rate over time, it is clear that some form of replacement is needed to both stabilize and optimize cocoa producers’ annual returns over time. Thus, the baseline is established using ORR and IRY, implying that producers are acting in a profit-maximizing manner before the CLP training is implemented. This assumption yields a more conservative NPV increase estimation by using a higher baseline.

It is assumed that the yield benefits estimated in Equation 1 are attributable to the various training programs (FBS, FFS, and Input Promoter) and could represent a constant percentage gain associated with each level of training, above those cocoa producers who did not participate in the various CLP trainings (baseline scenario) over the life of the cocoa tree. That is, the constant gain would increase yields at each stage of growth by that percent. That is, at a 10% yield increase level, 100 kg/ha at year 10 would increase to 110 kg/ha while 200 kg/ha at year 20 would increase to 220 kg/ha.

Figure 2. Cocoa yield in a Low Input Land Race Cocoa (LILC) Production System in Ghana
Source: Gockowski et al. (2009) and Victor et al. (2010).
The calculations for net future value, and net present value were made as follows. Net Future Value (NFV) is equal to

\[ NFV_t = Yld_t \cdot (1 + X\%_t) \cdot P_t (1 + r)^t - C_t (1 + r)^t \]

\[ \text{(2)} \]

Where: \( NFV_t = \) Net future value in period t.
\( Yld_t = \) Yield (kg/ha) of cocoa in period t for a given hectare, and depends upon the age distribution of trees on that hectare.
\( (1 + X\%) = \) Yield increase with various CLP training. X=0 represents the baseline yield.
\( P_t (1 + r)^t = \) Cocoa price in period t compounded by inflation rate r.
\( C_t (1 + r)^t = \) Cost of cocoa production in period t compounded by inflation rate r.

The NPV for the 25-year productive life of the tree is computed as

\[ NPV = \sum_{t=1}^{50} \frac{NFV_t}{(1 + r_d)^t} \]

\[ \text{(3)} \]

where \( r_d \) is the discount rate and t runs from year 1 to year 50, or two cocoa production cycles.

The annual average return is calculated by dividing the NPV by 50, giving the annual average present value of profit per hectare per year. The model assumes no salvage value for cocoa trees consistent with Ward and Faris (1968) and Tisdell and De Silva (2008).

Baseline Labor was fixed at GH₵ 3.5 per day per laborer or USD $2.37 (2010 dollars) as estimated in Gockowski, Victor, Dziwornu, and Fredua-Agyeman (2009). Fertilizer, insecticide, and fungicide prices were respectively fixed at GH₵ 14.7 per 50kg or USD $9.98, GH₵ 16.8 per liter or USD $11.40, GH₵ 1.8 per sachet or USD $1.2 (all in 2010 dollars). By setting inflation at 10.26% per year, the prices of labor and inputs would rise at this rate. The baseline exchange rate was held constant at GH₵ 1.47 per USD, per the 2010 average (Mahrizal et al., 2013). A baseline NPV (no CLP training or X=0) is estimated from Gockowski, Victor, Dziwornu, and Fredua-Agyeman (2009) and the optimal ORR and IRY calculated by Mahrizal et al. (2013) of 6% and year 9, respectively.

The baseline production practice chosen for the study was classified as Low Input Landrace Cocoa (LILC) production system (Mahrizal et al., 2013). The system uses unimproved, local landrace cocoa varieties with pesticides and fungicides over the life cycle, but no inorganic fertilizer. Costs and returns are estimated for one hectare of unimproved cocoa planted at 3 x 3 m spacing (1,100 plants per hectare). No nursery costs are incurred as the farm is directly seeded with unimproved LILC cocoa varieties. Typical of most Ghanaian farmers, it is assumed that there is no use of agrochemicals other than those provided by the Government of Ghana’s mass spraying program, which is subsidized by COCOBOD. The amount of pesticides and fungicides used on average for LILC is 0.11 liters of Confidor per year and 31.68 sachets (50 grams) of Ridomil per year, respectively provided by the government. Prices for these inputs were obtained from Victor et al. (2010). The study also assumes that shade levels for the LILC system are 70 shade trees per hectare. The LILC production system is chosen as the baseline because it is popular with impoverished producers who cannot obtain financing for inputs, the very target of the CLP program. Thus, the baseline scenario portrays those producers who implement LILC cocoa production using the optimal ORR and IRY to maximize NPV, but who have had no CLP training. Once a producer has finished input training (IP), it is assumed that they would have access to inorganic fertilizer and fungicide, thus production costs would need to increase as well.
To account for this, all producers who have input training (IP) have associated higher costs of production. Cost estimates for High Input Medium Shade Cocoa (HIMSC) were obtained from Victor et al. (2010). The only difference between the cost estimates of LILC and HIMSC is the use of inorganic fertilizer, fungicide, and herbicide. From these new cost estimates, a more accurate profit can be estimated because the large theoretical yield increases associated with IP should be associated with higher input costs.

Revenue was calculated by multiplying yield in kilograms per hectare for time period \( t \) by the price of cocoa in time period \( t \) in USD per kilogram. Given the COCOBOD marketing board pricing structure, Ghanaian farmers received 76.04% of the FOB price in 2012 so cocoa price was set at USD $2,513.72 per metric ton of beans or 76.04% of the ICCO price of USD $3,305.79 (2011 dollars) per metric ton of beans as observed on May 2, 2011. The COCOBOD retains a portion of the FOB price to reinvest in the cocoa economy in the forms of educational scholarships, input and supply subsidies, and research in an attempt to increase yields and decrease costs. Inflation was estimated at 10.26% based on the annual average inflation in December 2010 (Bank of Ghana, 2011a). The discount rate was 10.67% using Treasury bill rates for a six-month period (Bank of Ghana, 2011b).

**Benefit Cost Ratio**

The difference between the baseline NPV (no training) and the CLP training program estimated NPV in Equation 3 would be the discounted benefits of the training program. Thus, the benefit-cost ratio (BCR) would be equivalent to

\[
BCR = \frac{B_x}{C_{0x}}
\]

where \( B_x \) represents the discounted benefits of the CLP training program minus the baseline NPV (no training) in USD per hectare and \( C_{0x} \) is the total cost of the training program per person assuming all costs of training are incurred at time 0. Training costs for the CLP program in Ghana were assumed to all occur in the first year (year one) of the program. The World Cocoa Foundation estimated costs of the farmer field school (FFS) and the farmer business school (FBS) to be USD $36 and USD $16, respectively, per participant (2010 dollars). WCF also stated that the input promoter training costs USD $200 (2010 dollars) per producer to implement. Therefore, the total cost of farmer training is USD $252.

**Results**

**Regression**

Table 2 presents a summary of average variable values, divided by district. The average farm size was 3.2 hectares. Juaboso had the largest average farm size at 4.2 hectares, while Birim South had the smallest at 2.3 hectares. The average yield in kilograms per hectare was 562.6. Sefwi Wiawso had the largest yield with 854.9 kilograms per hectare. Atwima Nwabiagya had the smallest at 213.2 kilograms per hectare. Of the sample, 68.9% were male, 76.5% completed FFS, 72.1% completed FBS, and 11.5% completed IP. The 11.5% who had completed IP were concentrated in Juaboso and Sefwi Wiawso. It should be noted that the regression data was collected in year two of CLP, a five-year program. By 2014, 100% of participants will have completed all three training courses.
Table 2

*Descriptive statistics for regression analysis.*

<table>
<thead>
<tr>
<th>District</th>
<th>Atwima Nwabiagya</th>
<th>Juaboso</th>
<th>Sefwi Wiawso</th>
<th>Birim South</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Participants $n$</td>
<td>16</td>
<td>59</td>
<td>32</td>
<td>76</td>
<td>183</td>
</tr>
<tr>
<td>Average Yield (kg/ha)</td>
<td>213.2</td>
<td>681.9</td>
<td>854.9</td>
<td>420.3</td>
<td>562.6</td>
</tr>
<tr>
<td>Farmer Field School Training ($FFS$) % (1=trained, 0=not trained)</td>
<td>50</td>
<td>93.2</td>
<td>68.8</td>
<td>72.4</td>
<td>76.5</td>
</tr>
<tr>
<td>Farmer Business School Training ($FBS$) % (1=trained, 0=not trained)</td>
<td>93.8</td>
<td>64.4</td>
<td>96.9</td>
<td>63.2</td>
<td>72.1</td>
</tr>
<tr>
<td>Input Promoter Training ($IP$) % (1=trained, 0=not trained)</td>
<td>0</td>
<td>32.2</td>
<td>6.3</td>
<td>0</td>
<td>11.5</td>
</tr>
<tr>
<td>Gender % (1=male, 0=female)</td>
<td>81.3</td>
<td>55.9</td>
<td>71.9</td>
<td>75</td>
<td>68.9</td>
</tr>
<tr>
<td>Average Farm Size (ha)</td>
<td>2.9</td>
<td>4.2</td>
<td>3.8</td>
<td>2.3</td>
<td>3.2</td>
</tr>
<tr>
<td>Farmer-Based Organization ($FBO$) Membership % (1=FBO membership, 0=no FBO membership)</td>
<td>50</td>
<td>59.3</td>
<td>28.1</td>
<td>32.9</td>
<td>42.1</td>
</tr>
<tr>
<td>Inorganic Chemical Fertilizer ($Fert$) % (1=used inorganic fertilizer, 0=did not)</td>
<td>12.5</td>
<td>84.7</td>
<td>62.5</td>
<td>48.7</td>
<td>59.6</td>
</tr>
<tr>
<td>Fungicide ($Fung$) % (1=used fungicide, 0=did not)</td>
<td>18.8</td>
<td>93.2</td>
<td>59.4</td>
<td>68.4</td>
<td>70.5</td>
</tr>
<tr>
<td>Herbicide ($Herb$) % (1=used herbicide, 0=did not)</td>
<td>6.3</td>
<td>22</td>
<td>25</td>
<td>44.7</td>
<td>30.6</td>
</tr>
<tr>
<td>Insecticide ($Insect$) % (1=used insecticide, 0=did not)</td>
<td>18.8</td>
<td>88.1</td>
<td>53.1</td>
<td>57.9</td>
<td>63.4</td>
</tr>
<tr>
<td>Using Improved Varieties ($ImprVar$) % (1=used improved varieties, 0=did not use)</td>
<td>18.8</td>
<td>66.1</td>
<td>46.9</td>
<td>55.3</td>
<td>54.1</td>
</tr>
<tr>
<td>Certified Seed Source %</td>
<td>18.8</td>
<td>30.5</td>
<td>12.5</td>
<td>36.8</td>
<td>29</td>
</tr>
<tr>
<td>Friend's Farm Seed Source %</td>
<td>68.8</td>
<td>40.7</td>
<td>37.5</td>
<td>23.7</td>
<td>35.5</td>
</tr>
<tr>
<td>Own Farm Seed Source %</td>
<td>12.5</td>
<td>27.1</td>
<td>50</td>
<td>34.2</td>
<td>32.8</td>
</tr>
</tbody>
</table>

*Due to missing observations, $n=138$ for the regression model estimates and percentages.*
Table 3 presents the results of the regression analyses. The R-squared was 0.36, meaning the model can describe 36% of the variability. Seven of the 16 variables (not counting the constant term) were statistically significant at the 10% level or better. Gender was statistically significant at the 5% level, demonstrating that being male was associated with a 33% increase in yield\(^1\), all other variables held constant. This may be correlated with the social status of males versus females in West African societies, particularly with banking access or land ownership, as well as the physical labor demands of cocoa farming. Farm size (measured in natural logs) with an estimated coefficient of -0.04 was significant at the 1% level, meaning that for every 1% increase in farm size, production decreased 0.04%. Considering a farmer’s labor resources are typically finite, it would be expected that yield in kilograms per hectare would decrease as hectares increase, since the farmers have fewer resources to provide to each tree. Fertilizer and insecticide use were also statistically significant at the 5% and 10% levels with a 54% increase and 34% in yield, respectively. Yield would be expected to increase with use of these inputs, given that fertilizer improves soil quality and pests like mirids can cause a 30-40% yield loss.

Table 3. Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>4.92 (7.16)***</td>
<td>Insect</td>
<td>0.29 (1.78)*</td>
</tr>
<tr>
<td>FFS</td>
<td>0.57 (0.86)</td>
<td>Herb</td>
<td>-0.20 (-1.19)</td>
</tr>
<tr>
<td>FBS</td>
<td>0.022 (0.14)</td>
<td>ImprVar</td>
<td>0.13 (1.00)</td>
</tr>
<tr>
<td>IP</td>
<td>0.56 (3.38)***</td>
<td>FrieFarm</td>
<td>-0.19 (-1.53)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.29 (2.31)**</td>
<td>CertSeed</td>
<td>-0.26 (-1.62)</td>
</tr>
<tr>
<td>FarmSize</td>
<td>-0.037 (-3.34)***</td>
<td>Atwima</td>
<td>-0.63 (-3.07)***</td>
</tr>
<tr>
<td>FBO</td>
<td>0.11 (0.74)</td>
<td>Juaboso</td>
<td>0.12 (0.60)</td>
</tr>
<tr>
<td>Fert</td>
<td>0.43 (2.36)**</td>
<td>Sefwi</td>
<td>0.45 (2.88)**</td>
</tr>
<tr>
<td>Fung</td>
<td>-0.094 (-0.53)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: \(n=138\) and \(R^2=0.36\).

*** Denotes statistically significant at the 1% level.
** Denotes statistically significant at the 5% level.
* Denotes statistically significant at the 10% level.
Parentheses denote t-ratio.

\(^1\) Note that the estimated coefficient of Gender is 0.29. Because the dependent variable (yield) is in natural logs, the coefficient of any given variable is the continuous change rate for a one-unit change in the associated independent variable for a dependent variable. But for a binary variable like Gender, the full impact of going from zero to one in a discrete jump requires exponentiating the coefficient, subtracting one, and multiplying this difference by 100 to get the full percentage change when a binary variable goes from zero to one.
The training coefficient estimates provide the most interesting feature of the regression analyses. Attending FFS (farmer field school) was associated with a 77.2% increase in yield, but it was not statistically significant. FBS (farmer business school) had a positive coefficient (2.2% increase in yield with training); however, it was also not statistically significant. The only training that was statistically significant was IP (input promoter), which was significant at the 1% level and associated with a 75.24% increase in yield.

**Net Present Value**

Table 4 presents the annual NPV estimates for the (a) baseline analysis from Mahrizal et al. (2013), (b) for the 75.24% yield increase associated with the statistically significant input promoter (IP) training course found on Table 3, and (c) a sensitivity analysis to provide reference and break-even points. Given that input promoter (IP) is the capstone training course, the percentage yield increase associated with its completion can be recognized as the total yield increase for completing the CLP farmer training program.

<table>
<thead>
<tr>
<th>Yield Increase</th>
<th>Annual Net Present Value (NPV)*</th>
<th>NPV Change ($ per Ha)</th>
<th>Percent Change from Baseline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline**</td>
<td>$445.57</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>75.24%***</td>
<td>$846.57††</td>
<td>$401.00</td>
<td>90.00</td>
</tr>
<tr>
<td>50%</td>
<td>$652.89</td>
<td>$207.32</td>
<td>46.53</td>
</tr>
<tr>
<td>25%</td>
<td>$459.20</td>
<td>$13.63</td>
<td>3.06</td>
</tr>
<tr>
<td>23.25%</td>
<td>$445.57</td>
<td>$0.00</td>
<td>0</td>
</tr>
</tbody>
</table>

* Denotes net present value in 2010 USD per hectare per year.
† The discount rate is based on Ghanaian Treasury bill rates for a six month period in 2010, is 10.67%. (Bank of Ghana, 2011a).
** Equivalent to the Baseline Value in Mahrizal et al. (2013), which is a producer with no CLP training
*** Estimate obtained from Table 3.
††Includes the increased costs used on inputs assumed to be used after input training. Annual total cost increase from use of inputs is 54% or $163.73 per year.

The baseline NPV (Low Input Landrace Cocoa or LILC), as calculated from Mahrizal et al. (2013), was $445.57 per hectare per year for the 50 years of the two production cycles. The NPV associated with the completion of CLP training was estimated at $846.57 or a 90% increase from the baseline. This includes $163.73 per year in increased input costs, modeled after High Input Medium Shade Cocoa (HIMSC) in the study by Victor et al. (2010).

Given that output results could be inflated on an interview-based survey, a sensitivity analysis was also conducted to see how various levels of yield increases affected NPV and what the minimum level of yield increase was needed to at least break even and cover the costs of the increased inputs (see Table 4). Instead of using the estimated 75.24% yield increase as estimated from Table 3 for the completion of input promoter (IP) training, 50% and 25% yield increases...
were selected as reference points to calculate NPV percent gain from the baseline and to compare with the BCR associated with a 25% assumed training gain (1.087) as estimated in Victor et al. (2010). NPV increased 46.53% and 3.06% for the 50% and 25% yield increases, respectively. In these cases, cost increases (54%) were greater than yield increases and thus the NPV increase was smaller than the yield increases. Finally, the break-even yield increase was estimated, finding the yield increase level (23.25%) in which additional revenue would equal the increased input costs, thus yielding a 0% change in NPV. Given the large difference between the estimated 75.24% \( \Delta \text{IP} \) yield increase and the break-even yield increase of 23.25%, these results illustrate a high likelihood of increased producer profitability (Table 4). These figures also suggest farmers would need to artificially inflate their yield by 324% when being surveyed (75.24/23.25) for the additional input costs to negate the NPV gains from farmer training.

**Benefit Cost Ratio**

Table 5 presents the 50-year extrapolations (two cocoa production cycles) of the annual NPV calculations found on Table 4. As such, the table illustrates (a) the total NPV for the baseline scenario (LILC) from Mahrizal’s study (2013), (b) the total NPV for completing the training program (\( \text{IP} \)) utilizing the 75.24% yield increase associated with the statistically significant \( \text{IP} \) training course found on Table 3, and (c) a sensitivity analysis to provide reference points and the break-even point. By comparing the baseline scenario NPV and the training NPV, the NPV gain (benefit) associated with training can be approximated.

Table 5

*Sensitivity Analysis of the Benefit Cost Ratio for the Cocoa Livelihoods Program (CLP) Input Training Course in Ghana.

<table>
<thead>
<tr>
<th>Yield Increase</th>
<th>Total Net Present Value (NPV)*†</th>
<th>NPV Change From Baseline</th>
<th>Total Training Costs**</th>
<th>Benefit Cost Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline***</td>
<td>$22,279</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>75.24%****</td>
<td>$42,329††</td>
<td>$20,050</td>
<td>$252</td>
<td>79.56</td>
</tr>
<tr>
<td>50%</td>
<td>$32,645</td>
<td>$10,366</td>
<td>$252</td>
<td>41.13</td>
</tr>
<tr>
<td>25%</td>
<td>$22,960</td>
<td>$682</td>
<td>$252</td>
<td>2.70</td>
</tr>
<tr>
<td>23.89%</td>
<td>$22,531</td>
<td>$252</td>
<td>$252</td>
<td>1.00</td>
</tr>
</tbody>
</table>

* Denotes net present value in 2010 USD for one hectare over two cocoa production cycles (50 years).
† The discount rate is based on Ghanaian Treasury bill rates for a six month period, or 10.67% in 2010 (Bank of Ghana 2011a).
**Costs are not discounted because they are all incurred in year one of the program.
*** Equivalent to the Baseline Value in Mahrizal et al. (2013), which is a producer with no CLP training.
**** Estimate value obtained from Table 3.
†† Includes the increased costs used on inputs assumed to be used after input training. Annual total cost increase from use of inputs is 54% or $163.73 per year.
When extrapolated over 50 years, the LILC baseline scenario (no CLP training) NPV was $22,279, whereas the 75.24% yield increase (from completing IP) NPV was $42,329, a difference of $20,050 (2010 dollars) per hectare. Therefore, the benefit associated with training represents $20,050 per hectare. With a total training cost of $252 ($36 for FFS, $16 for FBS, and $200 for IP), the benefit-cost ratio (BCR) was calculated to be 79.56:1 (20,050/252). That is, for every $1 invested into the CLP farmer training program, the return on investment (increased NPV per hectare for small scale cocoa producers) was roughly 80 dollars, which is a large return based on any measure, and particularly when compared to the 1.087 BCR from Victor et al. (2010). It should be noted that this is a conservative estimate considering the calculations are on a per hectare basis. If a producer was to produce on two hectares, then the BCR would be 159.12:1 (20,050*2/252). One hectare was the unit of measurement for the NPV/BCR model because a selection criterion for CLP is having access to at least one hectare of land to establish a new cocoa farm planted with hybrid cocoa once training is complete. The BCR ratio provides a clear illustration of the strength of human capital development in poverty alleviation, instilling knowledge in the farmers that can be used well past the year of training while increasing incomes by $79.56 per hectare for every $1 invested in initial training.

A sensitivity analysis was also conducted at the 50% and 25% yield increase levels to provide BCR reference points. Compared to the baseline, LILC model calculated from Mahrizal et al. (2013), the 50% and 25% levels respectively resulted in NPV gains of $10,366 and $682 per hectare over the 50-year period. With a total training cost of $252, the BCR was calculated to be 41.13:1 and 2.70:1. These returns are still well above the break-even ratio of 1.0 and are well below the yield increases reported by producers leading to the notion that these returns are both robust and that investment in the CLP was worthwhile.

To further analyze the benefit cost ratio, a break-even yield increase was estimated that results in a BCR of 1:1. The break-even yield increase necessary for benefits to equal costs was estimated at 23.89%, which includes both the cost of training ($252) and costs of increased input use ($163.73 per year). Any training yield increase less than 23.89% per hectare results in a BCR less than one. Again, this would assume that a producer only produces one hectare. The BCR could be greater than one with a lesser yield gain if they produced on more than one hectare. While most cocoa producers are small scale in Ghana, they typically produce on more than one hectare.

Discussion

While a 50-year extrapolation of a NPV model appears to be too long of a time horizon to use for estimating the benefits of a training program with the analysis based on a one year, cross-sectional model, there are several reasons this time frame was chosen. First, as part of the CLP program, cocoa producers are taught the value of replacing trees instead of letting their yields decline to zero. Unlike most conventional annual crops, cocoa producers have to weigh the benefits and costs of replacing assets where productivity is plateauing or diminishing over time. Given that cocoa trees can yield fruit for up to 50 years but peak at a much earlier age, culling and replanting are considered necessary to maintain maximum orchard profitability over time. However, most impoverished cocoa producers find it difficult to forgo immediate income to enhance long-run revenue potential. Thus, by using a model that extends 50 years (which is typically the cycle of two cocoa trees at 25 years a piece), the model shows the effects that CLP can have on human capital knowledge and replacement rates while potentially providing low-income cocoa producers a higher and less volatile income stream. The importance of this is

2 This assumes there are not multiple people farming the same hectare.
illustrated in Figure 3, which shows that by allowing it to extend well past 25 years, the model is able to fully capture the benefits of the CLP training program in regards to revenue smoothing and eliminating negative profits through replacement training. Per the figure, the baseline is less attractive as it extends beyond year 25, displaying negative profits for years 25-28 and substantially lower profits than phased replanting during years 20-30.

Figure 3. Yearly Profit Per Hectare from Cocoa Production in Ghana Under Medium Shade High Input Production Practices Under Phased Replacement and Status Quo Production. (Status Quo denotes common practice in Ghana where producers simply let yields diminish to zero and then replant the entire orchard. Optimal replacement rate (ORR) denotes the optimal year and percentage of trees to be replaced to maximize NPV.) Source: Mahrizal et. al (2013).

In regards to the regression model, there are several reasons as to why FFS and FBS may not have yielded statistically significant results. FFS is the introductory program to CLP and provides foundational production practices that may not be implemented without additional inputs and sound financial management. Among other concepts, FFS covers safety practices, fermentation methods, and farm size estimation that could lead to a higher quality of life and a higher quality of cocoa bean, but may not necessarily increase yield per hectare.

Additionally, FBS stresses the importance of a balanced budget, demonstrates the benefits of crop diversification, analyzes the caloric intake of farm families, and reviews common farming measurements such as kilograms and hectares. A balanced budget and crop diversification will facilitate a healthier financial position, but like safety practices or fermentation methods with FFS, those practices may not manifest themselves in yield enhancements. It is assumed that ensuring families receive enough calories to subsist and have access to financial services would increase overall quality of life; however, this regression model does not seek to explain quality of life factors, so it is not surprising that FBS and FFS are not statistically significant.

Initially, it was expected that IP would be statistically significant, considering it is the capstone course of three training courses. It teaches farmers how to expand production through the use of chemical fertilizer, fungicides, and insecticides. Upon graduation farmers are able to access the human capital and knowledge base that they obtained from all three programs and,
perhaps more importantly, they qualify for microcredit loans via TechnoServe (>95% of graduates take out loans). The financial skills they attain during FBS could be fully realized if they are able to access credit, and the use of inputs could fully utilize the production skills obtained in FFS. For this reason, the yield increase associated with IP is used with the NPV model to approximate the overall value of training in comparison to the baseline scenario.

Initially, it would seem infeasible for yield to increase only 75% but the NPV to increase by 90%. This is explained by the fact that yield is increasing at a greater rate than cost, 75% compared to 54%. Thus, as long as yield increases at a rate of greater than 54%, NPV gain can be larger than yield gain. This would seemingly indicate that CLP training is an effective way of increasing producer revenue even with the associated new input costs for fertilizer, fungicide, and herbicide. If all 44,200 Ghanaian CLP participants were to experience this gain ($401.00 per hectare), that would result in an annual total gain of $17,724,200 in Ghana alone. For the 52% of the Ghanaian population living on $2 or less a day ($730.00 annually), $401.00 equates to a 54.9% increase in income, a considerable jump by most standards. For the poorest of the poor, the 27% of the population living on $1.25 or less per day ($456.25 annually), $401.00 results in an 87.9% increase in income. Roughly 2% of the Ghanaian population is comprised of poor cocoa farmers, indicating that cocoa production could be a road out of poverty. Based on the calculations from Table 4, it is clear that CLP training is helping to raise incomes for cocoa farmers, ideally leading to improved livelihoods and overall quality of life.

**Conclusion**

In Ghana, where approximately 52% of the population lives on USD $2 a day or less, 27% live on $1.25 or less per day, and 19% of rural households produce cocoa, agricultural development in the cocoa sector has the potential to increase incomes for the poorest of the poor. While billions of dollars flow into low-income countries each year to alleviate poverty, assessing the full impact of these programs can be difficult. For studies that do measure the benefits of development programs, many lack a temporal dimension because they measure costs and benefits in a single, static year or do not account for the full benefit of human capital development. Farmer training programs can provide skill development that is utilized long after the training is complete. Given that the primary intent of the CLP is to increase cocoa yield and farmer quality of life through training in production practices, financial management, and input use, calculating the costs and benefits that extend beyond the five years of the program generates information to more efficiently invest scarce resources.

Using primary data collected in summer 2011 from the 2010-2011 growing season and a baseline model from Mahrizal et al. (2013), the goal of this study was to estimate the NPV of CLP training over a 50-year period—two cocoa production cycles. Using multiple regression analysis to determine the effect of CLP on yield and thus NPV, it was estimated that cocoa yield rose 75.25% per hectare after completing all CLP training. This resulted in an annual NPV gain of $401.00 per hectare or a 90% increase in annual NPV compared to the baseline model. When extrapolated over 50 years to account for human capital development, training is associated with a $20,050 per hectare total increase in NPV. With a total training cost of $252, the BCR of the CLP was 79.56:1, meaning that for every $1 invested in the program, farmers’ income increased by $79.56 per hectare, a considerable increase by most standards.

These results should be considered to represent a conservative estimate given that the costs are fixed at $252; however the benefits vary by farm size. That is, this study assumed that producers only produced one hectare of cocoa. If they produced more than one hectare, the costs remain fixed at $252 per person but the benefits increase, thus increasing the BCR. As noted
previously, the average farm size of the sample was 3.2 hectares. WCF also estimates that training costs decrease over time as training networks are established. The higher costs of the trial programs allow for a more conservative NPV estimate for training. Furthermore, it was assumed that farmers were already maximizing income stability through an optimal tree replacement rate and an optimal initial year of replacement. Farmers who were not optimizing replacement would have lower yield values than the baseline scenario, and thus receive a greater NPV gain after training.

Nevertheless, there are some limitations to this study. Farmers were reported to either use specific inputs or not, but the input application rate was not known. A more accurate study would include specific rates to better compare input use and yield. Collecting this data would likely result in a higher R-squared value. Additionally, the age of the trees was not gathered because of farmers’ inability to recall the ages and replacement rates of all of their plots. Future research should also incorporate a control group that is completely unaffiliated with the training program and that has received no prior training, even training that could not have an effect on yield. This is significant for the self-selection issues that exist within communities that receive training and for the ability of farmers to share CLP skills with other farmers in the community. Finally, the NPV and model are based on one year’s CLP data. Ideally having multiple years with a measure of yield variability would be preferred. This would allow for a range of BCR’s as well as a best and worst case scenario. These four limitations exist largely from the financial infeasibility of conducting a study in West Africa with perfect information on agricultural practices, yield, and cost.

These results can be used by development NGOs to illustrate the potential of skill attainment in alleviating poverty, particularly when encouraging prospective donors, technical partners, or governments to provide financial support. Moreover, by measuring costs and benefits beyond the years of the program, this study provides an established standard in estimating the net present values of other development programs, ideally providing citizens of low-income countries more opportunities to lift themselves out of poverty and contribute to the global economy.
References


ENHANCED TRIBOLOGICAL PROPERTIES OF SURFACES PATTERNED WITH SU8/DLC MICROSTRUCTURES

By Luke Osborn
Department of Mechanical Engineering

Faculty Mentor: Dr. Min Zou
Department of Mechanical Engineering

ABSTRACT

In tribology (the study of friction, wear, and lubrication), it is known that micro-textured surfaces can reduce friction due to decreased contact area between two surfaces. The problem with many micro-textures, however, is their inability to withstand significant amounts of wear, thus limiting their potential uses and overall effectiveness. In this study, connected and isolated microstructures of SU8, a negative tone photoresist, coated with and without diamond-like carbon (DLC), were designed to investigate the effects of these microstructures and DLC coating on the tribological performance of a surface. Friction and wear studies were conducted using a tribometer; surface topography and wear were characterized by scanning electron microscopy (SEM), optical microscope, and contact profilometry. Results indicate that the coefficient of friction and surface wear were significantly reduced by texturing the surface with connected microstructures coated with DLC. Therefore, these microstructures show potential to be used for tribological applications.

KEY WORDS: microstructure, tribology, surface durability, SU8, DLC.

ACKNOWLEDGEMENTS

This research was supported in part by the National Science Foundation grant CMMI-0645040, the Arkansas Department of Higher Education SURF grant, as well as the NSF REU: Nanomaterials and Nanomechanics Program at the University of Arkansas. The authors would like to thank Dr. Ronghua Wei from Southwest Research Institute for depositing DLC films on the samples and Samuel Beckford, Robert Fleming, and Corey Thompson for their support during this study.
Introduction

Over the past several years, there has been a significant expansion in the field of complex micro and nano-devices. Commercially, many of these devices are used as sensors, such as accelerometers for airbag deployment in automobiles. There is a wide variety of uses in consumer, military, and even biomedical applications (Zhang, Tan, & Gong, 2001). Tribological issues, such as friction and wear, have become key factors in determining the functionality and reliability of micro-devices. As systems become smaller, even relatively minute frictional forces can prevent the proper operation of moving parts, resulting in system failure. As the applications of micro and nanotechnologies increase, it is imperative that the friction and wear of these miniaturized devices is curtailed due to their governing nature at the micro-scale.

Silicon (Si) is the most widely used material in micro and nano-devices. Its hydrophilic nature causes high adhesion between microstructures due to large meniscus forces. It is also brittle and cannot withstand applied normal and shearing loads (Williams & Le, 2006). These characteristics translate into Si’s relatively poor tribological characteristics, thus creating a need to develop surfaces with enhanced tribological properties such as reduced friction and increased wear resistance (Liu, Ahmed, & Scherge, 2001; Yoon, Singh, Oh, & Kong, 2005).

Recently, polymers have been studied for use in microsystems because of their low friction, corrosion resistance, and easy fabrication characteristics (Blanco Carballo, Melai, Salm, & Schmitz, 2009; He, Chen, & Wang, 2008; Li, Liu, & Chen, 2005; Yoon et al., 2006). SU8, a negative tone photoresist, has taken the place of Si in novel micro and nanosystems; the reason for this change being that SU8 offers many benefits such as hydrophobicity, biocompatibility, and structural strength (Abgrall, Conedera, Camon, Gue, & Nguyen, 2007; Foulds & Parameswaran, 2006; Lorenz et al., 1997; Seidemann, Butefisch, & Buttgrenbach, 2001). SU8 can easily be fabricated into various patterns using photolithography. It contains a Bisphenol A Novolak epoxy oligomer and triarylsulfonium hexafluoroantimonate salt (Del Campo & Greiner, 2007; Lorenz et al., 1997). Each molecule contains eight reactive epoxy groups, which create a large degree of cross-linking within the polymer once exposed to UV light and developed. This in turn generates strong mechanical, as well as thermal, stability for the SU8 structure (Del Campo & Greiner, 2007). Researchers have found that while SU8 might be very practical for the structural components of microsystems, its tribological characteristics could benefit from improvement, such as adding nanoparticles or texturing the surface (Jiguet, Judelewicz, Mischler, Bertch, & Renaud, 2006; Jiguet, Bertsch, Judelewicz, Hofmann, & Renaud, 2006; Seidemann et al., 2001).

Surface texturing has been studied extensively to enhance the tribological properties of engineered surfaces (Due, Na, Yang, Kim, & Yoon, 2009; He et al., 2008; Morton, Wang, Fleming, & Zou, 2011; Yoon et al., 2006; Zou, Seale, & Wang, 2005; Zou, Cai, Wang, Yang, & Wyrobek, 2005; Zou et al., 2006). It is well known that textured surfaces are beneficial in that they reduce the levels of adhesion and frictional forces due to the decreased areas of contact between two surfaces (Due et al., 2009; He et al., 2008; Menezes, Kishore, Kailas, & Lovell, 2010; Singh et al., 2007; Singh, Duc-Cuong, Kim, Yang, & Yoon, 2009; Zou et al., 2005; Zou et al., 2006). However, there are problems due to the relatively poor wear characteristics of many textured surfaces. These textures often deform or even break when subject to tribological testing (Due et al., 2009; Yoon et al., 2006). Although great progress has been made over the past several years in the area of surface texturing, there is still a need to engineer better, more durable surface textures (Bandorf, Luthje, Wortmann, Staedler, & Wittorf, 2003; He et al., 2008; Jiguet et al., 2006; Tay, Minn, & Sinha, 2011b).
Recent studies examined the tribological performance of patterned SU8 surfaces with and without perfluoropolyether (PFPE) coatings as it pertains to friction and wear (Singh, Siyuan, Satyanarayana, Kustandi, & Sinha, 2011; Tay, Minn, & Sinha, 2011a; Tay, Minn, & Shnha, 2011b). Although it was found that SU8 patterns could be developed to produce low COF, results showed that the surfaces undergo severe wear and damage, thus indicating a need to develop novel surface patterns with enhanced durability (Tay et al., 2011b). It is also expected that the addition of diamond-like carbon (DLC) coatings could enhance the tribological properties of such patterned surfaces because of the known wear resistance and low coefficient of friction of DLC (Ding et al., 2011; Paulkowski et al., 2008; Pettersson & Jacobson, 2004; Singh et al., 2009; Varma, Palshin, Meletis, & Fountzoulas, 1999).

In this study, the tribological performance of surfaces patterned with SU8 microstructures with and without DLC coating was investigated in an attempt to uncover novel engineered surfaces that maintain low frictional values while increasing overall durability through minimizing wear. Surface durability can be defined as the resistance of surface plastic deformation due to the sliding motion of a contacting surface. Textured surfaces with connected and isolated microstructures were examined using reciprocating friction and wear testing. The results show that, by designing and fabricating connected surface microstructures and covering them with a DLC coating, a highly durable surface can be produced for potential uses in novel micro-devices as well as many other applications.

**Experimental Details**

**Materials**

SU8-2050 (Microchem LTD) epoxy based, i-line, negative tone photoresist was used to create microstructures on Si substrates. DLC coating was deposited on top of the SU8 to act as an outer, protective layer to help improve the wear resistance. This concept is illustrated for both the connected and isolated microstructures in Figure 1.

![Figure 1](image.png)

*Figure 1. Schematics of (a) connected and (b) isolated SU8 microstructures covered with DLC (not to scale).*
Fabrication

Si (100) wafers were used as substrates for this study. The Si wafers were cleaned in an ultrasonic bath for 20 minutes with acetone and for another 20 minutes with isopropyl alcohol. The wafers were then rinsed with deionized water, blown dry with nitrogen gas, and baked at 200 °C for 10 minutes to remove any remaining liquid. SU8 films of thickness of about 45 μm were formed by spin coating the SU8 photoresist on the cleaned Si wafer according to the Microchem SU8-2050 recipe, which outlines various spin coating parameters for desired film thicknesses (MicroChem, 2013). A thin layer (< 500nm) of propylene glycol methyl ether (PGMEA) acted as an adhesion promoter between the Si and SU8 film and was spun at 3000 rpm on the Si wafer for 30 seconds. The SU8 film was spun at 2000 rpm for 25 seconds.

The fabrication steps for the SU8 microstructures are as follows: (1) the SU8 photoresist was pre-baked at 65 °C for 6 minutes, 95 °C for 10 minutes, and then cooled for 10 minutes at room temperature; (2) the pre-baked SU8 was then exposed to UV light using photolithography for 140 seconds; (3) the exposed SU8 was then post-baked at 65 °C for 2 minutes, 95 °C for 10 minutes, and then again at 65 °C for 9 hours; (4) the samples were then developed in MicroChem’s SU8 developing solution for 8 minutes to create micro-pillars; (5) the developed SU8 was then hard-baked at 175 °C for 1 hour. Since the SU8 photoresist is to remain as part of the finished product, it was necessary to hard-bake the samples to further cross-link the material to help maintain the good mechanical properties of SU8 (MicroChem, 2013); (6) finally, the samples were annealed at 250 °C for 10 minutes in an attempt to create curvature on the top of the microstructures. By heating the SU8 polymer above its glass transition temperature, which is approximately 200 °C, the microstructures would soften and develop curvature to form a stable, low surface energy geometry (MicroChem, 2013; Tanaka, Umbach, & Blakely, 1997).

Two types of SU8 microstructures were patterned on the Si wafers through photolithography by using plastic dark field masks with circles that had a diameter of 25 μm and center-to-center distances (pitches) of 50 μm and 75 μm, respectively. These values were chosen based on the resolution available for creating the dark field masks. When the plastic mask containing circles with a 50 μm pitch was placed over a Si wafer coated with SU8 film, the circle openings in the mask were close enough that, when the wafer was exposed to UV light, some of the SU8 between the circles also absorbed the UV light. Since SU8 is a negative photoresist, it begins to cross-link and harden when exposed to UV light. As a result, the SU8 film between the microstructures became partially hardened, thus linking the microstructures together. The mask used for the isolated microstructures had a large enough pitch that it did not allow UV rays to harden the SU8 between the pillars, which kept them isolated from one another.

Samples were then coated with DLC using plasma immersion ion deposition (PIID), a plasma-enhanced chemical vapor deposition (PECVD) process. A detailed description of these processes and parameters can be found elsewhere (Wei, 2010; Yukimura & Wei, 2004; Zou et al., 2011).
Tribological Testing and Analysis

An automatic friction abrasion analyzer (Triboster, Kyowa Interface Science Co., LTD), also known as a tribometer, was used to perform friction testing and measure the dynamic coefficients of friction (COF). This machine provides a horizontal, linear reciprocating motion of the sample surface relative to a fixed frictional counterpart.

A 7 mm chrome steel ball (SUJ-2) was used as a counterpart for the friction tests. For the testing, a stroke of 10 mm was used with a sliding speed of 2.5 mm/s. The applied normal load was 0.1962 N, which was achieved by placing a 20 g mass on the balance arm of the tribometer. Each friction test was run for 1000 cycles. The Hertzian model was used to calculate the maximum contact pressure of the ball applied on the flat sample surfaces (Table 1).

The Hertz’s equation is given as:

\[ P_o = \frac{1}{\pi} \left( \frac{6F_n E_c^2}{R^2} \right) ^{\frac{1}{3}} \]  

(1)

where \( P_o \) is the maximum contact pressure, \( R \) is the effective radius of curvature, \( F_n \) is the normal load, and \( E_c \) is the reduced modulus of the two materials in contact and is calculated by:

\[ \frac{1}{E_c} = \frac{1-v_1^2}{E_1} + \frac{1-v_2^2}{E_2} \]  

(2)

\( E_1, E_2 \) and \( v_1, v_2 \) are the elastic modulus and Poisson’s ratio of the two contacting materials, respectively (Hertz, 1881). Poisson’s ratio and the elastic modulus of the chrome steel ball (SUJ-2) are 0.3 and 210 GPa, respectively (Yagi, Kyogoku, & Nakahara, 2003). The Poisson’s ratio of the SU8 was taken to be 0.22 and the elastic modulus as 5.00 GPa (Al-Halhouli, Kampen, Krah, & Buttgenbach, 2008; MicroChem, 2013; Teh, Durig, Drechsler, Smith, & Guntherodt, 2005). For Si, the value used for Poisson’s ratio was 0.27 and the elastic modulus was 160 GPa (Staedler & Schiffmann, 2001b). For the microstructured sample surfaces, Hertzian contact analysis shows that the contact area is more than the cross section of a single microstructure. Therefore, based on the pitch of the SU8 microstructures, the number of microstructures per unit area was calculated and then multiplied by the estimated contact area between a flat surface and the chrome steel ball. This gave the number of microstructures in contact with the friction counterpart. By multiplying the number of contacting microstructures with their top area, the actual contact area was estimated for the connected and isolated microstructure samples. It should be noted that there is a degree of estimation in the calculated area and the actual area of the microstructure cross sections due to the variations in each individual microstructure. This
calculated contact area was used to determine the approximate maximum contact pressure, which is also shown in Table 1.

An optical microscope (ME300TZ-2L-9M, AmScope) and a scanning electron microscope (SEM, JSM-6335F, JEOL) were used to characterize the topography and wear of the samples. A Dektak surface profilometer (DekTak 150, Bruker) was also used to analyze sample topography.

Table 2
Sample Description

<table>
<thead>
<tr>
<th>Sample</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Si</td>
<td>Smooth Silicon</td>
</tr>
<tr>
<td>SU8 Film</td>
<td>45 μm thick SU8 Film on Silicon</td>
</tr>
<tr>
<td>Connected Microstructures</td>
<td>25 μm diameter with 50 μm pitch on Silicon</td>
</tr>
<tr>
<td>Isolated Microstructures</td>
<td>25 μm diameter with 75 μm pitch on Silicon</td>
</tr>
<tr>
<td>Si + DLC</td>
<td>Smooth Silicon + 500 nm DLC</td>
</tr>
<tr>
<td>SU8 Film + DLC</td>
<td>45 μm thick SU8 Film + 500 nm DLC</td>
</tr>
<tr>
<td>Connected Microstructures + DLC</td>
<td>25 μm diameter with 50 μm pitch + 500 nm DLC</td>
</tr>
<tr>
<td>Isolated Microstructures + DLC</td>
<td>25 μm diameter with 75 μm pitch + 500 nm DLC</td>
</tr>
</tbody>
</table>

Results

Sample Topography and Characterization

Table 2 describes the samples and their classifications. Figure 2 is an SEM image of the connected and isolated microstructures. The diameter of the surfaces’ structures was intended to be 25 μm; however, the connected and isolated microstructures, as seen in Figure 2, have diameters of 32 μm and 30 μm, respectively. The connected microstructures have a slightly larger base than the isolated micro-pillars. The SU8 thickness of 45 μm was chosen to ensure that the Si substrate had minimal effect on the tribological results of the SU8 surface. The DLC film thickness is approximately 500 nm and was measured using a profilometer (DekTak 150, Bruker). The DLC thickness was chosen to maximize surface durability while minimizing COF. Profilometer scans, as seen in Figure 3, show the topographical makeup of the samples textured with microstructures. The non-vertical slope of the structures in Figure 3 is a result of the conical profilometer tip and its limited resolution. The isolated microstructures are not actually touching one another as Figure 3(b) might suggest. The tip itself has a radius of 12.5 μm, which causes the microstructures to appear closer to each other than they really are (compared to Figure 2). The profilometer scans are useful in that they can accurately measure the height of the microstructures, which is about 40 - 45 μm as seen in Figure 4, despite being limited in accurately determining the spacing between the structures. It should be noted that the textures’ non-uniform horizontal dimensions, both in spacing and top diameter, as seen in Figure 4 is a result of the profilometer scan and not complete alignment with the rows of the microstructures. This slight misalignment of the scanning tip causes each texture to have a different diameter and spacing between its neighboring textures when viewed as a 2D scan. This also explains why the dips between the microstructures, as measured by the probe, are not at the same level with each other.
Figure 2. SEM images of (a) connected and (b) isolated microstructures.

Figure 3. Profilometer scans of the connected (top) and isolated (bottom) microstructure samples before DLC coating.
The connected microstructures [Figure 1 (a)] are linked together by SU8 film near the top of the microstructures, thus creating a lower texture aspect ratio than the isolated microstructures [Figure 1(b)]. This extra connection offers structural reinforcement for the microstructures, which in turn enhances the overall durability of the surface.

**Friction**

Friction tests were conducted in a laboratory setting in ambient air with a temperature of 21 ± 2 °C and relative humidity of 40 ± 5%. Multiple runs were taken for each type of surface and the results were averaged. Figure 5 is a plot of COF versus testing cycles for all samples without DLC coating. Both micro-textured samples show lower COF than those of the smooth Si and SU8 film, which have very high COF at about 0.8. Initially, both connected and isolated microstructure samples show significantly lower COF values. The connected microstructure surface shows a very steady, yet increasing COF curve. The COF of this surface approaches the values of both Si and the SU8 film after about 800 cycles. On the other hand, the isolated microstructure surface behaves erratically, in terms of COF, despite its slight, steady decline after a surface failure around 150 cycles.

DLC is known for its tribological benefits, such as slickness and hardness, and has a major effect on friction of the samples as seen in Figure 6. Si with DLC coating shows a significant reduction in COF as it drops from 0.8 to 0.4. However, the COF has large variations with rubbing cycles. Although it benefits from the DLC deposition, the SU8 film has a COF slightly higher than that of Si + DLC. The connected microstructures showed the lowest COF values while maintaining relatively steady COF over the entire 1000 cycles. On the other hand, the isolated microstructure sample, even when coated with a protective DLC layer, has higher, slightly increasing COF.

**Figure 4.** Cross-sectional profilometer scans of the connected (top) and isolated (bottom) microstructure samples.
Figure 5. COF vs. sliding cycles for samples without DLC coating.

Figure 6. COF vs. sliding cycles for samples with DLC coating.
Wear

For the purposes of this study, wear is quantified as the plastic deformation, including destruction and texture removal, of a surface. Although it is difficult to summarize the entire wear track of a particular surface with an individual image, the figures chosen best represent the overall wear performance of a sample. Optical and SEM images of the wear track were taken to characterize the wear performance of each sample. Figure 7 shows optical and SEM (insets) images of the Si and SU8 film samples both with and without the DLC coating. Figure 8 shows a magnified SEM image comparison of SU8 and SU8 + DLC wear tracks. The wear on the connected and isolated microstructure surfaces is seen in Figure 9 and Figure 10.

Si and SU8 film have similar friction values, but the wear images suggest that the SU8 film is tribologically more beneficial. The smooth Si shows a wear track width of approximately 200 μm with large amounts of damage and debris [Figure 7(a)]. On the other hand, SU8 film has a much smaller wear track width of about 70 μm and less debris [Figure 7(c)]. There is more surface damage on the Si compared to the SU8 film because of Si’s brittle nature. Once coated with DLC, Si shows reduced wear track width as well as reduced damage and debris, as seen in Figure 7(b). Conversely, SU8 + DLC shows higher amount of wear debris and its wear track is narrower than the uncoated SU8 film due to increased surface hardness resulting from the DLC coating [Figures 7(d) and 8].

Figure 7. Optical and SEM (insets) images of (a) Si, (b) Si+DLC, (c) SU8 film, and (d) SU8 film+DLC.
The connected microstructure sample, Figure 9(a), shows exceptional wear resistance with a very miniscule wear track (~70 μm wide) with small amounts of debris. Even when zoomed in to a magnification of 1000x in Figure 10(a), the structures appear undamaged. Likewise, when coated with DLC [Figures 9(b) and 10(b)], the connected microstructures show excellent wear resistance. The smaller wear track (~19 μm) and noticeably less debris can be attributed to the DLC layer on top of the connected SU8 microstructures. The wear track for the connected microstructures with DLC is nearly impossible to see in the optical image, Figure 9(b), because it is so small. The SEM is able to capture what little wear there is on this sample, as seen in Figure 10(b). Although the connected microstructures prove to be very wear resistant, it is obvious that the additional layer of DLC is still beneficial in that it reduces both wear and friction.

Unlike the connected microstructures, the wear track of the isolated microstructure sample is very wide, stretching to about 1 mm [Figures 9(c) and 10(c)]. Even when coated with DLC, the isolated microstructures were very much damaged [Figures 9(d) and 10(d)]. The DLC helped prevent complete removal of the structures, as seen by the slight bumps at the bases of where some structures once stood, but was overall ineffective in keeping the structures standing. The optical image of the uncoated isolated microstructure sample [Figure 9(c)] shows several fallen microstructures inside the wear track.

*Figure 8. Magnified SEM wear track images of (a) SU8 and (b) SU8 film+DLC.*
Figure 9. Optical wear track images of (a) connected, (b) connected + DLC, (c) isolated, and (d) isolated + DLC microstructures.

Figure 10. SEM wear track images of (a) connected, (b) connected + DLC, (c) isolated and (d) isolated + DLC microstructures.
Discussion

Smooth Si and SU8 Film

The brittle nature of Si causes a higher degree of wear in comparison to the SU8 film. However, the DLC coating has a more profound effect on Si than it does on SU8 film, as seen in Figure 7. The noticeably higher amounts of debris on SU8 + DLC [Figures 7(d) and 8(b)] are caused by a higher contact pressure due to the reduced area of contact that is generated by what is known as a plate bending effect; this occurs when a hard coating is placed on top of a softer substrate, in this case the DLC on SU8 film (Staedler & Schiffmann, 2001a; Staedler & Schiffmann, 2001b). The addition of DLC to the smooth surfaces is beneficial in terms of reducing frictional forces; however, it provides better wear resistance improvements for the Si surface when compared with the SU8 film surface.

Connected and Isolated Microstructures

The smaller COF exhibited by the micro-textured samples is because of the reduced area of contact [Figure 5]. By successfully reducing the area of contact between the sliding counterpart and the sample, the COF was reduced. As the test progresses, the surface can produce small amounts of debris due to wear, which causes COF to rise. In the case of the connected microstructure sample, the surface most likely developed enough debris to affect the sliding counterpart and increase the COF. While it is very difficult to see the wear track in Figure 9(a), Figure 10(a) illustrates a clearer buildup of debris along the wear track. Also, SU8’s relatively soft nature leads to a large deformation, which causes an increase in surface contact area. This is reflected from the COF data since the connected microstructure sample begins to approach the friction behavior of SU8 film.

In the case of the isolated microstructure sample, the initial positive slope and erratic behavior indicates a surface failure soon after testing. Surface characterizations after the friction test clearly indicate the isolated SU8 microstructures were not able to withstand the horizontal shearing of the counterpart, as seen in Figures 9(c) and 10(c). The area of adhesion of SU8 on the Si substrate is directly related to the mechanical strength of the microstructures. In a recent study, it was found that the adhesion strength of SU8 on Si grows not linearly but almost exponentially with an increased contact area between a substrate and an SU8 structure (Blanco Carballo et al., 2009). In fact, it was shown that a square texture with a side of 100 μm could be sheared off an aluminum (Al) substrate with a force of as little as 20 gf (Blanco, Carballo, et al., 2009). Even though SU8 has better adhesion to Si than Al, the sheer force required to remove a small structure, such as the isolated microstructures in this study, would be relatively small (MicroChem, 2013). This being the case, the isolated microstructures were easily removed from the Si substrate mainly because of the small contact area for adhesion between the two materials. On the other hand, the connected microstructures are structurally strong because of the larger adhesion area between the SU8 and the Si as well as the microstructure connections to each other.
It should be noted that the COF curve of the isolated microstructure sample begins to decrease after the initial surface failure. This is likely due to the removed microstructures, as seen in Figures 9(c) and 10(c), rolling along the surface in the wear track, thus almost acting as roller bearings with much reduced friction for the point contact. In order to confirm the assumption of the removed structures acting as roller bearings, it would be necessary to evaluate video footage of the friction tests themselves. Unfortunately, this is beyond the current capabilities for this study. Also, despite the increased contact pressure, this would in turn further reduce the contact area between the sliding counterpart and the sample and thus reduce friction even more. Similar to the isolated microstructure sample without DLC, the sliding point contact has damaged and removed the DLC coated isolated micro-textures, as seen in Figures 9(d) and 10(d). However, in the case of isolated microstructure coated with the DLC layer, there was most likely a larger plowing effect caused by the counterpart as it pushed broken and damaged textures along the wear track, which would cause higher COF than the smooth surfaces and the connected microstructure surface [Figures 6 and 9(d)]. This was observed by a large buildup of removed structures at the edges and ends as well as along the wear track. Although textures were destroyed and removed, the relatively high friction values for the isolated microstructure sample resulted from no broken microstructures rolling beneath the sliding point contact to act as roller bearings during these friction tests.

It should be noted that the connected microstructure surface appears different when comparing optical and SEM images. This visual discrepancy is because of the greater depth of focus an SEM has over an optical microscope. The combination of optical and SEM images of the wear tracks provides multiple perspectives of the wear tracks to give a better overall understanding of the samples’ tribological performance.

Conclusion

It was confirmed that the poor friction and wear characteristics of Si are greatly improved with a layer of DLC. This coincides with similar findings in various other studies (Liu et al., 2001; Singh et al., 2009; Varma et al., 1999). In addition, SU8 film has better wear characteristics compared to Si because of its relatively soft nature. The application of a DLC layer on the SU8 film reduces COF and wear track width but is not as tribologically beneficial as Si + DLC.

The isolated microstructures were completely removed from the surface after wear testing. The DLC coating did not help prevent structure damage. The connected microstructures showed an initial reduction in frictional forces before they began to steadily rise toward values similar to Si and SU8 film. The deposition of DLC helped the connected microstructure surface maintain a low COF. Also, the connected microstructures provide a sturdy, reinforced surface that shows minimal wear. Overall, surfaces patterned with connected SU8 microstructures coated with DLC exhibit low COF, while at the same time creating very little wear or debris. Surfaces such as these show promise for lowering frictional forces and simultaneously preventing extreme wear.
References


LIST OF FIGURE CAPTIONS

Figure 1. Schematics of (a) connected and (b) isolated SU8 microstructures covered with DLC (not to scale).
Figure 2. SEM images of (a) connected and (b) isolated microstructures.
Figure 3. Profilometer scans of the connected (top) and isolated (bottom) microstructure samples before DLC coating.
Figure 4. Cross-sectional profilometer scans of the connected (top) and isolated (bottom) microstructure samples.
Figure 5. COF vs. sliding cycles for samples without DLC coating.
Figure 6. COF vs. sliding cycles for samples with DLC coating.
Figure 7. Optical and SEM (insets) images of (a) Si, (b) Si + DLC, (c) SU8 film, and (d) SU8 film + DLC.
Figure 8. Magnified SEM wear track images of (a) SU8 and (b) SU8 film + DLC.
Figure 9. Optical wear track images of (a) connected, (b) connected + DLC, (c) isolated, and (d) isolated + DLC microstructures.
Figure 10. SEM wear track images of (a) connected, (b) connected + DLC, (c) isolated and (d) isolated + DLC microstructures.
PREPARATION OF NEW VERSATILE AND IMPLANTABLE TITANATE NANOFIBER-BIOSCAFFOLDS VIA EFFICIENT CATION EXCHANGES

By Keaton Piper
Department of Chemistry and Biochemistry

Faculty mentor: Dr. Z. R. Tian
Department of Chemistry and Biochemistry

Abstract

For the first time in the implantable biomaterials field, efficient hydrothermal cation exchange methods were developed suitable for mass-production of implantable material. The cation exchange conditions were performed by various methods of temperature, pressure, and time exposure. To confirm both chemical and physical alteration of the scaffold, XRD shift analysis, EDX atomic composition, and SEM imaging were conducted. The optimal results are to use aqueous solutions of 0.5 M chloride salts in sealed heated containers at 180 °C for eight hours. Cation exchange capacities are determined for K⁺, Na⁺, Li⁺, Ca²⁺, and Sr²⁺; future studies demanding durable, highly specific composition of a biocompatible material are needed.

Acknowledgements

I would first like to extend my gratitude to Dr. Z. Ryan Tian for his faculty mentorship over the past two years of this project. I would also like to recognize the diligent efforts of my colleagues, Alexandra Chilton, Brady McLaughlin, Cale White, and Jared Hopkins, who assisted with data collection and the interpretation of findings in this project. Finally, thank you to the University of Arkansas Chemistry and Biochemistry Department at Fayetteville for facility use. Funding of this project was made possible by the University of Arkansas Honors College Research Grant.
Introduction
Titanium alloys are currently the most common orthopedic and dental implant material (Elias, Lima, Valiev, & Meyers, 2008). In biomaterials science, titanium materials are frequently used; they have been studied extensively because of their low toxicity, anticorrosion properties, and their capacity for photocatalytic sterilization (Bavykin, Friedrich, & Walsh, 2006). One current barrier to in vivo titanium use is poor long-term cell adhesion to the unaltered smooth titanium (Elias et al., 2008). While the use of bone cement and coatings has emerged as a temporary solution, fatal complications have not been eradicated. New obstacles of nerve root damage and pulmonary embolism from bone cement leakages have caused the Federal Drug Administration (FDA) to issue a strong advisement against the use of bone cement in load bearing application (Schultz, 2004). The need for a new multifunctional implant material is pressing and the titanate nanofiber-bioscaffold in this study shows great promise.

Various modified nanostructures, such as those with rigid surfaces that provide enhanced focal adhesions, have been observed to promote cellular adhesion and proliferation, leading to longer lasting implants (Biggs, Richards, & Dalby, 2010). The unique physical and chemical characteristics offer advantages extending beyond medical applications to many industrial applications such as solutions that need quick and specific exchanges of cation (Larsson et al., 1996).

One such structure is a titanate nanofiber-bioscaffold that is thermodynamically stable, low cost, robust, and capable of supporting cell-proliferation due to the nanofiber-assembled macroporous scaffold (Dong et al., 2007). Such titanate nanofiber-bioscaffold can be seen in Figure 1(a); it is made of titanate nanofibers grown on titanium in a manner to form a macroporous structure. The titanate nanofiber-bioscaffold has been previously shown to sustain mesenchymal stem cell and osteoblast cell growth as well as exhibit drug release properties, further propelling it as a future implant material (Dong et al., 2006). However, the titanate surface has a high basicity that is deadly for cells right after the alkali hydrothermal synthesis. In order to improve upon this bioscaffold, variables in the synthesis and post-synthesis procedures have been explored. In addition, assays for controlled chemical alterations of the scaffold were developed that can be useful in applications demanding highly specific chemical composition of a durable biocompatible material (Biggs et al., 2010; Henrotin et al., 2001).

Reported here is an efficient washing method for the preparation of NayH2-yTinO2n+1 (xH2O) nanofiber-bioscaffold and the subsequent efficient cation exchange with Li+, Na+, K+, Ca2+, and Sr2+ using various chloride salts at concentrations similar to that at plasma and cellular conditions. The hydrothermal reactions between NayH2-yTinO2n+1 (xH2O) and chloride salts in an autoclave-liner proved to be the most efficient method for cation exchange. These results will allow future researchers to accomplish more precise alterations needed for studies demanding unique, stable, and precise control over the material and environment because of ion sensitivity of cellular proliferation (Henrotin et al., 2001). For example, osteoblast cells have been known to flourish in environments of elevated Sr2+ and Ca2+, yet Sr2+ is a known toxin at high levels; therefore an implant with controllable amounts of Sr2+ offers a great advantage over other implant materials in monitoring the growth of osteoblast cells (Health Effects, 2008; Henrotin et al., 2001; Marie & Hott, 1986).
Methods

Synthesis of Titanate Nanofiber-bioscaffolds
Prior to nanofiber-bioscaffold synthesis, a titanium foil 0.127 mm thick from Alfa Aesar was cut into approximately 1cm² piece and then sonicated in an acetone-filled, Teflon-lined container for 20 minutes. Subsequently, the container and substrate were rinsed with deionized water and blotted dry. The substrate was submerged in 1.0 M NaOH aqueous solution. The container was then placed in a digestion bomb, sealed tightly and heated at 240 °C for four hours. After three hours of cooling, the vessel was removed from the oven. The treated titanium substrate was removed from the alkali solution and rinsed with deionized water until the rinsed water had a pH around 7 as measured by a Mettler Toledo SevenEasy pH Meter. Ethanol washing showed a quicker pH reduction and efficient washing but was not used in order to keep the scaffold washing method constant throughout the whole study. Acetone, also tested, was found to be very inefficient as a washing solution. The substrate with titanate nanofiber-bioscaffold was then left to air-dry at 25°C.

Cation Exchange
In this study, several assays for cation exchange of the nanofiber-bioscaffold were explored for the first time using aqueous salt solutions at varying temperatures, time, and stimuli. The aqueous salt solutions used for cation exchange are solutions of Sr(OH)₂, Ca(OH)₂, SrCl₂, CaCl₂, KCl, LiCl, and NaCl with molarities up to 2 M. The temperature used in the different methods ranged from 25°C to 260°C. The time of each method varied depending on the method, ranging from 30 minutes to 30 days. The varying time between methods is because of the

Figure 1. Characterizations of starting Sodium titanate nanofiber-bioscaffold. (a) SEM image low magnification. (b) SEM high magnification. (c) XRD pattern with titanate nanofiber peaks labeled ; other peaks are titanium. (d) EDX plot of atomic ratio.
preliminary results, not shown explicitly here, that determined when the scaffold reached an
equilibrium, preventing further cation exchange. The various stimuli used to induce cation
exchange consisted of soaking, stirring, centrifugation, and heat.

Cation exchange was first attempted by soaking the titanate nanofiber-bioscaffold in
aqueous solutions in separate beakers from an overhead clip submerged in the middle of the
solution to create a point of reference and to determine if cation exchange occurred. The stirring
method also used this same set-up on top of a stir plate with a magnetic stirring from one hour to
21 days at a speed of up to 200 rpms. After the stirring experiment, the stirring method was
repeated with the addition of heat through an oil bath up to 75 °C. Centrifugation and vibration
consisted of 1, 2, & 5 minute intervals in a Precision Centrifuge Durafuge 100 followed by 30
second intervals of vibrations on a VWR mini-Vortexer.

Pressurized Teflon-lined autoclaves filled with aqueous solutions were used at
temperatures ranging from 80°C to 260°C over 1-24 hour time ranges. All samples were washed
by dilute salt solution at one-half the concentration used in each particular trial after the cation
exchange trial and before the characterizations.

Characterizations
A Rigaku Miniflex II Desktop X-Ray Diffractometer (XRD) was used to confirm the
titanate nanofiber-bioscaffold composition after first synthesis and by peak shifts after each
cation exchange attempt. Quorum Technologies SC7620 Mini Sputter Coater was used to apply a
gold coating on the scaffold to increase surface conductivity for further characterization using a
Tescan Vega-II Scanning Electron Microscope (SEM) and FEI Nova Nanolab 200 High
Resolution SEM & Energy Dispersive X-Ray (EDX). The SEM was used to confirm the
structural integrity before and after the cation exchange, along with EDX measurements for the
determination of the extent of cation exchange of each cation in the various methods.

Results & Discussion
Soaking the titanate nanofiber-bioscaffold in the various salt solutions served as a
benchmark for comparing the effect of different variables on the rate and effectiveness of
spontaneous cation exchange. Although ionic exchange of this scaffold has not been previously
attempted, our group has had success (in unpublished data) using similar nanofibers grown in
solution as opposed to on a substrate. To check whether the structural integrity of the scaffold is
maintained, SEM images were taken before and after cation exchange for every sample.
Strontium cation exchange was used as a representative treatment and shows the SEM images
before [Figure 1(a) & 1(b)] and after cation exchange [Figure 2(a) & 2(b)] while all other image
data are not shown due to similar imaging results that can be portrayed by these images.
Importantly, the structural integrity was not affected by the elevated pressure and temperature
method of cation exchange while other less ideal methods, such as high concentration or long
time exposure, allowed alterations in the form of precipitate on the surface that were determined
from SEM images (not shown here) of distorted scaffolds. The extent of cation exchange was
estimated by using EDX to calculate atomic ratios. As a representative example, note the
different EDX plots for Na⁺ [(Figure 1(d)] following exchange with Sr²⁺ [Figure 2(d)].
The titanate XRD pattern shown in Figure 1c provided independent confirmation of cation exchange via peak shift measurements. The titanate peak (110) at 24.30° 2θ was compared to an internal standard of pure titanium at 63.73° 2θ because the interplane (d-space) was on the same h, k, l plane with miller indices of (110). Using Bragg’s Law, the 2θ measurements observed are not reported; however, the resulting most optimal d-space alteration calculations are shown in Table 1 for exposure in 0.5M chloride solutions at 180 °C for eight hours. The difference in calculated d-space before and after various treatments proved to be a reliable indicator of cation exchange because the greatest d-space changes corresponded with the changes in atomic radii between the original cation and exchanged cation.

Figure 2. Characterizations of strontium titanate nanofiber-biosecaffold after 8 hours autoclaved in 0.50 M SrCl₂ at 180°C. (a) SEM image low magnification. (b) SEM high magnification. (c) XRD pattern peak shift indicating d-space change. (d) EDX plot of atomic ratio.
The XRD pattern of the titanate nanofiber-bioscaffold in Figure 1(c) corresponds to $\text{Na}_x\text{H}_{2-y}\text{Ti}_n\text{O}_{2n+1} (x\text{H}_2\text{O})$ as opposed to the previously suggested $\text{H}_2\text{Ti}_n\text{O}_{2n+1} (x\text{H}_2\text{O})$ (Dong et al., 2007). The authors’ interpretation differed from the results reported here most likely due to differing washing methods or other variables factoring between synthesis and XRD scans. If the originally proposed hydrogen titanate scaffolds were true, one would not expect the drastic d-spacing decrease by 0.03 nm seen when exposed to deionized water at room temperature; one would also not expect the reduction of sodium as reported by EDX results in Table 3. By using various anhydrous cations to exchange, the sodium titanate form is further confirmed because Li$^+$ and Ca$^{2+}$, with smaller atomic radii than sodium, both decreased the d-space while Na$^+$ caused little change and K$^+$ and Sr$^{2+}$, with larger atomic radii, increased the d-space. Table 1 depicts the results from eight hours of exposure to 0.5 M chloride salts at 180 ºC.

As a form of titanate with variable cation composition, the electrostatic interaction between the cations and anionic titanate strongly suggests an opportunity for cation exchange. The initial nanofiber-bioscaffold is purposed as being a mixture of sodium and hydrogen ions due to the NaOH aqueous solution used in synthesis and thus from the known ionic radii of hydrogen and sodium the original d-space is believed to be between 0.053 nm and 0.095 nm (Evans & Weber, 1999).

From a control study, soaking the nanofiber-bioscaffold in deionized water, a faster rate of proton exchange can be seen than other cation exchanges with larger ions, not shown explicitly in the figures but observed by XRD analysis. This finding is most likely due to the original basic surface of the scaffold and the spatial arrangement kinetically favoring the integration of smaller ions, such as protons, with the release of the larger sodium ions into solution (Evans & Weber, 1999). The scaffold control study soaked in deionized water began to reach equilibrium and near maximum exchange for sodium between days 3 and 14, apparent by the leveling off of the d-space change at a decrease of 0.03 nm. In contrast, the other samples soaked in chloride salt solutions continued cation exchange at a near linear rate with no inclination of becoming fully exchanged with the cations within 30 days. However, these samples did show gradual interplane modifications indicating a very slow transfer of ions between the solution and scaffold as can be seen by the SrCl$_2$ solution used in Figure 3(a).

Strontium cations were used to compare the different cation exchange assays because of the distinguishable ionic radius differences from sodium and hydrogen, in addition to inherent advantages of using strontium titanate scaffolds for bone implants. After the assays were refined, further cations (Li$^+$, Na$^+$, K$^+$, Ca$^{2+}$) were explored. Strontium chloride was the electrolyte examined as the chloride ions did not appear to affect the scaffold compared to the strontium and calcium hydroxides that presented topical surface modifications seen by SEM imaging. The XRD peak analysis of the strontium and calcium hydroxide treated samples resemble the unaltered crystalline structure, similar to that in Figure 1(a), suggesting the alterations were most

### Table 1

**XRD determined d-space shifts after cation exchange**

<table>
<thead>
<tr>
<th>Electrolyte (aq) solutions</th>
<th>KCl</th>
<th>SrCl$_2$</th>
<th>CaCl$_2$</th>
<th>NaCl</th>
<th>LiCl</th>
</tr>
</thead>
<tbody>
<tr>
<td>d-space shifts</td>
<td>0.0368 nm</td>
<td>0.0455 nm</td>
<td>-0.0153 nm</td>
<td>0.0122 nm</td>
<td>-0.01744 nm</td>
</tr>
</tbody>
</table>

The XRD pattern of the titanate nanofiber-bioscaffold in Figure 1(c) corresponds to $\text{Na}_x\text{H}_{2-y}\text{Ti}_n\text{O}_{2n+1} (x\text{H}_2\text{O})$ as opposed to the previously suggested $\text{H}_2\text{Ti}_n\text{O}_{2n+1} (x\text{H}_2\text{O})$ (Dong et al., 2007). The authors’ interpretation differed from the results reported here most likely due to differing washing methods or other variables factoring between synthesis and XRD scans. If the originally proposed hydrogen titanate scaffolds were true, one would not expect the drastic d-spacing decrease by 0.03 nm seen when exposed to deionized water at room temperature; one would also not expect the reduction of sodium as reported by EDX results in Table 3. By using various anhydrous cations to exchange, the sodium titanate form is further confirmed because Li$^+$ and Ca$^{2+}$, with smaller atomic radii than sodium, both decreased the d-space while Na$^+$ caused little change and K$^+$ and Sr$^{2+}$, with larger atomic radii, increased the d-space. Table 1 depicts the results from eight hours of exposure to 0.5 M chloride salts at 180 ºC.

As a form of titanate with variable cation composition, the electrostatic interaction between the cations and anionic titanate strongly suggests an opportunity for cation exchange. The initial nanofiber-bioscaffold is purposed as being a mixture of sodium and hydrogen ions due to the NaOH aqueous solution used in synthesis and thus from the known ionic radii of hydrogen and sodium the original d-space is believed to be between 0.053 nm and 0.095 nm (Evans & Weber, 1999).

From a control study, soaking the nanofiber-bioscaffold in deionized water, a faster rate of proton exchange can be seen than other cation exchanges with larger ions, not shown explicitly in the figures but observed by XRD analysis. This finding is most likely due to the original basic surface of the scaffold and the spatial arrangement kinetically favoring the integration of smaller ions, such as protons, with the release of the larger sodium ions into solution (Evans & Weber, 1999). The scaffold control study soaked in deionized water began to reach equilibrium and near maximum exchange for sodium between days 3 and 14, apparent by the leveling off of the d-space change at a decrease of 0.03 nm. In contrast, the other samples soaked in chloride salt solutions continued cation exchange at a near linear rate with no inclination of becoming fully exchanged with the cations within 30 days. However, these samples did show gradual interplane modifications indicating a very slow transfer of ions between the solution and scaffold as can be seen by the SrCl$_2$ solution used in Figure 3(a).

Strontium cations were used to compare the different cation exchange assays because of the distinguishable ionic radius differences from sodium and hydrogen, in addition to inherent advantages of using strontium titanate scaffolds for bone implants. After the assays were refined, further cations (Li$^+$, Na$^+$, K$^+$, Ca$^{2+}$) were explored. Strontium chloride was the electrolyte examined as the chloride ions did not appear to affect the scaffold compared to the strontium and calcium hydroxides that presented topical surface modifications seen by SEM imaging. The XRD peak analysis of the strontium and calcium hydroxide treated samples resemble the unaltered crystalline structure, similar to that in Figure 1(a), suggesting the alterations were most
likely due to a precipitate on the surface or further hydroxyl attachment to the scaffold. These alterations may be an indication of undesirable layered double hydroxides (Braterman, Xu, & Yarberry, 2004). The uses of hydroxides are thus less favorable than chloride salts for cation exchange.

On the basis of agitating the electrostatic interaction between the original hydrogen and sodium cations within the titanate, a stirring assay was first attempted at up to 21 days. XRD analysis of the titanate peak shifts showed similar d-space changes to the scaffolds soaked in electrolyte solution without any agitations while maintaining the scaffold structure after stirs at 200 rpm. EDX of the sample showed a reduction in atomic composition of sodium from 6.24% to 2.28% and an increase in strontium from 0% to 3.57%, indicating some cation exchange. Since mechanical agitation by stirring had little effect on the scaffold ionic composition, annealing was hypothesized to have a stronger effect in disrupting the chemical bond within the nanofibers. This was explored by elevating temperatures of the aqueous solutions in an oil bath in addition to the stirring. Figure 3(a) compares the results of stirring, stirring with heat, and the undisturbed soaking assays, depicting a more dramatic improvement with even mild heat.

In addition to the heating and stirring, another method was explored through centrifugation and vibration repeats, which have been effective for the solution-based nanofibers. However, for the solid state nanofiber-bioscaffold in this study, this method quickly revealed poor results due to uneven pressure upon the scaffold by the centrifuge that revealed inconsistent XRD patterns, indicating degraded parts of the scaffold and confirmed by SEM images.

Due to the faster cation exchange rate at elevated temperatures, an autoclave was used to induce cation exchange at varying temperature, time, and concentration of electrolyte. From trials using 0.10 M, 0.50 M, and 1.0 M SrCl₂ solutions for four hours at different oven temperatures, 180°C yielded the most optimal rate of cation exchange without any structural corruption. Below 180°C the change was not as drastic, while above 180°C the SEM pictures showed superficial deterioration or electrolyte precipitate due to the elevated pressure and heat.

Figure 3a & b. (a) Time study of cation exchange in 0.50 M SrCl₂ at varying temperatures, without the use of autoclave. (b) Time study of autoclave induced cation exchange using SrCl₂ at 180°C.

Inquiry: The University of Arkansas Undergraduate Research Journal, Vol. 15 [2013], Art. 1
As seen in Figure 3(b), depicting the d-space changes of autoclave experiments over various times, at least eight hours is needed for optimal cation exchange with the cations in solution while at times, over eight hours, an equilibrium appears to develop between the solution and the scaffold as no further d-space change could be observed when using strontium for the exchange cation. Table 2 showing data from 0.5 M SrCl₂ at 180°C supports this conclusion as the atomic ratio of the titanate nanofiber-bioscaffold showed no further increase in strontium after eight hours in an autoclave and a complete reduction of sodium.

Table 2  
*Atomic ratio of strontium exchanged titanate nanofiber-bioscaffold*

<table>
<thead>
<tr>
<th></th>
<th>4 hours</th>
<th>8 hours</th>
<th>12 hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxygen</td>
<td>0.6371</td>
<td>0.6415</td>
<td>0.6211</td>
</tr>
<tr>
<td>Titanium</td>
<td>0.2986</td>
<td>0.2901</td>
<td>0.3157</td>
</tr>
<tr>
<td>Sodium</td>
<td>0.02510</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Strontium</td>
<td>0.0392</td>
<td>0.0684</td>
<td>0.0632</td>
</tr>
</tbody>
</table>

The concentration dependence of strontium chloride had less effect on the cation exchange rates than the elevated temperature and pressure as well as time exposed to solutions shown by the mild disparities between the incorporated concentrations illustrated in Figure 3(b). However, at concentrations above 0.5 M, there were often unwanted discrepancies in the scaffold structure similar to the trial above 180°C and extra peaks emerging on XRD analysis around a 2θ of 32.1°, indicating a precipitate may be forming on top of the scaffold under elevated pressures. This accounts for the less predictable d-space change seen in Figure 3(b) at 1M SrCl₂. Further tests were completed with the same three varying concentrations of NaCl, KCl, CaCl₂, LiCl, and H₂O at four and eight hours to confirm these results. These samples were then examined through XRD to determine the optimal d-space change due to the respective cation exchange shown in Table 1. SEM analyses showed that the scaffold’s surfaces were constant before and after the autoclave treatment using salt solutions of 0.5 M, confirming the high-temperature stability of the scaffold during cation exchange under such conditions.

Finally, the samples were inspected via EDX in order to determine the exact amount of cationic exchange shown in Table 3, depicting the results of 0.5 M chloride aqueous solutions at 180 °C in autoclave liners for eight hours. The atomic composition of the scaffold fluctuated per sample and location quantified because top layers of nanofibers contain higher atomic composition of oxygen compared to those closer to the basal portion of the scaffold with higher concentrations of titanium. The results contain the percent atomic composition of gold and carbon that can be ignored, because they are present due to the sample preparation from coating and adhesion, respectively. The smaller atoms, such as lithium and hydrogen, could not be detected because the spectrometer’s beryllium window cannot detect any atom smaller than sodium (Piburn & Barron, 2012). With all of these varying factors taken into account, one can determine the extent of cation exchange by the absence of chloride, the decrease in sodium present, and the elevated presence of the cation correlating to the electrolyte solution used.
Table 3
Atomic ratio of cation exchanged titanate nanofiber-bioscaffold

<table>
<thead>
<tr>
<th></th>
<th>control</th>
<th>Na(^+)</th>
<th>Li(^+)</th>
<th>K(^+)</th>
<th>Ca(^{2+})</th>
<th>Sr(^{2+})</th>
<th>H(_2)O</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oxygen</td>
<td>0.6339</td>
<td>0.6344</td>
<td>0.6305</td>
<td>0.5884</td>
<td>0.6303</td>
<td>0.6415</td>
<td>0.6591</td>
</tr>
<tr>
<td>Titanium</td>
<td>0.2997</td>
<td>0.2994</td>
<td>0.3695</td>
<td>0.3507</td>
<td>0.3126</td>
<td>0.2901</td>
<td>0.3409</td>
</tr>
<tr>
<td>Sodium</td>
<td>0.0664</td>
<td>0.06620</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Hydrogen*</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Lithium*</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Potassium</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.06090</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Strontium</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0684</td>
<td>0</td>
</tr>
<tr>
<td>Calcium</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0671</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

*EDX cannot detect lithium and hydrogen atoms

The comparison of products obtained from treating scaffolds using SrCl\(_2\) in the autoclave for 4, 8, and 12 hours shows a distinct decrease in sodium present over time with optimum cation exchange predicted between four and eight hours; no further increase in strontium indicates that not only is all of the sodium removed from the d-space but also all of the hydrogen has more than likely been exchanged with strontium. These results show how to alter the chemical composition of the titanate nanofiber-bioscaffold while maintaining the structural integrity. Further studies may be able to use these methods to create materials that need distinct chemical composition or to apply these cation exchange techniques to similar nanostructures.
References


