QRBD

Quarterly Review of Business Disciplines

May 2018

Volume 5
Number 1

A Journal of International Academy of Business Disciplines
Sponsored by University of North Florida
ISSN 2334-0169 (print)
ISSN 2329-5163 (online)

Blank inside front cover

QRBD - QUARTERLY REVIEW OF BUSINESS DISCIPLINES

A JOURNAL OF INTERNATIONAL ACADEMY OF BUSINESS DISCIPLINES

FROM THE EDITORS

This issue of *Quarterly Review of Business Disciplines,* begins with Robert A. Page, Jr. and Louis K. Falk, exploring the ultra-competitive industry of education and the use of branding for business schools in order to distinguish one from another as applications decline, demographics change, and the quality of some non-traditional offerings are questioned. Wonseok Choi, Lawrence E. Zeff, and Mary A. Higby test the assumption that students' experiences with, and preferences for, increased/enhanced technology is factual. They compare virtual meetings and social media with face-to-face group member interaction.

Tamirat T. Abegaz and Bryson R. Payne present the National Security Agency/National Science Foundation GenCyber project to heighten awareness of and eventually graduate more university students with degrees in cyber security, computer education, et al. The research of Vincent J. Shea, Bobby E. Waldrup, Helen Xu, and Steven Williamson tests the customer profitability differences between complex and simple activity-based costing (ABC) systems. And our final paper by Kenneth R. Walsh and Sathiadev Mahesh, explores whether artificial intelligence can truly learn and whether business practice and the legal environment will allow a 'machine' to operate autonomously.

This is truly an interesting issue of QRBD.

Margaret A. Goralski, *Quinnipiac University,* Editor-in Chief
Charles A. Lubbers, *University of South Dakota,* Associate Editor

QRBD - Quarterly Review of Business Disciplines

A Journal of International Academy of Business Disciplines

The *Quarterly Review of Business Disciplines (QRBD)* is published by the International Academy of Business Disciplines quarterly in February, May, August, and November.

Manuscript Guidelines/Comments. *QRBD* is a blind peer-reviewed journal that provides publication of articles in all areas of business and the social sciences that affect business. The Journal welcomes the submission of manuscripts that meet the general criteria of significance and business excellence. Manuscripts should address real-world phenomena that highlight research that is interesting and different – innovative papers that begin or continue a line of inquiry that integrate across disciplines, as well as, those that are disciplinary. The Journal is interested in papers that are constructive in nature and suggest how established theories or understandings of issues in business can be positively revised, adapted, or extended through new perspectives and insights. Manuscripts that do not make a theoretical contribution to business studies or that have no relevance to the domain of business should not be sent to *QRBD.* Submissions to *QRBD* must follow the journal's Style Guide for Authors, including length, formatting, and references. Poorly structured or written papers will be returned to the authors promptly. Manuscript length is approximately 16 – 20 single-spaced pages. Acceptance rate is 25-28%.

Description. The *Quarterly Review of Business Disciplines* is a quarterly publication of the International Academy of Business Disciplines (IABD); a worldwide, non-profit organization established to foster and promote education in all of the functional and support disciplines of business. The objectives of *QRBD* and IABD are to stimulate learning and understanding and to exchange information, ideas, and research studies from around the world. The Academy provides a unique global forum for professionals and faculty in business, communications, and other social science fields to discuss and publish papers of common interest that overlap career, political, and national boundaries. *QRBD* and IABD create an environment to advance learning, teaching, and research, and the practice of all functional areas of business. *Quarterly Review of Business Disciplines* is published to promote cutting edge research in all of the functional areas of business.

Submission Procedure. An electronic version of the manuscript must be submitted in MS Word to the Editor-in-Chief, Dr. Margaret A. Goralski at Margaret.Goralski@Quinnipiac.edu. Upon completion of a review by expert scholars who serve on the *QRBD* Editorial Review Board, the first author will be informed of acceptance or rejection of the paper within a one to two-month timeframe from the submission date. If the paper is accepted, the first author will receive a formal letter of acceptance along with the *QRBD* Style Guide for Authors. IABD members and authors who participate in the IABD annual conference are given first priority as a matter of courtesy. For additional information, please visit www.iabd.org.

Subscription. The annual subscription price for QRBD is US$100 plus postage and handling. Single issue price is $35 per issue plus postage and handling.

QRBD - QUARTERLY REVIEW OF BUSINESS DISCIPLINES

A JOURNAL OF INTERNATIONAL ACADEMY OF BUSINESS DISCIPLINES

INTERNATIONAL ACADEMY OF BUSINESS DISCIPLINES

MISSION STATEMENT

The organization designated as the International Academy of Business Disciplines is a worldwide, non-profit organization, established to foster and promote education in all of the functional and support disciplines of business.

WWW.IABD.ORG

The Quarterly Review of Business Disciplines (QRBD) is listed in *Cabell's Directory of Publishing Opportunities.*

QRBD - Quarterly Review of Business Disciplines

A Journal of International Academy of Business Disciplines

Volume 5 Issue 1 May 2018

ISSN 2329-5163 (online)
ISSN 2334-0169 (print)

Contents

ARTICLES

This page has been intentionally left blank.

BUSINESS SCHOOLS, BRAND INNOVATION AND ARCHETYPAL TRADEOFFS

Robert A. Page, Jr., Southern Connecticut State University

Louis K. Falk, University of Texas Rio Grande Valley

ABSTRACT

In today's ultra-competitive education industry many business programs may be in danger of closing within the next 20 or 30 years. As universities face enrollment, funding, and non-traditional student support difficulties - the pressure increases. These troubles stem from the growth in the popularity of business degrees among employers, while applications decline, demographics change, and the quality of non-traditional offerings are questioned. The use of academic branding has emerged as a tool in this struggle for viability/sustainability. The ultimate goal of branding for a business school is to provide an impression leading to a positive reaction. Given the importance of adaptation and change, the authors propose that brand innovativeness is becoming an increasingly important criterion in academic marketing. This paper explores types of brand innovativeness by adapting the model developed by Beverland, Napoli and Farrelly to business schools. Further, Mark & Pearson's (2001) 12 Jungian archetypes can be added to the mix to give these innovative brands a face, a persona and marketing appeal. While brand archetypes are commonly utilized in other industries, the application of brand archetypes to business schools has just begun to be explored. From this perspective the potential tradeoffs between business school branding strategies and their attendant brand marketing initiatives become clear.

Keywords: Business Schools, Brand Innovation, Brand Archetypes, Marketing, Advertising

INTRODUCTION

Higher education has entered a period of discontinuous change, which may result in the closure of one third of business programs in higher education within the next 20 to 30 years (Christensen & Eyring, 2011). The academy is shrouded in a series of paradoxes. The popularity of business degrees among employers is increasing while student enrollments are decreasing. While pressures for a quality, marketable education rise, the numbers of non-traditional offerings, which have been challenged for not meeting those standards also grow. As an increasingly non-traditional student population needs more services and support, public and private funding necessary to meet those needs is being cut. This turbulence is forcing a "shakeout" phase in academia, where institutions are scrambling to adapt and grow in uncertain times, or otherwise risk falling victim to predicted closures (Christensen & Eyring, 2011; Page & Forbus, 2018). Popular media outlets cover rising tuition costs and often ask whether the diplomas are worth the investment.

Academic branding has emerged as an increasingly essential tool in this struggle for long-term institutional viability/sustainability. The academic branding process has evolved into a concerted public relations campaign to meet the expectations of a variety of external constituencies, ranging from customers to funding sources to government regulators. Expectations for desired outcomes of applying and being accepted into the program involve program inputs (what the student experiences in terms of what goes into the program, including the quality of their fellow students and faculty), program processes outcomes (what the student will experience in the program) and program completion outcomes (what they will experience as the result of having graduated from the program) (Heslop & Nadeau, 2010).

BUSINESS SCHOOL BRANDING

Responding to these trends, business schools have developed and maintained academic brands to improve their public image and attract support. Branding involves the aggressive application of external promotional strategies such as advertising and sales promotion (Balmer, Liao, & Wang, 2010; Ali-Choudhury, Bennett, & Savani, 2009; Chapleo, 2010; Hinds, Falgoust, Thomas, & Budden, 2011; Judson, Aurand, Gorchels, & Gordon, 2008; Pinar, Trapp, Girard, & Boyt, 2014). Academic brands involve developing a clear institutional "image," "identity" and "personality" in universities and business schools (Page & Forbus, 2018). These institutional traits are “a manifestation of the institution’s features that distinguish it from the others, reflect its capacity to satisfy students’ needs, engender trust in its ability to deliver a certain type and level of Higher Education (H.E.), and help potential recruits to make wise enrollment decisions” (Bennett & Ali-Choudhury, 2009, p. 94). Academic brands must clearly differentiate their institution in an overcrowded business school education marketplace (Balmer et al, 2010; Bisoux, 2003, 2015; Chapleo, 2010; Cova, Ford, & Salle, 2009; Opoku, Abratt, & Pitt, 2006; Stephenson & Yerger, 2014). Branding has become critical for the growth and success of many institutions without a distinctive reputation.

The branding goal is to provide a big-picture, strategic overview for the business school to focus on internal goals, structures, systems and staff in restructuring, recruiting, resource allocation and other critical decision-making processes that build the brand (Page & Forbus, 2018). Initiatives that strengthen the brand are prioritized over those which undermine or dilute those efforts (Judson, et. al., 2008). Some key elements of the branding ecosystem include: student experiences as the driving force of the university branding strategies, and academic services as the core value creation activities in delivering student learning experiences (Pinar et al., 2014). This is where a brand transitions from an "image" to an implemented, credible reality (Balmer & Liao, 2007).

Strong academic brands tend to systematically make the changes necessary to align their institution behind the selected brand (Bisoux, 2015; Eaton, 2008; Rowley & Sherman, 2001). From this perspective, key elements of the branding community or branding "ecosystems" must mutually support one another to function effectively. Properly aligned, the people, strategies and systems involved in an academic ecosystem work together and complement one another, creating a foundation for a solid reputation and credible brand (Balmer & Liao, 2007). When organizational leaders and strategies, systems and structures, staff and partners, and shared value systems conflict instead of complementing each other, the dysfunctions often created by misalignment undermine effectiveness. These severe misalignments undermine the authenticity of the brand and the prospects for long-term growth, resulting in lost credibility (Page & Forbus, 2018).

In contrast, the academic brand is strengthened by internal alignment over time coupled with an external customer relationship marketing focus to "treat the university, with all of its stakeholders, as a brand community, and to pursue policies and programs to strengthen the relationships that define the community" (McAlexander, Koenig, & Schouten, 2004, p. 61; Page & Forbus, 2018). Provided these strategic initiatives and innovations are consistent, external stakeholders and students will recognize this commitment and begin to regard the brand as credible and authentic.

Increasingly, the long-term viability of strategic initiatives and their attendant brands, is linked to its approach to innovation and change (Beverland, Napoli, & Farrelly, 2010). Beverland, Napoli and Farrelly (2010) developed a typology of brand innovation to analyze this dynamic. The typology categorizes brands by their relationship with strategic change and innovation through two sets of competing values. One branding continuum contrasts a focus on the degree of change, ranging from small increments to radical disruptions (Beverland, Napoli & Farrelly, 2010; in academia see Chapelo, 2010; Christensen & Eyring, 2011; Judson et. al., 2008). The other contrasts branding efforts responding to external trends and pressures versus those which tend to drive trends (Beverland, Napoli & Farrelly, 2010; in academia see Balmer & Liao, 2007; Pinar et al., 2014). These relationships lead to four branding categories, as illustrated in Figure 1:

Figure 1: Brand Innovation Categories

Disruptive Change	**CATEGORY LEADERS** (radical and market driven) Dominate market share with "bold product initiatives"	**PRODUCT LEADERS** (radical and driving markets) Innovation to leading-edge pioneers
Incremental Change	**FOLLOWER BRANDS** (incremental, market-driven) Efficient, fast to market branding and info. systems	**CRAFT-DESIGN BRANDS** (incremental and driving market) Maintain an "aura" of quality and authenticity
	Market-driven	***Driving Marketing***

(adapted from Beverland, Napoli & Farrelly, 2010)

This linkage increases the utility of branding as a strategic tool, particularly over time.

APPLYING ARCHETYPES TO BUSINESS SCHOOLS

Authentic brands tend to cluster around different externally validated "brand archetypes," which symbolically represent the university and its distinctive strengths. Archetypes create an organizational "persona" suitable for storytelling through linkages with mythological, animal or other attributes (Herskovitz & Crystal, 2010; Lloyd & Woodside, 2013). The most commonly accepted brand archetypes are developed from Jungian personality archetypes signifying sets of fundamental desires, first applied to business schools by Mark & Pearson (2001). While there is some consensus concerning the 12 archetypes identified, there is little concerning how to cluster and categorize them (Hartwell & Chen, 2012; Page & Forbus, 2018).

Given the importance of adaptation and change, the authors propose that brand innovativeness is becoming an increasingly important criteria in academic marketing and recruitment (Beverland, Napoli & Farrelly, 2010; Morse & Brooks, 2017). Further Mark & Pearson's (2001) 12 Jungian archetypes can be used to give these innovative brands a face, a persona and marketing appeal. This approach results in four broad categories distinguished by strategic innovative intent: industry leadership (product leaders), targeted branding (category leaders), networked branding (followers/imitators) and craft-design branding. Here the model must be adapted for academia. Following the brand innovativeness typology, as adapted to accommodate Mark & Pearson (2001), these 12 archetypes can be categorized by their placement on two continua: order versus change, and an individual, internal focus versus an external group focus, as summarized in Figure 2:

Figure 2: Traditional Brand Archetypes

Order

External Focus (group)	***PRODUCT LEADERSHIP BRANDS*** ***Sage*** Understanding truth by analysis ***Ruler*** Control and domination ***Magician*** Power to resolve problems	***TARGETED FOLLOWER BRANDS*** ***Innocent*** Safety through competence ***Everyman*** Belonging and quality offerings ***Jester*** Enjoy now, worry later	**Internal Focus (individual)**
	CATEGORY LEADER NETWORKED BRANDS ***Explorer*** Freedom through expansion ***Creator*** Innovation and creative vision ***Lover*** Intimacy and attractiveness	***CRAFTSMEN BRANDS*** ***Hero*** Individual mastery and strength ***Caregiver*** Support, service and aid ***Outlaw*** Revolution and break the mold	

Change

Note that there is little consensus as to which quadrant to place each archetype (Hartwell & Chen, 2012; Page & Forbus, 2018). The assignment of the archetypes within figure 2 is designed to best reflect the needs of higher education.

1. Archetypal Product Leadership Brands

Leadership brand archetypes are about using innovation to reinforce externally driven order and projecting the power to establish accepted areas of innovation. They establish the dominant design in established and emerging fields that provide a basis for external validation (Christensen & Eyring, 2011). For institutions projecting an authoritative, "shock and awe" presence, the following archetypes are for you, as summarized in Table 1:

Table 1: Archetypal Product Leadership Branding

BRAND ARCHETYPE	ACADEMIC MARKET NICHE
THE SAGE **Desire:** To find truth **Goal:** To understand all	To understand the world with superior curricula, networks, and analysis **Example:** Prestige program
THE MAGICIAN **Desire:** To know how the world works and overcome obstacles **Goal:** To realize dreams	Understanding of transforming the world through technology, sustainability, etc. **Example:** Emerging Field program
THE RULER **Desire:** Stability through control **Goal:** Dominate relationships	Industry leadership in targeted fields maintained at all costs **Example:** Standard-setting programs

The Sage

The Sages are usually prestigious research centers or elite private colleges with a reputation for research, learning communities and academic "best practices." They achieve sage status as they become business school program exemplars and higher education industry leaders with relatively rare innovative program offerings. These programs build brands around unique, distinctive features of university life that other institutions would find very difficult to replicate, grounded in world-class research. Students are willing to pay a premium for a program with a prestigious academic brand offering better status and world-class skill sets. Beyond leading edge research, sages tend to offer selectivity, impressive overall campus ambiance and a nurturing academic community (Joseph, Mullen, & Spake, 2012). Most Ivy League programs epitomize "the sage."

The Magician

Magicians use technologies as the "silver bullet" to satisfy both the need of cost control and competitive advantage through process innovation. Given that technology driven cost management is, in and of itself, a marketable skill set, programs developing this expertise can claim advantages in complex systems integration, service-oriented architecture, and supply chain management. These lean and even "green" initiatives build a brand around cost management and control. For example, interactive websites offer considerable cost savings by minimizing waste and transferring activities involving relatively expensive professional staff and on-ground facilities to relatively cheap and online technologies. While such "high tech" brands require considerable investments up front, they more than pay for themselves over time (Cater, Michel, & Varela, 2012; Dodd, 2014; McDougall, 2015). For example, the University of Maryland offers a highly ranked online MBA program with a specialization in supply chain management.

The Ruler

On rare occasions, business schools become so renowned for an area of expertise they dominate the market share and set quality standards to the point they can be considered "rulers." This is rare but does happen. For example, Harvard Business School case studies have become ubiquitous throughout management education (Levy, 2015).

2. Targeted Follower Academic Branding

Targeted brand archetypes are about low risk, proven innovative enhancements to improve well-established brand offerings. For institutions projecting a "tried and true" presence, or merely the facade of one, the following archetypes are viable options, as summarized in Table 2:

Table 2: Archetypal Academic Targeted Follower Branding

BRAND ARCHETYPE	ACADEMIC MARKET NICHE
THE EVERYMAN **Desire:** To connect with others **Goal:** To belong, fit in	Connects with students left out due to access and convenience **Example:** Online program
THE INNOCENT **Desire:** Peak experiences **Goal:** Personal fulfillment	Quality, effective programs offer reliability, consistency and safety **Example:** Standard 2 yr. program
THE JESTER **Desire:** Enjoy the moment **Goal:** Do not be too serious	Encourage customer spontaneity and fun - amusing, ironic, mischievous, playful **Example:** Superficial "Lite" program

The Everyman

The Everyman, or more accurately, Everyperson archetype is about access and convenience. The most successful access branding efforts in this arena involve online degree programs (Chapleo, 2010). Online degree programs comprise distance learning work and are steadily increasing in business schools (Cater, Michel, Varela, 2012; Gilardi & Guglielmetti, 2011; Lewin, 2013; Nelson, 2013). "Online Learning" lacks a common definition and is broadly defined, as content delivered exclusively online or hybrid - a mix of online and face-to-face (f2f) (Cater, Michel, Varela, 2012; Page, Williams, & McCarthy, 2009). One classification system bases the label on the proportion of web content, reserving the term "online" for programs with 80%+ online courses with little to no f2f interaction (Allen & Seaman, 2013). The popularity of this brand is Impressive - Devon Haynie notes: "With a total of 2.9 million graduate students in the U.S., 22% of them studied exclusively online. Among undergraduates, 11% pursued distance education exclusively" (2014, p. 2). Southern New Hampshire State University epitomizes a quality online business program where most students never set foot on campus.

The Innocent

Innocent archetypal branding involves an academic "field of dreams" mantra - if you build a quality program, students will come. Many business programs of state universities have embraced this archetype, offering generic business programs focused on an extensive core curriculum (Davis, 2014; Page, Williams, & McCarthy, 2009).

The Jester

Jesters develop the appearance of a tried and true program but undermine it when cost-cutting and profit-maximizing measures make their learning experiences superficial. While their branding efforts claim quality and depth, these claims are debatable. Also known as fools, tricksters, and practical jokers, jesters focus on giving their audience a good time while maintaining appearances, having fun versus rigorous culture, and not taking their degree programs too seriously (Batey, 2012). These types of programs are sometimes labeled "MBA-lite" in comparison with conventional two-year programs, regardless of the type of school (Petit, 2011). The compromised nature of "lite" programs is epitomized by conversions of general MBA programs into specialized MBA programs through little more than a series of relatively superficial changes - a set of "quick fixes" such as asserting that two elective classes in the field of study makes a student a subject matter expert (Rasche & Gilbert, 2015).

3. Category Leader Academic Branding

Network brand archetypes establish new product and service categories, as driven by market forces. This involves emergent fields, new alliances, and delving deeper into the subject as a distinctive advantage. For institutions projecting a more creative or real-world presence, the following archetypes are for you, as summarized in Table 3:

Table 3: Archetypal Academic Category Leader Network Branding

BRAND ARCHETYPE	ACADEMIC MARKET NICHE
THE EXPLORER **Desire:** Exploring the world with freedom **Goal:** To lead a fulfilling life	Exciting, risk-taking, seeking fulfillment linking campus and real-world discovery through a variety of media. **Example:** Applied program
THE CREATOR **Desire:** Make enduring beauty and value through innovation **Goal:** To give vision form	Emphasizes quality over quantity, customizes degrees to match areas of student interest **Example:** Specialized program
THE LOVER **Desire:** To attain intimacy and pleasure **Goal:** To have loving relationships	Alliances increase belonging, connection, and commitment to boundary spanning offers **Example:** Networked / Allied programs

The Explorer

Explorer brands reverse the old "teach students career skills in the classroom" mantra to its antithesis "use real jobs/internships to impart skills students need." Explorers find ways to teach students on the job, thus reducing costs to improve revenues. Employers call for more emphasis in these key areas: critical thinking, complex problem-solving, written and oral communication, and applied knowledge in real-world settings (Hart Research, 2013; Wilson, 2015). Thus, Explorer branding is based on the contention that the program structure and curriculum features process innovations that better prepare their students by offering typically neglected knowledge and skill sets needed for graduates to be employable. (Davies, Fidler, & Gorbis, 2011; Hart Research, 2013; Moskal, Ellis, & Keon, 2008).

Schools with strong job placement programs are increasingly attractive to potential MBAs - 21.87% in 2015, versus 18.8% in 2013 (GMAC, 2015). Real-world application also improve outcomes such as higher retention and greater influx of transfer students (Arum & Roska, 2011; Brownell & Swaner, 2009; Kuh, 2008). The degree to which a program is applied can be pictured as a continuum ranging from vicarious experiences to case studies to internships to immersion (students are embedded into a business and are exposed to real-world problems throughout the course of their degree) (Miethe, 2014). For example, Babson's entrepreneurship degrees are famous for their extensive experiential learning components where students are embedded in actual organizations and consulting projects for a significant portion of their degree training.

The Creator

Creators base their branding on their capacity for a specialized focus on process innovation in a particular field or industry instead of general degree (Davis 2014; Hanover Research 2013; McLeod 2013; Wilson, 2015). The premise is simple - if a student knows exactly where she or he wants to work, specialized programs will focus on that particular industry or market. Thus, taking better care of them than a more general program ever could. Students will emerge with focused skills and abilities, both conceptual and applied, making them subject matter experts in their specialization, qualifying them beyond a mere degree. Many schools find specialization useful for attracting students and cementing their brand identities. In recent years, specialization is becoming increasingly common even at high-profile institutions (Levy, 2008). Dan LeClair, Executive Officer and Chief Operating Officer of the Association to Advance Collegiate Schools of Business AACSB International summarizes: “Business schools are looking to differentiate themselves in an increasingly crowded market, so they’re ‘taking it up a notch’ by offering specialized master’s degrees" (quoted in Bisoux, 2015, p. 3). The number of specialized degrees has increased by 10%, and student interest in specialized programs has been steadily rising for the last five years, from 13 percent in 2009 to 20 percent in 2013 (Bisoux, 2015).

There are four types of degree programs with the creator brand: the specialized MBA, the specialized master’s, the customized degree and the dual MBA, which includes a separate specialized master’s degree (Bisoux, 2015; Hanover Research 2013; McLeod, 2013; Wilson, 2015). New developments in science and technology offer first-mover advantages to those programs that nimbly adapt traditional business degrees to a more customized approach. This includes establishing alliances with specific industries and incorporating innovation within specific niche markets to develop targeted curricula together. For example, Southern Connecticut State University recently announced an MBA program in public utilities management.

The Lover

Lovers embrace creative outsourcing and alliances - expanding and cost-cutting by moving away from highly skilled and expensive academic faculty and staff to a more open boundary model where external partners replace their internal counterparts. This type of network organization allows universities to share risk, conserve resources, increase in size and scope, and grow enrollments particularly on a global level (Bisoux, 2003, Davis, 2014).

> Outsourcing (a synonym for “privatizing) means that vendors outside the institution are exclusively handling services and functions that once were the domain of the institution’s staff. Progressively since the 1980s, campuses have outsourced bookstores, food services, print services, health services, information technology, building, planning, renovation, staff recruitment, and custodial services. They have also outsourced, to a lesser extent, security, housing, libraries, mail delivery, mental health services, the management of summer conferences, fundraising, admissions, retention planning, transportation, and alumni relations. (Milestone, 2010, p. 2)

Outsourcing is also increasingly common in the classroom:

- External contractors replace internal faculty (adjuncts)

- Professionally versus academically prepared instructors (University of Phoenix)
- Redefined coursework in terms of technology partners (SAP University Alliance)
- Cross-university curricular requirements with other institutions. University of North Carolina's (UNC) MBA program offers "cross-university" courses with George Washington, University of Southern California, University of Washington, and American University.
- Practical applications of lessons with business and community partners (Babson's 80 partners ships with local businesses in the Greater Boston area.)

4. Archetypal Academic Craftsman Branding

Craftsmen represent a brand archetype protecting a distinctive "march to the beat of a different drummer" pattern of process innovation. They master academic brand strategies that are failing for many institutions of higher learning and find creative ways to breathe new vitality into them. For institutions projecting and protecting an alternative, nontraditional presence, the following archetypes are for you, as summarized in Table 4:

Table 4: Archetypal Academic Craftsman Branding

BRAND ARCHETYPE **Mark and Pearson (2001)**	**ACADEMIC MARKET NICHE**
THE HERO **Desire:** To prove worth **Goal:** To master positively	Problem solve through acceleration with quality **Example:** Accelerated program
THE CAREGIVER **Desire:** To protect from harm **Goal:** To help others	Helpfulness, harmony and care for students creates a learning community **Example:** Affinity program
THE OUTLAW **Desire:** Revolution **Goal:** Defy rules / status quo	To challenge assumptions and restrictions as stifling and limiting **Example:** "Alternate" accreditation

The Hero

Hero branding tries to preserve the status quo through process innovations that accelerate the traditional program. Hero branding tries to preserve the status quo through process innovations that accelerating it. Program acceleration minimizes the time needed to complete degree requirements by combining undergraduate and graduate degrees, waiving requirements, granting credit for relevant life experience, and/or compressing class scheduling into all day or all week formats (Datar, Garvin, & Cullen, 2010, Page & Forbus, 2018). Given increased student flow, revenue growth offset the need for cost-driven staff reductions, making administrators heroes in the eyes of internal stakeholders. Done well, learning is accelerated without compromising quality.

The pressure to accelerate the speed of degree completion has become relentless on every degree level (Bogoslaw, 2012; Byrne, 2012; Hanover Research, 2013; Singh & Martin, 2004):

- Dual bachelor's and MBA degrees are routinely offered as a 5-year program, (4+1), and in the case of Quinnipiac University, as a 4-year program (3+1)
- MBA degrees are routinely reduced from the traditional two years to 18 months
- Specialized master's degrees target students who aren't willing and/or able to take the time necessary for a full MBA. They tend to favor a targeted educational experience at lower cost and faster completion (one year to 18 months) (Bisoux, 2015)
- Dual degrees (MBA and specialized master's) often are reduced from three to two years or less. Bentley offers dual degrees in one year.
- Doctor of Business Administration (DBA) programs are beginning to shrink their four-year programs into three years (South Florida, Creighton, South Alabama, Temple, etc.)

These programs are popular with students because of the additional costs associated with spending more time in school. Thus, perceived savings are decoupled from tuition and fee reductions. Further, staffing cuts can be minimized due to growing enrollments. Provided quality is not compromised, all major stakeholders benefit, and this brand can be implemented with minimal conflict (Hanover Research, 2013).

The Caregiver

Caregiver branding is a tried and true offering which is no longer in vogue. Aside form elite Liberal Arts and Sciences Colleges (Skidmore, Swarthmore, Amherst, etc.), academia is moving away from such intensive service in the name of efficiency and cost control. Programs which innovate continue to provide such services and support despite the supposedly unsustainable costs. These programs offer superior support for under-served student subpopulations. Beyond the famous Historically Black Colleges and Universities, Berea College provides a unique, friendly and supportive atmosphere for students from rural Appalachia (Irwin, 2014). The University of Texas Rio Grande Valley has been recognized by the White House Initiative of Educational Excellence for Hispanics as a bright spot in Hispanic education (UTRGV 2018). Caregiving begins with recruitment, by careful matching program features with student capabilities and/or needs - particularly targeting nontraditional students. Recruiting the right kind of students, with realistic expectations, and matching them with the right kind of program that will address those expectations is critical for student success and retention (Bisoux, 2015; Hanover Research, 2013). For example, academic leaders in 2012 note that only 11.2% of online learners are the type of disciplined, self-starters suited for online programs, down from 20% in 2007 (Allen & Seaman, 2013).

The Outlaw

Outlaw archetypes are dedicated to breaking all the rules and defying academic conventions. To such an extent that these programs end up compromising core business content - they run afoul of accrediting agencies such as the Association to Advance Collegiate Schools of Business (AACSB).

Without the external accreditation validation, business programs may have credibility issues, compromising graduates' employability and return-on-investment. Trump University attempted to substitute positive motivational programs for traditional curricula, and was subsequently condemned as a scam, a fraud and a diploma mill (Tuttle, 2016).

However, when the outlaws are right, and their non-traditional content and/or pedagogy proves itself, they do not remain outlaws for long. Their disruptive innovations slowly spread to become the new norm (Christensen & Eyring, 2011). Online education transcended outlaw status and is now ubiquitous, but with a twist. In distance learning, hybrid or blended approaches are outperforming those programs that are completely on-ground or online, resulting in a blurring of the definitions of online education. Students prize flexibility and the convenience of online classes, but not to the point of completely sacrificing face-to-face interaction, particularly in discussing complex issues (Moskal, Dziuban, & Hartman, 2013). The following programs demonstrate this trend:

- Regular face to face meeting with fellow students and faculty (UNC)
- Week-long orientation immersions (University of Indiana)
- Weekend or mid-semester retreats (Carnegie Mellon, University of Florida)
- Multi-day consulting projects onsite (Babson)
- Telecommuting for synchronous online classes (all students at the same time) or recorded lectures for asynchronous classes (UNC) (Byrne, 2013).

BRAND INNOVATIVENESS IMPLICATIONS

The focus of branding for business schools is ultimately to be attractive and "cool" (Warren & Campbell, 2014), while remaining authentic and credible (Ibarra, 2015). The premise of this paper is that brand innovativeness and its accompanying archetype are critical components of these cool and credible brands. Their effectiveness however, can be threatened and undermined.

Unfortunately, there is little consensus on what innovativeness specifically refers to. The root of this problem lies in the ambiguity of the concept. Given the value stakeholders place on innovativeness, one might assume that innovative brands and their attendant archetypes associated with change would be the most cool and attractive. However, this may not be the case. From an external stakeholder perspective, different stakeholders have different priorities. Faculty and accreditors tend to focus on order (knowledge-based product innovation), while students and employers value change (process innovations). Consequently, the definition of brand innovativeness becomes dependent on the type of data used in the study. A Google search of "the most innovative university" rankings reveals three major rating agencies using different criteria. Reuters and QS Stars rely on objective indicators gleaned from public records, such as published science and technology research (particularly industrial research, awards, patents, spinoffs, etc.). Usually these indicators have a technology bent (Ewalt, 2015; QS Stars, 2014). In contrast US News and World Report uses subjective opinion data from university administrators on a broad range of topics, including facilities, technology, curriculum, faculty, students and campus life (Morse & Brooks, 2017).

The lack of quality, objective data on academic process innovations skews ratings in terms of public perceptions, giving external market driven brands and their archetypes an advantage because they are most responsive to public feedback. Stakeholders clearly place innovativeness in a science and technology context. From this perspective, the academic process innovations characteristic of targeted follower and craftsman brands are a much harder sell as being attractive and cool, particularly when they are overlooked in university rankings. The complexity of different types of innovativeness as perceived by different stakeholders is clearly a topic for future research.

CONCLUSION

While the power of brand innovation and brand archetypes have been clearly established as a valuable marketing tool in industry, they are just beginning to be empirically researched, particularly in academia (Page & Forbus, 2018). Understanding their power, and the pitfalls that render them impotent is clearly worthwhile. In terms of Higher Education Branding, each of these described archetypes point to a different type of program. Brand dilution can threaten strategic effectiveness. For the savvy brand manager, "ideal" business programs often involve a variety of presentation formats, flexibility in systems and structures and adaptable curricular goals - when and where they need it (Moskal, Dziuban, & Hartman, 2013). Simultaneously pursuing a variety of brand archetypes is counterproductive if resources, faculty and facilities are spread too thin and undermine perceived quality.

Brands pursued simultaneously often compete with each other for attention and resource allocation. Brand images are undermined when they are under-resourced (Alajoutsijärvi, Juusola, & Siltaoja, 2014; Pinar et. al., 2011). In business schools, academic branding as jack-of-all-trades is often the master of none (Rasche & Gilbert, 2015). It is not uncommon, however, to find different archetypes being pursued for different programs within the same school - full-time versus part-time versus accelerated versus dual-degrees (Page & Forbus, 2018). Chances are one of these branding initiatives will be successfully implemented while the other falters. While some archetypes are complementary the differences can be enough to confuse the end consumer (student). In addition, the pursuit of any of the competing archetypes would also most likely lead to added uncertainty for the consumer. It is highly recommended that no two archetypes be used within the same domain during the same time period. Otherwise they are not effective.

If credibility is lost widespread cynicism may develop concerning the authenticity of the branding. This issue is now being confronted by many private equity for-profit schools (Jevons, 2006; Page & Forbus, 2018; Temple, 2006). The key lies in aggressiveness of brand marketing campaigns, which are often at best, overly enthusiastic, and at worst, fraudulent. Despite being tainted by their sub-standard student outcomes in student retention, degree completion, subsequent unemployment (Morse, 2015; Quinton, 2014) and being condemned by The National Bureau of Economic Research as "agile predators" (Deming, Goldin, & Katz, 2011) - that often target vulnerable low-income and disadvantaged students (Morse, 2015; Webley, 2012), for-profits continue to attract students. They give innovative branding a superficial aftertaste, and brand archetypes a bad name.

Many topics introduced by this paper call for future research. The most available research in this area is on MBA programs, neglecting undergraduate curricula. Further, while international comparisons with European business schools are beyond the scope of this paper, they present another fertile future research stream. Methodologically, most empirical research in this area is in organizational trait theory, which typical can only distinguish about a half dozen traits.Consequently, there is no consensus on how to cluster archetypes, nor to link them with different types of competitive strategies. Nor is there any guarantee, based on the dynamic environment of higher education that this analysis will remain valid in the future. Limitations notwithstanding, the strategic and marketing advantages of archetypal innovation are undeniable and merit consideration.

REFERENCES

Alajoutsijärvi, K., Juusola, K., & Siltaoja, M. (2014). The legitimacy paradox of business schools: losing by gaining? *Academy of Management Learning & Education*, *14*(2), 277-291.

Ali-Choudhury, R., Bennett, R., & Savani, S. (2009). University marketing directors' views on the components of a university brand. *International Review on Public and Nonprofit Marketing*, *6*(1), 11-33.

Allen, I. E., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Arum, R., & Roksa, J. (2011). *Academically Adrift: Limited Learning on College Campuses*. Chicago: University of Chicago Press.

Balmer, J. M., & Liao, M. N. (2007). Student corporate brand identification: an exploratory case study. *Corporate Communications: An International Journal*, *12*(4), 356-375.

Balmer, J. M. T., Liao, M. N., & Wang W. Y. (2010). Corporate brand identification and corporate brand management: how top business schools do it. *Journal of General Management*, *35*(4), 77-102.

Batey, M. (2012). *Brand meaning*. Psychology Press.

Bennett, R., & Ali-Choudhury, R. (2009). Prospective students' perceptions of university brands: An empirical study [electronic version]. *Journal of Marketing For Higher Education,* *19*(1), 85-107.

Beverland, M. B., Napoli, J., & Farrelly, F. (2010). Can all brands innovate in the same way? A typology of brand position and innovation effort. *Journal of Product Innovation Management, 27*, 33–48.

Bisoux, T. (2015, September 4). Mastering the Mix. *Biz Ed,* Retrieved from http://bized.aacsb.edu/articles/2015/09/focus-on-specialized-programs-mastering-the-mix

Bisoux, T. (2003). The zen of b-school branding. *Biz Ed*, *3*(1), 24-30.

Bogoslaw, D. (2012, July 10). One-year MBA programs gain more traction. Retrieved from: http://poetsandquants.com/2012/07/10/one-year-mba-programs-gain-more-traction/3/2 "2012 QS TopMBA.com Applicant Survey."

Brownell, J. E., & Swaner, L. E. (2009). *Five high impact practices*. Washington, DC: Association of American Colleges and Universities.

Byrne, J. (2012, February 6). Kellogg to shrink two ‑ year MBA program. *Poets and Quants*, Retrieved from http://poetsandquants.com/2012/02/06/kellogg-to-shrink-two-year-mba-program/

Byrne, J. A. (2013, May 29) The online MBA comes of age. *Fortune*. Retrieved from http://fortune.com/2013/05/29/the-online-mba-comes-of-age/

Cater, J. J., Michel, N., & Varela, O. E. (2012). Challenges of online learning in management education: An empirical study. *Journal of Applied Management and Entrepreneurship*, *17*(4), 76-96.

Chapleo, C. (2010). What defines successful university brands? *The International Journal for Public Sector Management*, *23*(2), 169–183.

Christensen, C. M., & Eyring, H. J. (2011). *The innovative university: Changing the DNA of higher education from the inside out*. John Wiley & Sons.

Cohen, P. (2015, October 12). For-profit colleges accused of fraud still receive U.S. funds. *New York Times*. Retrieved from http://www.nytimes.com/2015/10/13/business/for-profit-colleges-accused-of-fraud-still-receive-us-funds.html? r=0

Cova, B., Ford, D., & Salle, R. (2009). Academic brands and their impact on scientific endeavour: The case of business market research and researchers. *Industrial Marketing Management*, 386, 570-576.

Datar, S. M., Garvin, D. A., & Cullen, P. G. (2011). Rethinking the MBA: Business education at a crossroads. *Journal of Management Development*, *30*(5), 451-462.

Davies, A., Fidler, D., & Gorbis, M. (2011). *Future work skills 2020*. San Francisco: Institute for the Future for the University of Phoenix Research Institute.

Davis, E. (2014, July 18). Changing program composition. AACSB blog. Retrieved from http://aacsbblogs.typepad.com/dataandresearch/2014/07/ changing-program-composition-.html

Deming, D. J., Goldin, C., & Katz, L. F. (2011). *The for-profit postsecondary school sector: nimble critters or agile predators?* (No. w17710). National Bureau of Economic Research.

Dodd, D. (2014, October 29). Three ways institutions are using technology to improve efficiency and effectiveness. *Evolution*. Retrieved from: http://evolllution.com/opinions/ways-institutions-technology-improve-efficiency-effectiveness/

Eaton, J. S. (2008). Attending to student learning. *Change*, *40*(4), 22-27.

Ewalt, D. (2015 September 15). The world's most innovative universities. *Reuters*. Retrieved from https://www.reuters.com/article/idUSL1N11K16Q20150915

Ford, R. C., & Bowen D. B. (2008). "A Service=Dominant Logic for Management Education." *Academy of Management Learning and Education, 7*(2), 224-243.

Gilardi, S., & Guglielmetti, C. (2011). University life of non-traditional students: Engagement styles and impact on attrition. *The Journal of Higher Education*, *82*(1), 33-53.

Graduate Management Admissions Council (GMAC) (2015). *mba.com Prospective Students Survey, 2015 survey report*. Retrieved from http://www.gmac.com/market-intelligence-and-research/research-library/admissions-and-application-trends/2015-mbacom-prospective-students-survey-report.aspx

Hanover Research (2013). Competitive MBA programming - Trends and best practices. Retrieved from http://www.hanoverresearch.com/insights/competitive-mba-programming/?i=higher-education

Hart Research (2013). *It Takes More than a Major: Employer Priorities for College Learning and Student Success*. Washington, DC: Association of American Colleges and Universities.

Hartwell, M., & Chen, J. C. (2012). *Archetypes in branding: A toolkit for creatives and strategists*. How Books.

Haynie, D. (2014, June 12). New government data sheds light on online learners. *U.S. News & World Report.*

Herskovitz, S., & Crystal, M. (2010). The essential brand persona: storytelling and branding. *Journal of Business Strategy 31*(3), 21-28.

Heslop, L. A., & Nadeau, J. (2010). Branding MBA programs: the use of target market desired outcomes for effective brand positioning. *Journal of Marketing For Higher Educatio*n, *20*(1), 85-117.

Hinds, T., Falgoust, D., Thomas Jr, K., & Budden, M. C. (2011). Examining the perceptions of brand images regarding competing MBA programs. *American Journal of Business Education*, *3*(12), 7-18.

Ibarra, H. (2015, Jan - Feb). The authenticity paradox. *Harvard Business Review*, 54-59.

Irwin, N. (2014). Why colleges with a distinct focus have a hidden advantage. *New York Times*. Retrieved from http://www.nytimes.com/2014/09/04/upshot/why-colleges-with-a-distinct-focus-have-a-hidden-advantage.html?emc=eta1&_r=0&abt=0002&abg=1

Jevons, C. (2006). Universities: a prime example of branding going wrong. *Journal of Product & Brand Management*, *15*(7), 466-467.

Joseph, M., Mullen, E. W., & Spake, D. (2012). University branding: Understanding students' choice of an educational institution. *Journal of Brand Management*, *20*(1), 1-12.

Judson, K. M., Aurand, T. W., Gorchels, L., & Gordon, G. L. (2008). Building a university brand from within: university administrators' perspectives of internal branding. *Services Marketing Quarterly*, *30*(1), 54-68.

Kuh, G. D. (2008). *High-impact educational practices: What they are, who has access to them, and why they matter.* Washington, DC: AAC&U.

Levy, F. (2008, July 25). The spread of specialized MBA programs. *BusinessWeek*. Retrieved from http://www.bloomberg.com/bw/stories/2008-07-25/the-spread-of-specialized-mba-programsbusinessweek-business-news-stock-market-and-financial-advice

Levy, F. (2015, April 29). Harvard Business School has the market cornered on case studies. *Bloomberg Businessweek*. Retrieved from https://www.bloomberg.com/news/articles/2015-04-09/harvard-s-case-study-monopoly

Lewin, T. (2013, August 17). Master's degree is new frontier of study online. *New York Times*. http://www.nytimes.com/2013/08/18/education/masters-degree-is-new-frontier-of-study-online.html

Lloyd, S., & Woodside, A. G. (2013). Animals, archetypes, and advertising (A3): The theory and the practice of customer brand symbolism. *Journal of Marketing Management*, *29*(1-2), 5-25.

Mark, M., & Pearson, C. S. (2001). *The hero and the outlaw: Building extraordinary brands through the power of archetypes*. McGraw Hill Professional.

Martin, A., & Lehren, A. W. (2012). A generation hobbled by the soaring cost of college. *The New York Times*, Retrieved from http://www.nytimes.com/2012/05/13/business/student-loans-weighing-down-a-generation-with-heavy-debt.html?pagewanted=all

McAlexander, J. H., Koenig , H. F., & Schouten, J. W. (2004). Building university brand community: The long-term impact of shared experiences. *Journal of Marketing for Higher Education*, *14*(2), 61–79.

McDougall D. (2015). Five ways for colleges and universities to get smart about energy. Retrieved from http://www.universitybusiness.com/article/five-ways-colleges-and-universities-get-smart-about-energy

McLeod, H. (2013). Significant strides in specialized master's enrollment. Retrieved from http://aacsbblogs.typepad.com/dataandresearch/2013/12/significant-strides-in-specialized-masters-enrollment-.html

Miethe, B. (2014, December 19). Business students thrive in marketplace. *Campus News*. Lee University.

Milestone, D. (2010). Outsourcing services in higher education: Consider the campus climate. *The Bulletin*, *78*(2), 1-5.

Morse, S. N. (2015). For-Profit Schools: A History of Abuse and the Need for Reform. *Brigham Young University Education and Law Journal, 2*, 585-595.

Morse, R., & Brooks, E. (2017 September 11). Most innovative schools methodology. *U.S. News and World Report.* Retrieved from https://www.usnews.com/education/best-colleges/articles/most-innovative-schools-methodology

Moskal, P., Dziuban, C., & Hartman, J. (2013). Blended learning: A dangerous idea? *The Internet and Higher Education*, *18*, 15-23.

Moskal, P., Ellis, T., & Keon, T. (2008). Summary of assessment in higher education and the management of student-learning data. *Academy of Management Learning & Education*, *7*(2), 269-278.

Nelson, C. (2013, January 24). Growth in accessibility of AACSB-accredited online degrees. Retrieved from http://aacsbblogs.typepad.com/dataandresearch/2013/01/growth-in-accessibility-of-aacsb-accredited-online-degrees.html

Opoku, R., Abratt, R., & Pitt, L. (2006). Communicating brand personality: are the websites doing the talking for the top South African business schools? *Journal of Brand Management*, *14*(1), 20-39.

Page, R. A. Jr., & Forbus, R. (2018). Business schools, MBAs, and brand archetypes. *Journal of Educational Advancement and Marketing. Forthcoming*

Page, R. A. Jr., Williams, K., & McCarthy, M. (2009). The on-line MBA: A competing values matrix. In R.A. Oglesby and H.P. LeBlanc (Eds.) *Business Research Yearbook*, *16* (pp. 416-424). Beltsville, MD: International Graphics.

Petit, F. (2011). Rethinking executive MBA programs. *MIT Sloan Management Review*, *53*(1), 19-25.

Pinar, M., Trapp, P., Girard, T., & Boyt, T. E. (2011). Utilizing the brand ecosystem framework in designing branding strategies for higher education. *International Journal of Educational Management*, *25*(7), 724-739.

Pinar, M., Trapp, P., Girard, T., & Boyt, T. E. (2014). University brand equity: an empirical investigation of its dimensions. *International Journal of Educational Management*, *28*(6), 616-634.

QS Stars (2014, July 31). Rating universities on innovation. Retrieved from https://www.topuniversities.com/qs-stars/rating-universities-innovation-qs-stars

Quinton, S. (2014). Will a for-profit degree get you a job? Even if it does, the benefit will likely be diminished by high student debt. *National Journal*. Retrieved from http://www.nationaljournal.com/next-america/education/will-a-for-profit-degree-get-you-a-job-20140324

Rasche, A., & Gilbert, D. U. (2015). Decoupling responsible management education why business schools may not walk their talk. *Journal of Management Inquiry*, *24*(3), 239-252.

Rowley, D. J., & Sherman, H. (2001) *From strategy to change: Implementing the plan in higher education.* San Francisco: Jossey-Bass.

Singh, P., & Martin, L. R. (2004). Accelerated degree programs: Assessing student attitudes and intentions. *Journal of Education for Business*, *79*(5), 299-305.

Stephenson, A. L., & Yerger, D. B. (2014). Does brand identification transform alumni into university advocates? *International Review on Public and Nonprofit Marketing*, *11*(3), 243-262.

Temple, P. (2006). Branding higher education: illusion or reality? *Perspective*, *10*(1), 15-19.

Tuttle, I. (2016, February 26). Yes, Trump University was a massive scam. *National Review*. Retrieved from http://www.nationalreview.com/corner/432010/trump-university-was-massive-scam.

U.S. Department of the Treasury with the Department of Education (2012). *The economics of higher education*. Retrieved from http://www.treasury.gov/connect/blog/Documents/20121212_Economics%20of%20Higher%20Ed_vFINAL.pdf

Warren, C., & Campbell, M. C. (2014). What makes things cool? The role of bounded autonomy. *Journal of Consumer Research, 41*, 1-22.

Webley, K. (2012). Can an online degree really help you get a job? *Time*. Retrieved from http://nation.time.com/2012/10/18/can-an-online-degree-really-help-you-get-a-job/

Wilson, A. M. (2015, February 23). Five MBA trends for 2015. *The Independent*. Retrieved from http://www.independent.co.uk/student/postgraduate/mbas-guide/five-mba-trends-for-2015- 10060102.html

DESIGNING A GROUP ASSIGNMENT IN A DIGITAL ERA

Wonseok Choi, University of Detroit Mercy

Lawrence E. Zeff, University of Detroit Mercy

Mary A. Higby, University of Detroit Mercy

ABSTRACT

Millennials, also called "digital natives," are attracted to new communication technologies and may incorporate them into their daily activities. But do they? Much research considers how faculty can effectively use the newest education technology to enhance coursework materials. Faculty expectations are high regarding student use of these new technologies to complete and comprehend course assignments and materials. Using results from an exploratory study involving two focus group interviews, a 66-item questionnaire was developed to test these assumptions and begin to determine what student experiences suggest regarding their use and impact on group project performance. We collected responses from 82 students at a Midwestern United States university. Our research question is: do students' experiences with and preferences for increased/enhanced technology in the completion of group assignments support and encourage an increased emphasis on technology-based interaction by faculty?

We collected data regarding student experiences with both virtual meetings and social media and compared them with face-to-face (FTF) group member interactions in the completion of class assignments involving group projects. Direct student experience demonstrates that superior performance and satisfaction result from FTF meetings rather than technology-based interaction. Students also prefer face-to-face meetings over virtual meetings. This result is true whether comparing FTF with social media or virtual meetings. Surprisingly, students find neither virtual meetings nor social media to have significantly more positive impact on groups than face-to-face interaction along both performance and process dimensions. Recommendations for faculty are provided and suggestions for future research are included.

Keywords: Millennials, Group work, Face-to-face interaction, Technology-based interaction

INTRODUCTION

Classrooms have changed dramatically over time. We have gained knowledge about learning styles and how people learn. There is a continuous flow of new educational technology available to faculty members. Small group activities and assignments are used to break down the size of the class and more closely replicate a work environment many students already face and others will soon be entering. Group experiences provide opportunities for students to practice interpersonal and leadership skills, both of which transfer directly to the job. They also increase participation and student involvement, which have direct relationships to the learning process. Moreover, group projects allow students to try out new ideas on and to gain feedback from peers to improve contributions to project outcomes. In addition, students often experience accountability and group issues such as social loafing.

Today's students have been called "digital natives" (see, e.g., Roberts & Kidd, 2017). Faculty expect these students to prefer technology in all of its forms when presenting course materials in the classroom. Faculty use these new educational technologies to better communicate with and involve current students. These expectations often influence faculty perceptions regarding student preferences in completing class assignments. Research that studies the availability and use of technology usually takes the perspective of faculty. Changing the perspective to determine how students actually use technology requires different research.

The present study investigates student experiences in their use of and preferences for technology in the completion of group assignments. Our research question is: do students' experiences in the completion of group assignments support and encourage an increased emphasis on educational technology by faculty? Specifically, we compare student experiences when using face-to-face or technology-based meetings in completing group projects.

LITERATURE REVIEW

Two basic issues are included in this literature review, namely, possible discrepancies between the expectations by faculty members and students' actual use of technology in the completion of classroom assignments, and the outcomes produced by students when they actually use face-to-face interaction or digital technology while working in groups. In the past few decades, many instructors have moved away from a sole diet of traditional lectures to a class in which students are active participants in the learning process. Small-group work is among the most often-used approaches to get students engaged in the classroom (Davidson, Major, & Michaelsen, 2014). Group work has developed into an increasingly valuable component of higher education (Cheng & Warren, 2000) providing students with pseudo-workplace projects that allow the opportunity to gain valuable teamwork experience and has the potential to enhance abilities such as communication and group skills (McCorkle et al., 1999).

Group work is defined as students working together in a small enough group so that everyone can participate on a clearly assigned learning task (Cohen & Lotan, 2014). Of a larger scope than individual assignments, group projects allow for in-depth work that provides a more realistic work experience than typical coursework (McCorkle et al., 1999). However, many benefits of group work are possible only when students communicate and work collaboratively (Gordon & Connor, 2001).

Digital technologies are now an important component of the university student learning experience. These students have been known to use a lot of digital communication tools (Roberts & Kidd, 2017). Some researchers, however, suggest that digital technologies are clearly not transforming the nature of university teaching and learning, or even substantially disrupting the student experience (Henderson, Selwyn, Finger, & Aston, 2015). For instance, while today’s college students are immersed and fluent in social media, instructors expect their students to be proficient in and prefer to use any form of digital communication tools, particularly course-learning technology. Hence, they see no need to provide training in the use of these tools or group interaction skills (Kirschner, & De Bruyckere, 2017). So, while 65% of the instructors thought that students were tech savvy, only 42% of the student respondents felt that instructors provided students with adequate training and support in the use of instructional technology (Buzzard, Crittenden, Crittenden, & McCarty, 2011). Perhaps the disconnect here between faculty and student perceptions relates not to digital skills. Rather, it relates to social media or educational technology preferences of each. Buzzard and colleagues (2011) found that while faculty members prefer the use of advanced level educational technology, students prefer more traditional instructional methods for effective engagement and learning.

Many of the current discussions and debates over social media are also unclear as to what aspects of social media use actually relate to education, learning and knowledge (Selwyn, 2012). One study of United Kingdom students' use of Facebook suggested that the vast majority (around 95%) of students' interactions were completely unrelated to their university studies (Selwyn, 2009). The majority of social media uses are perhaps most accurately described as constituting 'the ordinary stuff of life' (Shirky, 2008), rather than creative, communal and convivial activities. Recent studies suggest university students use social media at a surprisingly low level of sophistication (Gunter, Rowlands, & Nicholas, 2009; McLoughlin & Lee, 2010; Waycott, Bennett, Kennedy, Dalgarno, & Gray, 2010). At best, many students' engagement can be called a "low bandwidth exchange" of information and knowledge (Crook, 2008; Selwyn, 2009).

What if digital technologies do not actually help students and instead prevent them from attaining the maximum learning potential provided by group interaction during coursework? While much of the earlier literature on digital technologies was optimistic about their potential to enhance students' learning (for a review, see Selwyn, 2016), recent studies have been more cautious in this regard (see, for example, Chu, 2014). Kvavik (2005) found that many of the students most skilled in the use of technology had mixed feelings about technology in the classroom. Despite their potential benefits, students' uses of digital technologies are not the most expansive ways that they could be used (Henderson et al., 2015). Hence, to allow digital communication to have a positive impact on the learning process and to improve learning through the use of group projects, research should account for the role digital technology actually plays in students' group work completion (see, for example, McKnight et al., 2016).

Discussing success levels by considering how group projects are completed, the literature provides substantial information. For instance, building trust within a team is recognized as a key ingredient for team success (e.g., Davis, Schoorman, Mayer, & Tan, 2000; De Jong & Elfring, 2010). Breuer, Hüffmeier, and Hertel (2016) suggest that trust facilitates specific risk-taking behaviors such as reducing defensive control, open discussion of conflicts and mistakes, mutual feedback, and sharing of confidential information, which in turn should lead to more efficient coordination of team members' resources (time, effort, knowledge, etc.).

Social presence theory (Short, Williams, & Christie, 1976) explains how FTF interactions provide more complete communication since both verbal and non-verbal cues are part of the social exchange process. Advances in information technology have created new challenges for team processes (Cramton, 2001; Driskell, Radtke, & Salas, 2003; Rains, 2005; Thompson & Coovert, 2003). Digital communication can limit direct personal observations that allow members to perform effective cognitive trust assessment (Robert, Denis, & Hung, 2009). Awareness of who is responsible for specific outcomes (Cui, Lockee, & Meng, 2013) and issues of accountability (Driskell et al., 2003; Reio & Crim, 2006) further reduce overall performance, while increasing frustration and dissatisfaction, and lowering participation. For instance, team members that exclusively rely on technology-based interaction will have no opportunity to see firsthand the amount of effort others are expending or participate in the informal interactions with team members. It has been found when social context cues are missing, increased depersonalization, lower cohesiveness, and less social conformity often result (Lu, Fan, & Zhou, 2016; Szeto & Cheng, 2013).

Media richness theory explains how face-to-face interaction is so rich since it enables not only the spoken language and other verbal cues, but also body language (Kennedy, Vozdolska, & McComb, 2010; Lantz, 2001). This gives the communicating parties a better basis for understanding each other compared to purely technology-based interaction (Lantz, 2001). In this regard, much of the literature concludes that FTF interaction at the beginning of a group project enhances the level of trust. Hambley, O'Neill, and Kline (2007), Horwitz and Horwitz (2007) and Lantz (2001), for example, advise project teams to have at least an initial FTF meeting before following up with virtual team interactions. Kennedy and colleagues (2010) found in their behavioral simulation study that mixed-media teams (i.e., first as FTF and second as digital communication) had improved participative decision making over only digital communication teams. Both high and low media richness levels are effective when matched with appropriate tasks. For example, media with lower richness are effective when used with more routine tasks and richer media are better matched with nonroutine, complex and ambiguous tasks (Denstadli, Julsrud, & Hjorthol, 2012).

RESEARCH METHODOLOGY

Questionnaire Development

Two focus group interviews were conducted in an exploratory study to investigate student experiences and preferences in the use of technology-based communication during the completion of group projects (Choi, Zeff, & Higby, 2017). Both interviews were transcribed and content analyzed. The researchers then discussed and considered how the experiences could be translated into questionnaire items. The 66 survey items were the result of this discussion. Issues from these experiences dealing with outcomes, processes and preferences were applied to each of three collaboration methods, namely, FTF, virtual and social media for comparative purposes.

Our focus groups found that traditional group interaction occurs in face-to-face meetings and involves two basic types of activities, namely, on-task (or the more formal activities occurring within a group) and off-task (or informal and more social types of activities). We wanted to compare experiences with face-to-face interaction and technology-based communication. To more directly deal with both types of activities, we split all digital communication forms into either virtual (more formal) or social media (more social or informal) types. A more detailed breakdown is described in the results section below. A four-point Likert scale ranging from 1, "Strongly Disagree" to 4, "Strongly Agree" was used for each question. A forced-choice questionnaire provides a more reasoned response (Smyth, Dillman, Christian, & Stern, 2006) and lessens the compromise effect, decreasing the relative proportion of an average response (Dhar & Simonson, 2003). All items were pre-tested to make sure they accurately reflected the comments from the two focus groups. A complete questionnaire can be found in the Appendix.

Sample

Data were collected during the 2016-17 academic year at an urban Midwestern United States university school of business with mainly commuter students. Students from four courses (2 undergraduate, 2 graduate) were invited by their instructors to complete a questionnaire investigating their experiences with group projects. After the instructor briefly introduced the purpose of the survey, questionnaires were distributed to students, 82 of whom volunteered and filled out the questionnaire. Demographic information indicates: 80.5% of these students are between the ages of 17 and 26; 58% are female; 60.7% are graduate students; and, 40.5% of the students have more than 30 hours out-of-class commitments per week. All but one student (98.8%) has access to and uses smartphones whereas only 56.1% report that they have access to and use tablets. Every respondent indicates he/she has access to and uses a computer.

Data Analysis

Our research question is: do students' experiences in the completion of group assignments support and encourage an increased emphasis on educational technology by faculty? To help answer our research question, we tested whether responses were significantly above or below the 2.5 neutral point of the Likert 4-point scale using a one-sample t-test. We also applied ANOVA tests to see if the students' responses were different between groups within the demographic dimensions of gender, graduate/undergraduate level and number of hours per week out-of-class commitments (e.g., number of hours working per week). We used SPSS (version 22.0, 2013) to analyze questionnaire results.

RESULTS

Results from the questionnaire completed by 82 students are presented below. They are grouped into the three basic sections covered by the questions, namely, FTF, virtual and social media distinctions. For each section, we present the results responding to questions that relate to: outcomes (project/grade results, satisfaction and efficiency); process (including trust, task-orientation, information exchange/effective communication, boredom/division of work/asking for help); and, overall preference for method of group interaction for assignment completion. Most of the student responses were found to be significantly different from the 2.5 neutral point (unless otherwise noted). We have included the t-test statistics for each item in the table below the paragraph presenting the data, while any data comparing item-to-item responses are included parenthetically in the body of the paper. In the tables that follow, a t-test with a negative value indicates disagreement with the item and a positive value indicates agreement.

Performance

Performance issues deal directly with the outcome both for the group and the individual. Examples include the completed project and the ultimate value (grade) of that project for the group and satisfaction for the individual.

Outcome. Performance is better when meeting face-to-face than when using social media or virtual meetings. In particular, a higher grade is earned when more face-to-face meetings are used in group projects (see item #17, Table 1, below). And, while both face-to-face interaction and social media increase group effectiveness, FTF interaction results in greater effectiveness (items #40 and #12). Moreover, virtual meetings do not lead to higher grades (item #43), nor do they lead to better outcomes than FTF meetings (item #46). However, more virtual meetings do not lead to lower grades (item #34). So, participating in virtual meetings neither increase nor decrease student grades.

There is also a direct link between being more comfortable with group members and the grades received on group projects (item #59). FTF meetings help students feel more comfortable with others (item #63) while social media interactions do not increase the comfort level with group members (item #65). Thus, higher grades are attained when groups have more face-to-face meetings.

Table 1. Performance: Outcome

		№	Question	Mean	SD	*T*	*p*	df
Outcome	FTF	24	Face-to-face meetings result in better outcomes than virtual meetings	3.11^{***}	.71	7.70	<.001	79
		17	I earn a higher grade when group has more face-to-face meetings	3.00^{***}	.74	5.99	<.001	79
		30	My class grade is improved with face-to face teamwork	3.13^{***}	.68	8.43	<.001	81
		21	My groups perform better when meeting face-to-face than using social media	3.09^{***}	.72	7.32	<.001	81
		40	Face to face interaction is a good way to improve group effectiveness	3.15^{***}	.76	7.60	<.001	79
		63	Face-to-face meetings help me to feel more comfortable with my group members	3.16^{***}	.68	8.59	<.001	78
		59	My grades on group projects are better when I feel more comfortable with my group members	3.09^{***}	.81	6.45	<.001	79
	Virtual	46	Virtual meetings result in better outcomes than face-to-face meetings	2.15^{***}	.59	-5.33	<.001	80
		43	I earn a higher grade when my group has more virtual meetings	2.21^{***}	.68	-3.81	<.001	80
		34	My grades suffer when the more virtual meetings are used	2.23^{***}	.61	-3.86	<.001	80
	Social Media	28	Social media helps groups work only after you get to know group members	2.70^{*}	.76	2.39	.019	80
		61	When I work in groups, we perform better with social media interaction	2.50	.69	0	1	79
		6	Our group performance improves most when only social media interactions are used	2.02^{***}	.65	-6.56	<.001	80
		12	The use of social media improves team effectiveness	2.75^{***}	.64	3.54	<.001	80
		65	Social media interactions help me to feel more comfortable with my group members	2.35^{*}	.63	-2.18	.032	80

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Satisfaction. Face-to-face meetings provide more satisfaction than virtual meetings (see items #37, #53, and #10 Table 2, below). In addition, more virtual meetings do not increase satisfaction with other group members (item #18), which further strengthens the relationship between FTF interaction and satisfaction. Likewise, virtual meetings do not provide a better experience than face-to-face interaction (item #36).

Table 2. Performance: Satisfaction

		№	Question	Mean	SD	*t*	*p*	df
Satisfaction	FtF	37	Face-to-face meetings provide more satisfaction than virtual meetings	3.05^{***}	.71	6.91	<.001	81
		53	My satisfaction with other group members is greater when we have more face-to-face meetings than virtual meetings	3.00^{***}	.76	5.86	<.001	79
	Virtual	10	Virtual meetings provide more satisfaction than face-to-face meetings	2.06^{***}	.65	-6.07	<.001	81
		18	My satisfaction with other group members is greater when we use more virtual meetings than face-to-face interactions	2.23^{**}	.75	-3.20	.002	81
		36	Virtual meetings provided better experience for me than face-to-face meetings	2.23^{**}	.72	-3.35	.001	81

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Efficiency. FTF meetings are more efficient than virtual meetings, although virtual meetings do not waste more time (items #14, #29, #56, and #5, Table 3, below). One face-to-face or virtual meeting is neither better nor worse than several meetings of the opposite type (items #39 and #60).

Table 3. Performance: Efficiency

		№	Question	Mean	SD	*t*	*p*	df
Efficiency	FtF	14	Face-to-face meetings are more efficient than virtual meetings	3.07^{***}	.75	6.84	<.001	80
		5	When my group members have a face-to-face meeting, we waste more time	2.27^{*}	.88	-2.28	.042	80
		60	One long face-to-face meeting is more effective than several virtual meetings	2.59	.72	1.08	.283	79
		66	Face-to-face meetings typically take less time than virtual meetings.	2.48	.81	-.27	.783	79
	Virtual	29	Virtual meetings are more efficient than face-to-face meetings	2.20^{***}	.69	2.39	<.001	81
		56	When my group members have a virtual meeting, we waste more time	2.48	.77	-.28	.775	79
		1	Project demands require more virtual meetings than face-to-face meetings	2.49	.72	-.15	.879	81
		39	Several virtual meetings are more effective than even one long face-to-face meeting	2.35	.73	-1.84	.069	77

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Process

Process issues deal with how the group functions to work on and complete the project assignment. We investigate the formation of trust, how members work together and the use of communication. Additional elements of process we consider include degree of boredom, division of work and asking for help.

Trust. Trust, as we saw in the literature review, is one of the more important issues that groups must resolve. Face-to-face meetings are effective in building trust (item #20, Table 4, below). It is also a more preferred method than social media (item #58). FTF meetings result in stronger relations and getting to know team members better than virtual meetings (items #54, #42, and #23). Social media are not helpful in building trust (item #16), getting to know other members (item #62), and improving group activities (item #47). So, face-to-face interactions are better than both social media and virtual meetings in this area.

Table 4. Process: Trust

		№	Question	Mean	SD	*t*	*p*	df
Trust	FtF	20	Face-to-face meetings are effective in building trust with group members	3.34^{***}	.57	13.33	<.001	81
		58	I prefer to build trust with group members during face-to-face meetings as opposed to social media interactions	3.22^{***}	.71	8.95	<.001	78
		54	Face-to-face meetings result in stronger relations between team members than virtual meetings	3.10^{***}	.80	6.66	<.001	79
		23	Face-to-face meetings help me to get to know my group members better	3.34^{***}	.74	10.28	<.001	81
		9	I find face-to-face interactions better than social media interactions	3.22^{***}	.70	9.26	<.001	81
	Virtual	42	Virtual meetings result in stronger relations between team members than face-to-face meetings	2.13^{***}	.71	-4.67	<.001	79
	Social Media	16	Social media (e.g., facebook, Instagram) are effective in building trust with group members	2.48	.76	-.21	.827	80
		62	Social media interactions help me to get to know my group members better	2.48	.63	-.35	.726	79
		47	Social media interaction improves group activities	2.59	.70	1.19	.237	78

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Task-orientation. Both face-to-face meetings and social media are effective in encouraging project-related interactions (see items #35 and #4, Table 5, below). While FTF meetings do not distract group members from project tasks, social media interactions do (items #33 and #7). It is possible, therefore, that social media have considerable noise attached to the communications, as they have both positive and negative impact on working toward and completing project tasks. Multitasking occurs more in virtual meetings than FTF sessions (items #51 and #25). Group members are more focused on a task during FTF meetings while they are not more focused during virtual meetings (items #15 and #44).

Table 5. Process: Task-orientation

		№	Question	Mean	SD	*t*	*p*	df
Task-orientation	FtF	35	Face-to-face meetings encourage project-related interactions between group members	3.05^{***}	.71	6.91	<.001	81
		33	Face-to-face meetings distract group members from project tasks	2.21^{***}	.74	-3.53	<.001	81
		51	In face-to-face meetings, people are more likely to multitask than in virtual meetings	2.26^{**}	.72	-2.93	.004	79
		15	Group members are more focused on a task during a face-to-face meeting.	2.91^{***}	.71	6.84	<.001	80
	Virtual	25	In virtual meetings, people are more likely to multitask than in face-to-face meetings	3.13^{***}	.66	8.67	<.001	81
		44	Group members are more focused on a task during a virtual meeting.	2.21^{***}	.68	-3.73	<.001	79
	Social Media	4	Social media encourage project-related interaction between group members	2.68^{*}	.63	2.58	.012	78
		7	Social media interactions cause distraction from group work	2.74^{**}	.72	2.93	.004	79

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Information exchange/effective communication. Face-to-face meetings are more effective than social media interactions in encouraging the exchange of ideas (see item #3, Table 6, below). Students gain and remember more project-related information from face-to-face meetings than they do from virtual meetings (items #2, #52, #49, and #22). In addition, they are not more confused after face-to-face than virtual meetings (item #38) nor is communication more effective in virtual than face-to-face meetings (item #11). Overall, therefore, communication is more effective in face-to-face than in virtual meetings.

Table 6. Process: Information exchange, effective communication

		№	Question	Mean	SD	*t*	p	df
Information exchange, Effective	FtF	3	Face-to-face meetings are more effective than social media interactions in encouraging the exchange of ideas	3.07^{***}	.76	6.77	<.001	81
		2	I gain more project-related information from face-to-face meetings than I do from virtual meetings	2.96^{***}	.86	4.84	<.001	81
		52	I remember more information from face-to-face meetings than I do from virtual meetings	2.99^{***}	.73	5.91	<.001	79
		38	I am often more confused after face-to-face meetings than I am after virtual meetings	1.99^{***}	.71	-6.44	<.001	80
	Virtual	49	I gain more project-related information from virtual meetings than I do from face-to-face meetings	2.26^{**}	.67	-3.16	.002	79
		22	I remember more information from virtual meetings than I do from face-to-face meetings	2.24^{**}	.77	-2.97	.004	81
		8	I am often more confused after virtual meetings than I am after face-to-face meetings	2.43	.77	-.86	.392	81
		11	Communication is more effective in virtual meetings than in face-to-face meetings	2.13^{***}	.73	-4.52	<.001	81
		55	Social media interactions increase the exchange of ideas related to the group project	2.56	.70	.78	.433	79

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Boredom/division of work/asking for help. Boredom does not occur more often in face-to-face than virtual meetings (see item #32, Table 7, below). Students are more likely to ask for help in FTF sessions, which are also a better way to divide project work than social media interactions (items #50 and #41). Social media interactions do not enhance understanding of teammates' strengths more than face-to-face meetings (item #27).

Table 7. Process: Division of work, help asking, and boredom

		№	Question	Mean	SD	t	p	df
Division of work	FtF	50	Face-to-face meetings are a better way to divide project work than social media interactions	2.90***	.70	5.07	<.001	79
	Social Media	27	Social media interactions help me understand my group members' strengths more than face-to-face meetings	2.11***	.70	-4.95	<.001	80
		13	Social media interactions are a better way to divide group work than face-to-face interactions	2.34	.75	-1.89	.062	81
Help	FtF	41	Group members are more likely to ask for help in face-to-face meetings	3.03***	.69	6.69	<.001	78
	Social Media	45	Group members are more likely to ask for help by using social media interactions than face-to-face interactions	2.41	.72	-1.15	.251	80
Boredom	FtF	32	I am more often bored or uninterested during face-to-face meetings than I am in virtual meetings	2.05***	.68	-5.90	<.001	80
	Virtual	57	I am more often bored or uninterested during virtual meetings than I am in face-to-face meetings	2.65	.73	1.83	.070	79

*** $p < .001$; ** $p < .01$; * $p < .05$

Overall Preference

These students overwhelmingly prefer face-to-face meetings over virtual meetings (see items #26 and #64, Table 8, below). As a bottom line issue, this preference, when combined with or perhaps because of their experiences, indicates that FTF interactions are both more effective and create a better process or environment for successful group completion of project assignments.

Table 8. Preference

	№	Question	Mean	SD	t	p	df
FTF	26	I prefer face-to-face meetings over virtual meetings	3.2***	.70	9.02	<.001	79
Virtual	64	I prefer virtual meetings over face-to-face meetings	2.1***	.79	-4.25	<.001	80

*** $p < .001$; ** $p < .01$; * $p < .05$

Demographic Comparisons

No differences were found when we analyzed male vs female responses. There were several differences between undergraduate and graduate students. Both graduates and undergraduates find that FTF meetings provide more information than virtual meetings and are more effective than social media interaction in encouraging the exchange of ideas (see items #2 and #3, Table 9, below). Undergraduates, however, more strongly support these results (item #2, $t(78) = 5.37$, $p = .023$; item #3, $t(78) = 3.99$, $p = .049$). Undergraduates also find that social media encourage project-related interaction (item #4). Graduates are not more confused after virtual than they are FTF meetings (item #8) and continue their awareness and use of virtual meetings. They do note, more so than undergraduates, that more multitasking occurs in virtual than FTF meetings (item #25, $t(78) = 4.37$, $p = .040$). Graduates see that grades do not suffer when more virtual meetings are used (item #34, $t(77) = 4.47$, $p = .038$). Finally, undergraduates are bored during virtual meetings while graduates are not (item #57, $t(76) = 4.99$, $p = .029$). Undergraduate students are comfortable with and use face-to-face sessions; and, graduate students, while preferring FTF sessions, are less uncomfortable with virtual meetings.

Table 9. Demographic comparisons: Grads and Undergrads

	№	Question	Grads			Undergrads			
			M	SD	n	M	SD	n	t
FTF	2	I gain more project-related information from face-to-face meetings than I do from virtual meetings	2.77	.90	48	3.22	.76	31	5.37*
	3	Face-to-face meetings are more effective than social media interactions in encouraging the exchange of ideas	2.94	.83	48	3.29	.64	31	3.99*
Virtual	8	I am often more confused after virtual meetings than I am after face-to-face meetings	2.25	.67	48	2.61	.84	31	4.51*
	25	In virtual meetings, people are more likely to multitask than in face-to-face meetings	3.25	.67	48	2.94	.63	31	4.37*
	34	My grades suffer when the more virtual meetings are used	2.10	.56	48	2.4	.67	30	4.47*
	57	I am more often bored or uninterested during virtual meetings than I am in face-to-face meetings	2.49	.75	47	2.87	.68	30	4.99*
Social Media	4	Social media encourage project-related interaction between group members	2.52	.59	46	2.93	.58	30	8.98**

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

Four significant differences were found when comparing students who had less than 30 and 30 or more hours of outside commitments. Three of these differences fit the picture drawn above for graduate and undergraduate students, i.e., students with less than 30 hours respond like undergraduates while those with higher commitment respond like graduate students. Respondents with less than 30 hours indicate that social media encourage project-related interaction (see item #4 in Table 10, below, $t(75) = 4.26$, $p = .042$) and agree more strongly that FTF meetings are more satisfying than virtual meetings (item #37, $t(78) = 4.02$, $p = .048$). Those with higher time commitments prefer knowing teammates before starting even more so than those with less (item #48, $t(76) = 4.78$, $p = .032$). The fourth difference contradicts this picture: higher commitment respondents note that social media improves team effectiveness (item #12, $t(77) = 5.81$, $p = .018$).

Table 10. Demographic comparisons: Outside Commitments

	№	Question	More than 30 hours			Less than 30 hours			
			M	SD	n	M	SD	n	t
FtF	37	Face-to-face meetings provide more satisfaction than virtual meetings	2.84	.80	32	3.17	.63	47	4.02*
	48	I like to be in teams where I know everyone beforehand	3.22	.75	32	2.89	.57	45	4.78*
Social Media	4	Social media encourage project-related interaction between group members.	2.48	.62	31	2.78	.59	45	4.26*
	12	The use of social media improves team effectiveness.	2.97	.54	31	2.62	.67	47	5.81*

$^{***}p < .001$; $^{**}p < .01$; $^{*}p < .05$

DISCUSSION AND CONCLUSIONS

This research began by trying to compare face-to-face human interactions with the large category of technology-based communication. Our initial focus was the role of technology (lumped all together) in the experiences and preferences of students. Face-to-face meetings utilize two sets of activities, on-task and off-task, to accomplish its purposes dealing with processes and outcomes. For technology-based interactions to be an effective surrogate, they must also successfully accomplish these purposes. Analyzing this data led to a series of conclusions about the relationship between face-to-face meetings and technology-based interactions. We found that technology-based interactions are divided into two distinct categories: virtual meetings, a more formal set of digital communication tools to complete task related activities; and, social media, a more informal set of digital communication interaction approaches that fulfill off-task activities. Virtual meetings and social media, alone or in combination, are not as successful as FTF meetings in fulfilling its purposes.

Our study finds face-to-face interaction brings unsurpassed results in group output. We recommend, therefore, that faculty create opportunities within their course structures for increasing student involvement and peer interaction through FTF meetings. Our results, consistent with the literature, indicate resolving trust issues early in the semester improves group processes. Since trust is most established through FTF interaction, rather than any other mode, we suggest considering the early use of student collaboration for this purpose. Moreover, face-to-face interaction is the best way for students to get to know and feel more comfortable with each other, which further improves group performance. These results are consistent with Choi and colleagues (2017) who describe a "U-shaped" curve, where FTF interaction is found to be more effective in the beginning and ending stages of group projects and less effective in the middle stages where technology-based interactions are most effective. Face-to-face meetings near the beginning of a semester also help groups more effectively divide up project work. These FTF meetings have a direct and positive impact on satisfaction, both as a desired outcome and as a facilitator in creating higher group performance, greater member interaction and better course experience. Group communication also improves when face-to-face meetings take place during the semester. That is, people have less confusion, gain and remember more knowledge, maintain greater focus, have a better exchange of ideas and create a more effective communication process.

Demographic information regarding the level of education shows that the MBA students see more benefits of virtual meetings perhaps because they are more used to this form of collaboration at work. Undergraduate students are less likely to have this experience and, therefore, see fewer benefits in virtual meetings. They have more face-to-face meetings on campus since time is available between classes to meet with other students.

Interestingly, one area in which we find results quite different from what we expected deal with the use of virtual rather than FTF meetings. We find that even part-time MBA students, not just undergraduates, have a more positive experience with face-to-face meetings over virtual meetings. Given that many part-time MBA students have jobs and additional commitments, we originally thought they would prefer virtual over FTF meetings due to higher time and travel constraints (see, for example, Denstadli, et al., 2012). Our survey results, however, indicate they prefer FTF over virtual meetings at the same level as undergrads. Cramton (2001) notes that typed communication in technology-based interaction is more time-consuming and includes response delays that decrease the efficiency of virtual meetings. In addition, he indicates how the lack of nonverbal communication reduces the actual amount of information in messages (Cramton, 2001). Our students, likewise, indicate FTF meetings are more efficient than virtual sessions (see items #14 and #29). The model presented in Denstadli, et al., (2012) suggests that as the work becomes more complex, there will be a higher preference for FTF meetings to better accomplish the task. A group project assignment is a complex task and preference is greater for use of face-to-face sessions. These MBA students are aware of the advantages of FTF meetings based on their educational and work experiences to date.

Our overall conclusion is: face-to-face interaction brings unsurpassed results in group output. This is true for all types of respondents, including graduate and undergraduate students, less or greater than 30 hours of time commitments outside of class, and gender. FTF meetings result in higher performance, improved group effectiveness, greater satisfaction, higher efficiency, greater trust, and enhanced overall communication. And, regardless of demographic group, FTF is the preferred method of interaction, as well as the preferred way to build trust. Surprisingly, neither virtual meetings nor social media have significantly more positive impact on groups than face-to-face interaction along any dimension studied. Moreover, technology is not a panacea. Therefore, faculty need to consider whether to include new technology into the context of a course. It should only be included if it enhances activities in improving the fulfillment of process and/or outcome results.

FUTURE RESEARCH

Future research will help in more fully identifying the categories of technology-based interactions, perhaps adding to the two types we included in this study. It may also study the specific roles each of these categories fulfill, as well as the level of effectiveness for each type. An earlier research study describes an effective integration of both FTF and technology-based interactions (Choi et al., 2017). This combination should be more fully investigated to provide additional information to faculty/trainers as they prepare students and managers in accomplishing their work-related tasks. It would also be beneficial to understand the impact of social media on satisfaction and other measures of outcome compared to virtual meetings.

We have assumed all undergraduate students would respond similarly. It is possible that undergraduate seniors are really more closely aligned with graduate students than they are with undergraduate freshmen or sophomores. This needs to be further tested.

Our research question of whether students' experiences in the completion of group assignments support and encourage an increased emphasis on educational technology by faculty is answered with an emphatic "No!" Additional research is required to determine how generalizable this response is and whether the answer needs to be modified in any way.

REFERENCES

Breuer, C., Hüffmeier, J., & Hertel, G. (2016). Does trust matter more in virtual teams? A meta-analysis of trust and team effectiveness considering virtuality and documentation as moderators. *Journal of Applied Psychology*, *101*(8), 1151-1177.

Buzzard, C., Crittenden, V. L., Crittenden, W. F., & McCarty, P. (2011). The use of digital technologies in the classroom: A teaching and learning perspective. *Journal of Marketing Education*, *33*(2), 131-139.

Cheng, W., & Warren, M. (2000). Making a difference: Using peers to assess individual students' contributions to a group project. *Teaching in Higher Education*, *5*(2), 243-255.

Choi, W., Zeff, L., & Higby, M. (2017). Students' Experiences with Face-to-face and Technology-based Communication in Group Assignments: A "U-Shaped" Model, *Quarterly Review of Business Disciplines, 4*(3), 261-277.

Chu, Hui-Chun. (2014) Potential Negative Effects of Mobile Learning on Students' Learning Achievement and Cognitive Load--A Format Assessment Perspective. *Journal of Educational Technology & Society, 17*(1).

Cohen, E. G., & Lotan, R. A. (2014). *Designing Groupwork: Strategies for the Heterogeneous Classroom*. New York, NY: Teachers College Press.

Cramton, C. D. (2001). The mutual knowledge problem and its consequences for dispersed collaboration. *Organization science*, *12*(3), 346-371.

Crook, C. (2008). *Web 2.0 technologies for learning: The current landscape–opportunities, challenges and tensions*. Retrieved from http://dera.ioe.ac.uk/1474/

Cui, G., Lockee, B., & Meng, C. (2013). Building modern online social presence: A review of social presence theory and its instructional design implications for future trends. *Education and information technologies*, *18*(4), 661-685.

Davidson, N., Major, C. H., & Michaelsen, L. K. (2014). Small-group learning in higher education—cooperative, collaborative, problem-based, and team-based learning: an introduction by the guest editors. *Journal on Excellence in College Teaching*, *25*(3), 1-6.

Davis, J. H., Schoorman, F. D., Mayer, R. C., & Tan, H. H. (2000). The trusted general manager and business unit performance: Empirical evidence of a competitive advantage. *Strategic Management Journal*, *21*(5), 563–576.

De Jong, B. A., & Elfring, T. (2010). How does trust affect the performance of ongoing teams? The mediating role of reflexivity, monitoring, and effort. *Academy of Management Journal*, *53*, 535–549

Denstadli, J. M., Julsrud, T. E., & Hjorthol, R. J. (2012). Videoconferencing as a mode of communication: A comparative study of the use of videoconferencing and face-to-face meetings. *Journal of Business and Technical Communication*, *26*(1), 65-91.

Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of marketing research*, *40*(2), 146-160.

Driskell, J. E., Radtke, P. H., & Salas, E. (2003). Virtual teams: Effects of technological mediation on team performance. *Group Dynamics: Theory, Research, and Practice*, *7*(4), 297.

Gordon, R., & Connor, R. (2001). Team-based learning in management education. In D. Boud, R. Cohen, & J. Sampson (Eds.), *Peer learning in higher education: Learning from & with each other* (pp. 85–98). London: Kogan Page Limited.

Gunter, B., Rowlands, I., & Nicholas, D. (2009). *The Google Generation: Are ICT innovations Changing information Seeking Behaviour?*. Oxford: Elsevier.

Hambley, L. A., O'Neill, T. A., & Kline, T. J. (2007). Virtual team leadership: The effects of leadership style and communication medium on team interaction styles and outcomes. *Organizational Behavior and Human Decision Processes*, *103*(1), 1-20.

Henderson, M., Selwyn, N., Finger, G., & Aston, R. (2015). Students' everyday engagement with digital technology in university: exploring patterns of use and 'usefulness'. *Journal of Higher Education Policy and Management*, *37*(3), 308-319.

Horwitz, S. K., & Horwitz, I. B. (2007). The effects of team diversity on team outcomes: A meta-analytic review of team demography. *Journal of Management*, *33*(6), 987-1015.

IBM Corporation (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corporation.

Kirschner, P. A., & De Bruyckere, P. (2017). The myths of the digital native and the multitasker. *Teaching and Teacher Education*, *67*, 135-142.

Kennedy, D. M., Vozdolska, R. R., & McComb, S. A. (2010). Team decision making in computer-supported cooperative work: How initial computer-mediated or face-to-face meetings set the stage for later outcomes. *Decision Sciences*, *41*(4), 933-954.

Kvavik, R. B. (2005). Convenience, communications, and control: How students use technology. In D. Oblinger & J. Oblinger. In *Educating the net generation* (pp. 7.1-20). Washington, DC: Educause.

Lantz, A., (2001). Meetings in a distributed group of experts: comparing face-to-face, chat and collaborative virtual environments. *Behaviour and Information Technology*, *20*(2), 111-117.

Lu, B., Fan, W., & Zhou, M. (2016). Social presence, trust, and social commerce purchase intention: An empirical research. *Computers in Human Behavior*, *56*, 225-237.

McCorkle, D. E., Reardon, J., Alexander, J. F., Kling, N. D., Harris, R. C., & Iyer, R. V. (1999). Undergraduate marketing students, group projects, and teamwork: The good, the bad, and the ugly? *Journal of Marketing Education*, *21*(2), 106-117.

McKnight, K., O'Malley, K., Ruzic, R., Horsley, M. K., Franey, J. J., & Bassett, K. (2016). Teaching in a digital age: How educators use technology to improve student learning. *Journal of research on technology in education*, *48*(3), 194-211.

McLoughlin, C., & Lee, M. J. (2010). Pedagogy 2.0: critical challenges and responses. In *Web 2.0-Based E-Learning: Applying Social Informatics for Tertiary Teaching: Applying Social Informatics for Tertiary Teaching* (pp. 43-69). Hershey, PA: Information Science Reference

Rains, S. A. (2005). Leveling the organizational playing field—virtually: A meta-analysis of experimental research assessing the impact of group support system use on member influence behaviors. *Communication Research*, *32*(2), 193-234.

Reio, T., & Crim, S. (2006). *The emergence of social presence as an overlooked factor in asynchronous online learning*. Retrieved from http://www.eric.ed.gov/ PDFS/ED492785.pdf

Robert, L. P., Denis, A. R., & Hung, Y. T. C. (2009). Individual swift trust and knowledge-based trust in face-to-face and virtual team members. *Journal of Management Information Systems*, *26*(2), 241-279.

Roberts, J., & Kidd, T. T. (2017). Millennials, Digital Natives, and the Emergence of New Educational Spaces. In *Handbook of Research on Instructional Systems and Educational Technology* (pp. 1-10). Hershey, PA: IGI Global.

Selwyn, N. (2009). The digital native – myth and reality. *Aslib Proceedings: New Information Perspectives*, *61*, 364–379.

Selwyn, N. (2012). *Social media in higher education*. Retrieved from http://sites.jmu.edu /flippEDout/files/2013/04/sample-essay-selwyn.pdf

Selwyn, N. (2016). *Education and technology: Key issues and debates*. London: Bloomsbury Publishing.

Shirky, C. (2008). *Here comes everybody: The power of organizing without organizations*. New York, NY: Penguin.

Smyth, J. D., Dillman, D. A., Christian, L. M., & Stern, M. J. (2006). Comparing check-all and forced-choice question formats in web surveys. *Public Opinion Quarterly*, *70*(1), 66-77.

Szeto, E., & Cheng, A. Y. (2016). Towards a framework of interactions in a blended synchronous learning environment: what effects are there on students' social presence experience? *Interactive Learning Environments*, *24*(3), 487-503.

Thompson, L. F., & Coovert, M. D. (2003). Teamwork online: The effects of computer conferencing on perceived confusion, satisfaction and postdiscussion accuracy. *Group Dynamics: Theory, Research, and Practice*, *7*(2), 135.

Waycott, J., Bennett, S., Kennedy, G., Dalgarno, B., & Gray, K. (2010). Digital divides? Student and staff perceptions of information and communication technologies. *Computers & education*, *54*(4), 1202-1211.

APPENDIX

Research Survey: Social Media and Group Performance

Part I: Demographics: Please check the appropriate box.

Gender: ☐ Male ☐ Female

Born: ☐ Before 1980 ☐ 1980-1989 ☐ 1990-1999 ☐ After 2000

Years of College completed: ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ Some Grad work

Out of class commitments/week ☐ 0-9 hrs ☐ 10-19 hrs ☐ 20-29 hrs ☐ 30+ hrs

I have access to and use: ☐ Smartphone ☐ Computer ☐ Tablet

Part II: Please respond to the following questions ***based on your experiences with group projects***.
Check the degree to which you agree or disagree with each of the following statements, where 1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree.

Please check the appropriate level of agreement	Strongly Disagree 1	Disagree 2	Agree 3	Strongly Agree 4
1. Project demands require more virtual meetings than face-to-face meetings				
2. I gain more project-related information from face-to-face meetings than I do from virtual meetings				
3. Face-to-face meetings are more effective than social media interactions in encouraging the exchange of ideas				
4. Social media encourage project-related interaction between group members				
5. When my group members have a face-to-face meeting, we waste more time				
6. Our group performance improves most when only social media interactions are used				
7. Social media interactions cause distraction from group work				
8. I am often more confused after virtual meetings than I am after face-to-face meetings				
9. I find face-to-face interactions better than social media interactions				
10. Virtual meetings provide more satisfaction than face-to-face meetings				
11. Communication is more effective in virtual meetings than in face-to-face meetings				
12. The use of social media improves team effectiveness				
13. Social media interactions are a better way to divide group work than face-to-face interactions				
14. Face-to-face meetings are more efficient than virtual meetings				
15. Group members are more focused on a task during a face-to-face meeting.				
16. Social media (e.g., facebook, Instagram) are effective in building trust with group members				
17. I earn a higher grade when group has more face-to-face meetings				
18. My satisfaction with other group members is greater when we use more virtual meetings than face-to-face interactions				
19. I prefer to work in groups formed by the instructor				
20. Face-to-face meetings are effective in building trust with group members				
21. My groups perform better when meeting face-to-face than using social media				
22. I remember more information from virtual meetings than I do from face-to-face meetings				
23. Face-to-face meetings help me to get to know my group members better				
24. Face-to-face meetings result in better outcomes than virtual meetings				
25. In virtual meetings, people are more likely to multitask than in face-to-face meetings				
26. I prefer face-to-face meetings over virtual meetings				
27. Social media interactions help me understand my group members' strengths more than face-to-face meetings				
28. Social media helps groups work only after you get to know group members				
29. Virtual meetings are more efficient than face-to-face meetings				

Please turn the page. ⇒

Please check the appropriate level of agreement	Strongly Disagree 1	Disagree 2	Agree 3	Strongly Agree 4
30. My class grade is improved with face-to face teamwork				
31. Communication is more effective in face-to-face meetings than in virtual meetings				
32. I am more often bored or uninterested during face-to-face meetings than I am in virtual meetings				
33. Face-to-face meetings distract group members from project tasks				
34. My grades suffer when more virtual meetings are used				
35. Face-to-face meetings encourage project-related interactions between group members				
36. Virtual meetings provided better experience for me than face-to-face meetings				
37. Face-to-face meetings provide more satisfaction than virtual meetings				
38. I am often more confused after face-to-face meetings than I am after virtual meetings				
39. Several virtual meetings are more effective than even one long face-to-face meeting				
40. Face to face interaction is a good way to improve group effectiveness				
41. Group members are more likely to ask for help in face-to-face meetings				
42. Virtual meetings result in stronger relations between team members than face-to-face meetings				
43. I earn a higher grade when my group has more virtual meetings				
44. Group members are more focused on a task during a virtual meeting.				
45. Group members are more likely to ask for help by using social media interactions than face-to-face interactions				
46. Virtual meetings result in better outcomes than face-to-face meetings				
47. Social media interaction improves group activities				
48. I like to be in teams where I know everyone beforehand				
49. I gain more project-related information from virtual meetings than I do from face-to-face meetings				
50. Face-to-face meetings are a better way to divide project work than social media interactions				
51. In face-to-face meetings, people are more likely to multitask than in virtual meetings				
52. I remember more information from face-to-face meetings than I do from virtual meetings				
53. My satisfaction with other group members is greater when we have more face-to-face meetings than virtual meetings				
54. Face-to-face meetings result in stronger relations between team members than virtual meetings				
55. Social media interactions increase the exchange of ideas related to the group project				
56. When my group members have a virtual meeting, we waste more time				
57. I am more often bored or uninterested during virtual meetings than I am in face-to-face meetings				
58. I prefer to build trust with group members during face-to-face meetings as opposed to social media interactions				
59. My grades on group projects are better when I feel more comfortable with my group members				
60. One long face-to-face meeting is more effective than several virtual meetings				
61. When I work in groups, we perform better with social media interaction				
62. Social media interactions help me to get to know my group members better				
63. Face-to-face meetings help me to feel more comfortable with my group members				
64. I prefer virtual meetings over face-to-face meetings				
65. Social media interactions help me to feel more comfortable with my group members				
66. Face-to-face meetings typically take less time than virtual meetings				

*** Thank you for your responses.**

This page has been intentionally left blank.

SECURING THE CYBER PIPELINE: TOWARD NATIONAL STRATEGIES FOR CYBER WORKFORCE DEVELOPMENT

Tamirat T. Abegaz, University of North Georgia

Bryson R. Payne, University of North Georgia

ABSTRACT

As nations race to build their cyber workforces, a critical shortage of highly skilled labor in cyber is hampering efforts and weakening defensive capabilities as rogue actors progressively grow their offensive capacity. A key element of national policy and strategy will be the development of an adequate pipeline of competent, qualified cyber professionals for the next twenty years and beyond. In one such effort, the United States' National Security Agency, in collaboration with the National Science Foundation, has developed and implemented a program targeted at pipeline development from primary school through college and is sharing information on the program with the international community. This paper presents the NSA-NSF GenCyber project, along with research related to the program's effectiveness, as one approach toward multiplying both cyber and broader related fields' career interest among students in primary and secondary schools as a means to bring forth significantly greater numbers of university graduates in cyber security, computer education, and related fields. Overall, this research suggests that cyber workforce development initiatives like the NSA-NSF GenCyber project can form the basis for building the next generation of cyber professionals and researchers.

Keywords: cooperative/collaborative learning, teaching/learning strategies, pedagogical issues, cybersecurity, gender studies

INTRODUCTION

Computer and Information Systems career interest in the 21st Century

Research indicates that the interest among K-12 students' in science, technology, engineering, and mathematics (STEM) career path is declining (Sadler, Sonnert, Hazari, & Tai, 2012). In fact, the lack of interest in STEM career paths has become a global educational concern. Today, most developing and developed countries realize that STEM is a high-demand career field, important for the sustainable growth of the near-future global economy. Countries have realized that investing in STEM training is extremely valuable for continual innovation, both to improve the livelihood of the world's population and to benefit society as a whole. For instance, in the US, the demand for personnel with cyber expertise in both the public and private sectors for STEM-related professionals exceeds the current supply (Chen, 2013; Sadler et al., 2012). Unfortunately, in almost all STEM-related career professions, women are under-represented (Clark, 2005; Wang, Eccles, & Kenny, 2013). Given the disproportionately low number of girls who are interested in STEM-related fields, it is important for society to encourage and provide the necessary assistance to citizens of all gender and ethnic origins to pursue STEM education. In other words, opening doors early to girls in STEM disciplines and careers, by empowering students with the necessary knowledge and skills in STEM-related fields, is extremely helpful for the sustainability of modern society. Arguably, this could be the best educational investment to enhance the pipeline for future STEM career paths. Meanwhile, the combined efforts by all stakeholders including parents, community leaders, teachers, and governments will help to create a strong and continuing career pathway for both girls and boys in STEM-related fields. Increasing the representation of skilled manpower in STEM-related scientific and technical fields is essential for today's increasingly interconnected global economy. Alternatively, as STEM skills become even more important, strengthening the workforce pipeline will be tremendously helpful for society and impact the potential progress of the world economy at large. Realizing this, various countries have set out long-term and short-term educational objectives to minimize the shortage of skilled human-power in STEM, computer science, cybersecurity, and related fields.

From the US perspective, the National Research Council (NRC), the National Science Foundation (NSF), the National Security Agency (NSA), the Department of Education (DOE), and the Department of Commerce (DOC) are working hand-in hand to prepare students for careers in STEM-related fields (Mau, 2016; Melissa & Bianca, 2015; NRC, 2011). For instance, a report from the NSF indicates that STEM-related workforce demand increased more than 21% and significantly outpaced other career fields (Lehming et al., 2010, Mau, 2013). With regard to K-12 STEM education, three goals were set forth by the US government to overcome current and future shortages of experts in STEM-related fields: "to increase advanced training and careers in STEM fields, to expand the STEM-capable workforce, and to increase scientific literacy for all students" (NRC, 2011). Particular, it is essential for the government to invest its resources at the family, school, district, state, and national levels to improve STEM education. But further changes in society's STEM culture are needed in order to attract more girls into STEM-related fields to minimize the gender gap. In addition, we must work together as a nation in order to change the attitudes of female students that stereotype STEM fields as careers for males. Recently, various government agencies have collaborated to increase the STEM pipeline for students in entering, completing, and persisting in STEM disciplines. Most recently, due to the efforts of various governmental and non-governmental stakeholders, the gender gap may be turning in the right direction (Hyde, Lindberg, Linn, Ellis, & Williams, 2008; Mau, 2016; NRC, 2011). While this has a positive impact, tremendous effort is required to reach to the expected results.

GenCyber training program

The GenCyber program states three main goals: "to increase interest in cybersecurity careers and diversity in the cybersecurity workforce of the Nation, to help all students understand correct and safe online behavior, and to improve teaching methods for delivering cybersecurity content for K-12 curricula" (GenCyber Program Director Guide, 2016). The GenCyber program sponsors free summer cyber camps for K-12 students across the country each summer. Each camp is expected to create opportunities for participants to gain a thorough understanding of cybersecurity principles and practices. Students are expected to leave the camp with a greater awareness of personal, organizational, and national cybersecurity issues, practical experience applying basic cyber hygiene, and the ability to research, analyze, and assess ethical issues in cybersecurity. Generally, the program focuses on delivering ten cybersecurity first principles: abstraction, process isolation, domain separation, resource encapsulation, information hiding, simplicity, least privilege, layering, modularization, and minimization. Achieving each principle is believed to enhance security in some respect.

For the sake of simplicity, we can categorize the first cyber principles into two sets: network and system. The network category includes domain separation, layering, resource encapsulation, and minimization. **Domain separation** is a mechanism to protect one functionality, task, or data from interfering with another to enforce security and protection. Similarly, **layering** encourages us to build multiple levels of defense to ensure resilience against attack. **Resource encapsulation** is another cybersecurity first principle that enables manipulation of resources only as intended by the resource owners to prevent unauthorized access. Lastly, from a cybersecurity point of view, the goal of **minimization** is to reduce the number of possible attack vectors, such as by turning off unused ports and unnecessary features. The system category includes abstraction, process isolation, least privilege, simplicity, modularization, and information hiding. From a software engineering point of view, **abstraction** is a design principle that enforces the minimization of unnecessary clutter from a system that can distract and possibly lead to complexity, which could make the system difficult to manage. **Process isolation** is a mechanism that enables systems to execute on the same platform without interfering with one another. **Least privilege** advocates a strategy of assigning minimum but sufficient power to manage system resources by ensuring correct operation, security, and protection. On the other hand, **simplicity of design** promotes the reduction of unnecessary details to in an effort to accomplish a reliable and secure system. **Information hiding** enforces secure coding by requiring programmers to expose only the necessary functions to external applications. Lastly, **modularity** emphasizes separation of functionality to enhance code security and protection. Overall, the first principles of cybersecurity are designated as the fundamental concepts of the GenCyber curriculum. A solid understanding of the first principles of cybersecurity is important to produce talented individuals for cybersecurity industry and government roles.

Research Objectives

The purpose of this research is to evaluate the impact of a representative GenCyber training program on students' future career paths. The GenCyber summer program aims to grow the future generation of computer science experts in general, and cybersecurity professionals in particular. The main goal of this research is to explore the impact of a GenCyber training program on the interest of K-12 students toward STEM fields, specifically in the area of computer science and cybersecurity. In addition, this research also evaluates its impact in addressing gender parity between boys and girls in their future career interests. Based on the above, this research identifies the following two questions: first, would participating in the GenCyber summer program impact K-12 students' interest in future STEM careers? And second, would participating in the GenCyber summer program minimize students' gender bias toward future STEM careers? The objective is to gain further information about factors that contribute to the underrepresentation of girls in STEM career pathways. The remainder of this paper is organized as follows: In section 2, related works relevant to this research are presented. Section 3 discusses the methodology used for this research. Section 4 presents the experimental results. Finally, a discussion of the results and their implications, and the conclusion, are presented in section 5.

RELATED WORK

Various studies have explored a number of models in their attempt to explain at what time in a person's life their career interest emerges, how it develops and matures, and how volatile or persistent the choice may be over time (Bandura, 1986; Coley, 2010; Crissey, 2009; Fouad 2007; Gibons, 2004; Lent, Brown, & Hackett, 1994). For instance, Boni et al. associated career choice with five aspects of learning: empathy (willingness to try new ideas, tools), integrative thinking (thinking outside the box), optimism (I believe I can do it), experimentation (let's experiment with this), and collaboration (ability and willingness to working with others) (Agogino 2007; Boni, Arthur, Laurie, & Shelley, 2009; Li 1999). On the other hand, the Social Cognitive Career Theory (SCCT) defines interest as a person's "pattern of likes, dislikes and indifferences" with regard to a particular field or subject matter (Lent & Brown, 2006). For instance, the SCCT attempted to explain the reason behind how people develop a specific career interest, how they reach a decision on making career choices, and how they deal with obstacles that hinder them form achieving their career goals (Lent, Brown, & Hackett, 1994). Overall, SCCT stated that career interests are potentially determined and regulated by self-efficacy, outcome expectations, and goals (Bandura, 1986; Christensen, Knezek, & Tyler-Wood, 2015; Gibons, 2004[Lent et al., 1994; Sadler et al., 2012).

Self-efficacy refers to an individual's attitude about their ability to successfully complete an assigned task. It is influenced by cognitive, social and situational factors. On the other hand, outcome expectations refer to the perceived results, either positive or negative, obtained from performing certain tasks. Finally, an individual's goals could depict the final decision as to whether to begin a particular career path. Particularly, research suggests that students would develop career interest in a particular field if the subject is engaging, if they feel that they possess personal competency and will experience positive outcomes. However, if they feel that they have low personal competency, they will tend to shy away from a particular career path. Therefore, barriers such as bias due to gender or ethnicity could create negative impacts on career interests. Overall, students' ability to self-categorize themselves as future STEM professionals is vital in shaping their success in achieving their career goals. This self-empathy (the is the act of giving oneself empathy, and optimism (positivity and confidence about oneself) could make a difference in shaping their college performance such as choosing a STEM major, persisting in their majors, and completing their college degree. Improving these contributing factors could considerably enhance the STEM pipeline. Therefore, school counsellors, family members, community leaders, and government officials should work hand-in hand in creating a conducive environment to encourage and engage students in STEM courses early and often throughout their education and training.

METHODS

Participants

The National Cyber Warrior Academy (NCWA) is a national cybersecurity awareness and ethical cyber operations training program based on the GenCyber framework, aimed to enhance interest in cybersecurity careers and help students understand how to protect themselves from cybercriminals. It is a two-week, residential cyber camp with over 80 hours of instruction, including more than 40 hours of hands-on labs in the area of computer science and cybersecurity. The academy is geared toward students interested in cybersecurity studies, and is expected to enhance the interest of high school sophomores, juniors and seniors in STEM-related careers. Participant recruitment took three main forms: printed brochures mailed to 212 high school principals in the university's 32-county service area, emails sent to over 2,000 high school advisement counsellors and instructors in the southeast region, and a program website and press releases from institutional university relations staff disseminated electronically. The majority of the applications received were from in-state applicants, primarily in the university's traditional 32-county service area, but a number of out-of-area and out-of-state students also applied.

Sample Characteristics

The program staff reviewed all 137 applications received and ranked the applications based on merit: by grade point average (GPA), students' self-reported computer interest as demonstrated by a written essay and student experience with computing or involvement in extra-curricular computing activities (programming, robotics, or cyber competition teams or related clubs). Due to the university's emphasis on global engagement and strategic languages, priority consideration was given to students with experience or proficiency in a Department of Defense (DoD) strategic language including Arabic, Mandarin Chinese, Dari, Hindi, Korean, Portuguese, Persian, Russian, Turkish, Swahili, and Urdu. The effect size of 0.5 was used to estimate the size of the participants using G*Power (Erceg-Hurn, & Vikki, 2008; Faul et.al, 2007; López et. al., 2015). Forty applicants with an average weighted GPA above 3.8, and highly diverse, with 24 males and 16 females (60% male, 40% female) were selected for the training program. As shown in Fig. 1, 55% students who self-identified as Caucasian, 20% as Asian, 12.5% as African American, and the remaining 10% as mixed ethnic groups, respectively. With regard to age, as expected all participants were between the ages of 14 and 17. More than 92% of the participants were between the ages of 15 and 17 while the remaining three (7%) were aged 14.

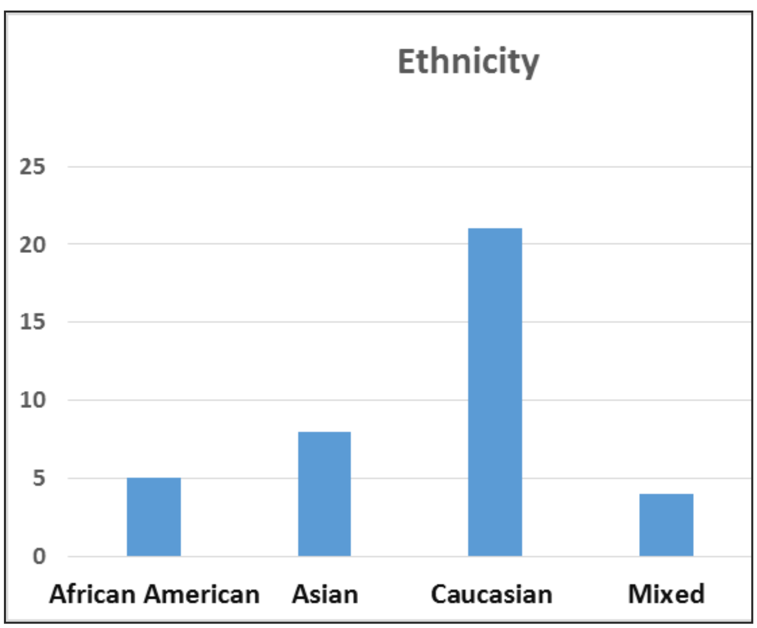


Figure 1. Ethnicity Distribution

Protocol and materials

NCWA GenCyber program began with parents dropping off students. Upon arrival, the participant was greeted, his or her identity was verified, and parents signed various release forms, including consent to participate in the IRB-approved research study. Once the researchers verified the parental consent form, then the PreGenCyber survey questionnaire was provided to the participants. This questionnaire was targeted to collect background information from the participant directly and to capture the career interest of the participants prior to attending the GenCyber summer training program. In addition, participants were asked to self-rate their computer science and cybersecurity skills before and after the GenCyber training program on a scale from 1 (less proficient) to 6 (extremely proficient). A copy of the informed consent document and STEM career interest questionnaire was provided to each participant upon request. After the participant completed the survey, he or she was then thanked and escorted to their living quarters by cadet counsellors. Each day of instruction, students participated in physical recreation activities before breakfast, not at the level of physical readiness training (PRT) for the Corps of Cadets, but enough to get their blood flowing and prepare their minds and bodies for intensive cyber training all day long. Class began at 9 AM, with lunch from 12-1 PM, lab instruction from 1-5 PM, followed by dinner and two to three hours of planned evening activities, including guest speakers and group activities such as drone programming, Sphero robot activities, car-hacking, 3D printing, capture-the-flag exercises, and NAO robotics.

The primary curriculum for the program consisted of the EC-Council's Certified Ethical Hacker (CEH) training material, specifically, the hands-on labs (EC-Council, 2016). The CEH curriculum consists of 18 modules, from hacking individual operating systems to web servers to mobile devices, and from cryptography to cloud computing to social engineering. The core focus of CEH is to look for weaknesses and vulnerabilities to assess the security of target systems and the lab manual includes over 700 pages of step-by-step security and vulnerability testing labs, with dozens of additional lab activities available through the EC-Council web portal. In addition, each day, one of the ten first principles was discussed in detail. In addition, students were asked to come up with a 3D printed object to embody each of the cyber first principles, and they used 3D printers to produce the objects and gave a presentation to help their fellow students understand why the particular object represented that concept or principle. Team-building activities were woven throughout the program. Finally, a PostGenCyber training survey questionnaire was given to the students before they left the academy.

Hardware and software

VMWare running nine virtual machines (Kali, Ubuntu and various Windows OS versions) from the CEH curriculum plus instructor-supplied materials served as the primary workstations. A wide variety of open-source and free software tools were used including Oracle VirtualBox VM, Kali/Metasploit, Wireshark, Snort, OpenGarages, and many more. The mini drones used for the drone programming/hacking exercises were Parrot Mini Cargo Drones (six total). The Sphero robot orbs (five total) were loaned to the program from one of the schools we partner with. The NAO robot was used for the last full evening of elective activities (owned by the computer science department). And, the 3D printers (three total) were XYZ Corp. Da Vinci Jr. 1.0W printers. All of these hardware items were in place before the GenCyber program, and are used in multiple programs at the university. A field trip to Georgia Tech Research Institute's (GTRI) security operations center (SOC) in Atlanta on the Saturday between the two weeks of instruction was conducted to enable students see real-time and aggregated information across ten 60-inch monitors in the unclassified level of GTRI's SOC.

Design

The experimental data points were collected using pre- and post-training surveys with a scale from 1 (less interested) to 6 (extremely interested). The data analysis in the study employs different categories of mixed factorial design. Each design includes gender & ethnicity as the between-subject design. Together with gender and ethnicity, two categories of future career interest were identified and investigated: STEM (STEM careers and computer science) and non-STEM (medical and social sciences). In addition, participants were also asked to self-rate their computer science area proficiency before and after the training. This was mainly used to capture their perceived leaning and self-confidence in perusing STEM careers. Based on the above description, five two-way mixed ANOVA factorial design analyses were conducted. Each factorial design is described in section 4 together with the analysis result.

RESULTS

As reported in section 3, five two-way mixed ANOVA factorial design analyses were conducted. Participants were asked to rate their future career interest in STEM, non-STEM, and computer science and cybersecurity related fields. In addition, participants were asked to self-rate their computer science and cybersecurity skills before and after the GenCyber summer program on a scale of 1 (less proficient) to 6 (extremely proficient). We predicted that GenCyber training would improve proficiencies in the area of computer science and related fields. Based on previous studies, we also predicted that there would be gender and ethnicity differences in computer proficiency ratings.

A preliminary analysis of ethnicity as a between-subject factor revealed no significant main or interaction effect, so it was omitted from further consideration. Therefore, this study was adjusted as a 2 (GenCyber training: PreGenCyber and PostGenCyber) by 2 (Gender: female and male) two-way, mixed analysis of variance (ANOVA) factorial design. The self-rated values are treated as numbers and the mean value was used instead of the median for better analysis as suggested by (David M & Mirosevich) to represent ratings of the participants for a given condition. ANOVA was used to analyse the statistical significance of each of the conditions.
The hypothesis, data analysis, and results are presented in the following subsequent sub-sections.

1. Impact of GenCyber training on future STEM career interest: data analysis and result

The hypothesis of the study of the impact of GenCyber training on future STEM career interest is stated as “GenCyber training enhances future career interest in STEM and related fields”. The first IV (independent variable) is gender, which includes two values: female and male. The second IV (i.e., GenCyber training) contains two conditions: PreGenCyber (baseline) and PostGenCyber training. The dependent variable (DV) is future career interest rating. The descriptive statistics of the influence of early computer science training on early future career interest in STEM field are provided in Table 1-1. It can be seen from Table 1-1 that the overall mean for future STEM career interest for females and males before and after the training are 24.53, 28.83, 23.00, and 29.13 respectively. This shows that the training shows an increase in career interest ratings for the male gender category. However, the mean STEM career interest for females before and after the training shows that the training has a negative impact for females. This is not consistent with the researchers' prediction. Further analysis was needed to determine the statistical significance of the means. As indicated, a two-way mixed ANOVA was performed to compare the effect of early GenCyber training on future STEM career interest of the participants. The result of the analysis is presented in Table 1-2.

Table 1-1. Descriptive Statistics of STEM career interest rate

GenCyber training	Gender	Mean	Std.Dev
PreGenCyber	Female	24.53	9.15
PreGenCyber	Male	28.83	5.65
PostGenCyber	Female	23.00	8.67
PostGenCyber	Male	29.13	7.60

As presented in Table 1-2, the analysis indicated there was no statistically significant difference either in the main effect of GenCyber training, $F(1, 37)=0.341$, $p>0.05$, or as an interaction effect between GenCyber training and gender, $F(1,37)=0.736$, $p>0.05$, on future career interest. Meanwhile, the result indicated that there is a statistically significant main effect of gender on future career interest, $F(1,37)=5.280$, $p(0.027)<0.05$, $\eta p2 =0.125$. However, the effect size was 0.125, which means that the effect of gender difference accounted for 12.5% of the between-group differences, suggesting that the impact of gender on future career interest is relatively minimal.

Table 1-2. Two-Way Mixed ANOVA for STEM career interest rate

Source of Variance	SS	df	MS	F	P(Sig)	η_p^2
GenCyber training	7.116	1	7.116	.341	.563	.009
Error (GenCyber training)	772.346	37	20.874			
Gender	501.603	1	501.603	5.280	.027*	.125
Error (Gender)	3515.346	37	95.009			
GenCyber training * Gender	15.372	1.000	15.372	.736	.396	.020
Error (GenCyber training * Gender)	772.346	37	20.874			

Note: SS=Sum of Square, MS=Mean Square, Result of 2 (GenCyber training : PreGenCyber, PostGenCyber) × 2 (Gender:Female, Male)

Figure. 2 presents the plot of the comparisons of estimated mean differences in future career interest ratings in the STEM fields by gender and GenCyber training. In general, GenCyber training has mixed impact on gender-based future career interest. Specifically, the plot in Figure. 2. shows that the GenCyber training program improves future STEM-related career interest for males. However, this pattern did not hold true for females.

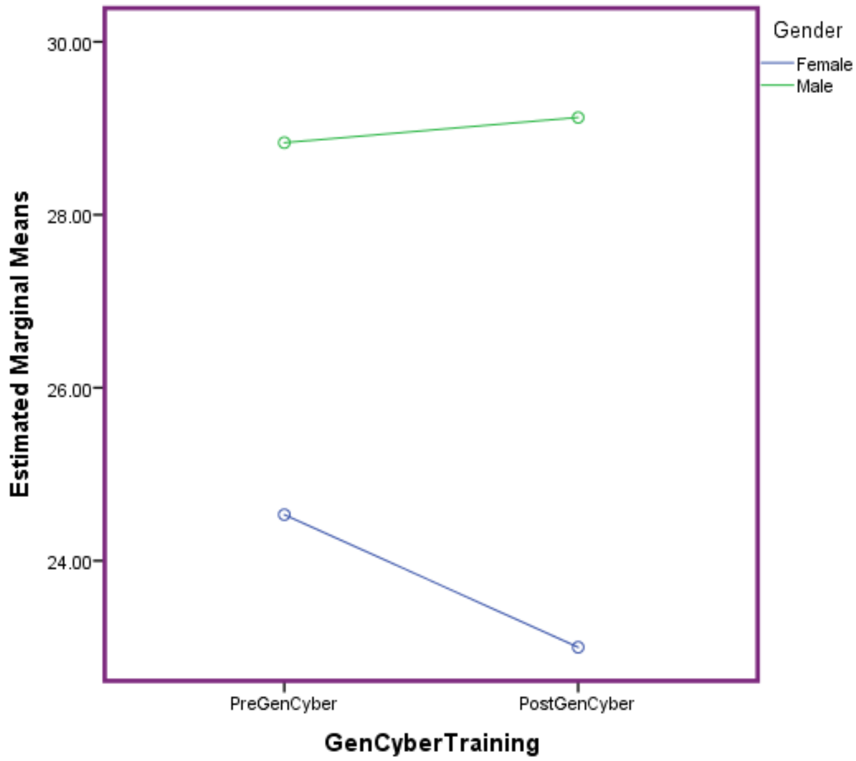


Figure 2 Estimated means for future STEM career interest of GenCyber training for each gender category.

2. Impact of GenCyber training on future Computer Science and Cybersecurity career interest: data analysis and result

The descriptive statistics of the influence of GenCyber training on future career interest in computer science and related fields is provided in Table 2-1. It can be seen from Table 2-1that the overall mean for future career interest in computer science before and after training for both females and males are 4.0, 4.87, 5.12, and 5.38 respectively. This shows that the training increases interest for the female gender group. Further analysis was needed to determine the statistical significance of the means. Early results were consistent with our prediction. Additional data analysis was required to investigate the statistical significance of each of the conditions. As stated, a two-way mixed ANOVA was used to analyse the statistical significance of the impact of GenCyber training on future computer science and cybersecurity career interest.

Table 2-1. Descriptive Statistics of Computer Science career interest rate

GenCyber training	Gender	Mean	Std.Dev
PreGenCyber	Female	4.00	1.77
PreGenCyber	Male	5.12	1.19
PostGenCyber	Female	4.87	1.36
PostGenCyber	Male	5.38	1.06

Table 2-2 presents the analysis results for the effects of GenCyber training on future career interest in computer science and cybersecurity. The analysis provides interesting findings for computer science career interest ratings for GenCyber training and gender categories. In addition, the result showed that there was no statistically significant difference in the interaction between GenCyber training and gender, $F(1,37)=1.619$, $p>(.211)>0.05$. However, the result indicated that there were statistically significant difference of main effect for both GenCyber training, $F(1,37)=5.308$, p (0.027)<0.05, ηp2=.125, and gender category, $F(1,37)=5.185$, p (0.029)<0.05, ηp2=0.123. Overall, the results indicated that there was a statistical difference between gender and GenCyber training on the future career interest in computer science and cybersecurity. However, the effect of gender and GenCyber training differences accounted for 12.5% (for gender) and 12.3% (for GenCyber training), suggesting that the impacts of both gender and GenCyber training in future career interest in computer science and related fields are relatively minimal.

Table 2-2. Two-Way Mixed ANOVA for computer science and related fields career interest rate

Source of Variance	SS	df	MS	F	P(Sig)	η_p^2
GenCyber training	5.755	1	5.755	5.308	.027*	.125
Error (GenCyber training)	40.117	37	1.084			
Gender	12.313	1	12.313	5.185	.029*	.123
Error (Gender)	87.867	37	2.375			
GenCyber training * Gender	1.755	1	1.755	1.619	.211	.042
Error (GenCyber training * Gender)	40.117	37	1.084			

Note: SS=Sum of Square, MS=Mean Square, Result of 2 (GenCyber training: PreGenCyber, PostGenCyber) * 2 (Gender: Female, Male

Figure. 3 shows the plot for the comparisons of estimated mean differences in future career interest rate in computer science and related fields by gender and GenCyber training. Overall, the plot in Figure. 3 shows that for GenCyber training program improves the future computer science and cybersecurity-related career interest for both females and males. This pattern is consistent with our predictions.

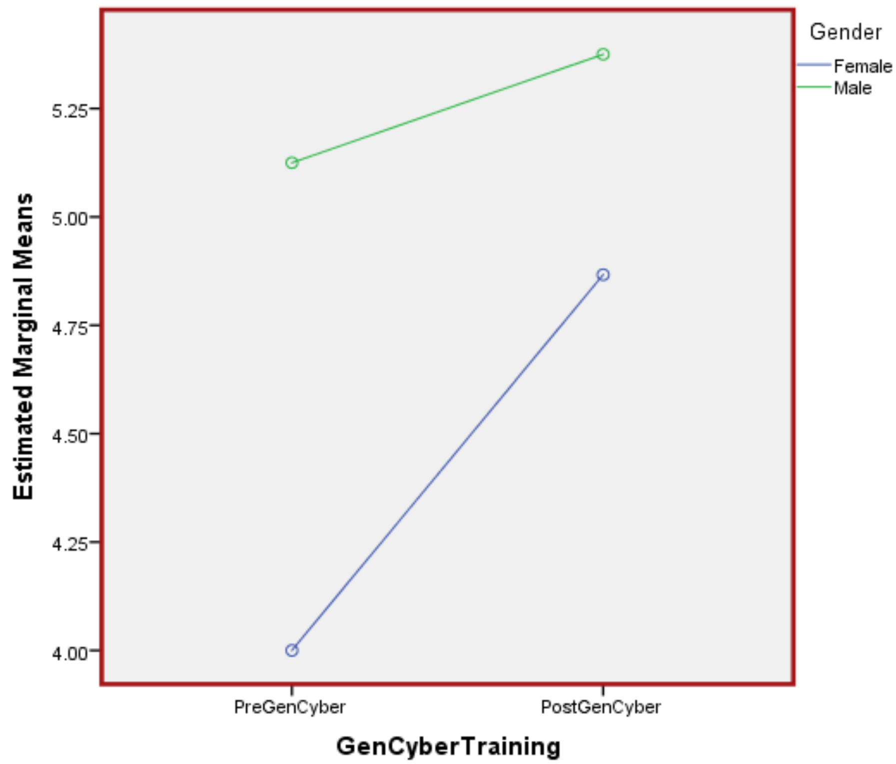


Figure 3. Estimated means for future computer science career interest of GenCyber training for each gender.

3. Impact of GenCyber training on future medical career interest: data analysis and result

The hypothesis for the impact of GenCyber training on future medical career interest is that GenCyber training enhances future career interest rate in medical and related fields and medical career interest rate as a dependent variable. The research is geared toward answering the following hypotheses: GenCyber training decreases future career interest rate in medical fields. The analysis result is presented in Table 3-1. It can be seen from Table 3-1 that the overall mean for future medical career interest for females and males before and after the training are 5.53, 5.13, 4.08, and 4.29 respectively. This shows that the training has a decrease in interest rate for female participants. However, the overall mean medical career interest for males before and after the training show that the training has a positive impact. Further analysis was needed to determine the statistical significance of the means. As indicated, a two-way mixed ANOVA was performed to compare the effect of early GenCyber training on future medical career interest of the participants.

Table 3-1. Descriptive Statistics of GenCyber of Medical career interest rate

GenCyber training	Gender	Mean	Std.Dev
PreGenCyber	Female	5.53	2.85
PreGenCyber	Male	4.08	1.77
PostGenCyber	Female	5.13	2.47
PostGenCyber	Male	4.29	2.20

The analysis results showed that there was no statistically significant differences in main effect of GenCyber training $F(1,37)=0.77$, $p>0.05$, and main effect of gender $F(1,37)=2.985$, $p>0.05$. Similarly, the interaction results indicated that there was no significant interaction effects gender and GenCyber training on future medical career interest impact, $F(1,37)=0.773$, $p>0.05$. Overall, the result indicated that there was no significant difference between GenCyber training and gender in influencing future medical career interest.

4. Impact of GenCyber training on future social career interest: data analysis and result

Likewise, the hypothesis of the impact of GenCyber training on future non-STEM (social) career interest is that GenCyber training enhances future career interest in non-STEM (social studies) and related fields. Similar to the medical career interest described in section 4.3, a preliminary analysis including Ethnicity as a between-subject factor revealed no significant main or interaction effects, so it was omitted from further consideration. Therefore, this study is adjusted as a 2 by 2, mixed factorial design ANOVA with two independent variable (IV). The first IV is gender, which includes two levels: female and male. The second IV (i.e., GenCyber training) contains two conditions: baseline (pre-training) and post-training. The dependent variable (DV) is future social science career interest level.

As can be seen from Table 4-1, the overall mean for future social career interest ratings for females and males before and after the training are 20.47, 18.80, 19.79, and 19.83 respectively. This shows that the training induces a decrease in interest rate for female participants. However, the overall mean social career interest for males before and after the training shows that the training has a positive impact for males. Further analysis was needed to determine the statistical significance of the means. As indicated, a two-way mixed ANOVA was performed to compare the effect of early GenCyber training on future social career interest of the participants. The result of the analysis is presented in Table 4-1.

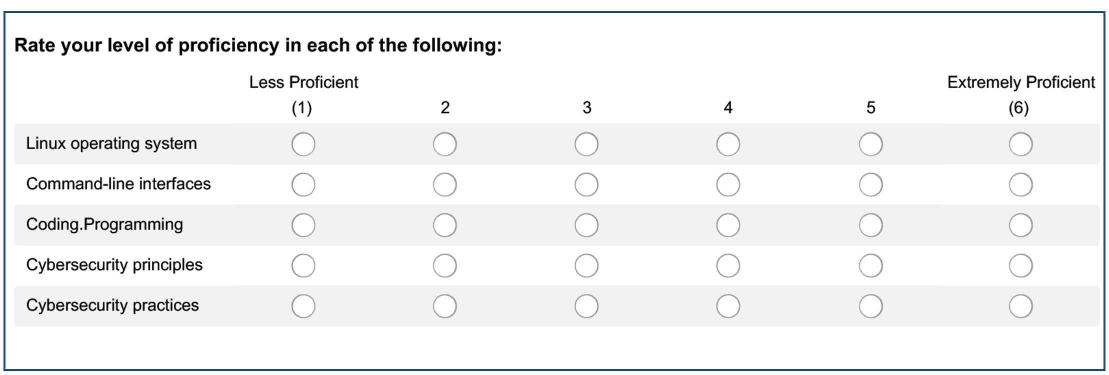
Table 4-1. Descriptive Statistics GenCyber of Social Science career interest rate

GenCyber training	Gender	Mean	Std.Dev
PreGenCyber	Female	20.47	7.54
PreGenCyber	Male	19.79	4.60
PostGenCyber	Female	18.80	5.77
PostGenCyber	Male	19.83	5.95

The analysis results showed that there was no statistically significant difference in main effects for both gender and GenCyber training $F(1,37)=0.776$, $p>0.05$, and $F(1,37)=0.11$, $p>0.05$, respectively. Similarly, the interaction results indicated that there was no significant interaction effects among gender and GenCyber training on future social science career interest, $F(1,37)=0.11$, $p>0.05$. Overall, the result indicated that there was no significant difference between GenCyber training and gender in influencing the future social science career interest.

5. Impact of GenCyber training on self-rating proficiency in computer science and related fields: data analysis and result

The hypothesis of the impact of GenCyber training on self-rated proficiency in computer science and related fields is stated as, “GenCyber training improves self-rated proficiency in computer science and related fields.” Figure 4. presents the self-rating proficiency questionnaire. It includes various technical areas in computer science and cybersecurity fields. The descriptive statistics of the influence of GenCyber training on self-rated proficiency in computer science and related fields is provided in Table 5-1. It can be seen from Table 5-1 that the overall mean for proficiency in computer science and related fields for females and males before and after the training are 7.04, 21.88, 13.88, and 20.15 respectively. This shows that the training has an increase in self-rated proficiency ratings for both female and male gender categories. This is consistent with our prediction. Further analysis was needed to determine the statistical significance of the means. As indicated, a two-way mixed ANOVA was performed to compare the effect of early GenCyber training on participants’ self-rated proficiency. The result of the analysis is presented in Table 5-2. As presented in Table 5-2 shows that the mean user proficiency ratings for both gender categories follow a similar pattern in that GenCyber training enhances perceived proficiency in computer science and related fields for both females and males. Early result is consistent with our prediction. Further data analysis is required to find out the statistical significance of each of the conditions of GenCyber training and gender.


Rate your level of proficiency in each of the following:

	Less Proficient (1)	2	3	4	5	Extremely Proficient (6)
Linux operating system	○	○	○	○	○	○
Command-line interfaces	○	○	○	○	○	○
Coding.Programming	○	○	○	○	○	○
Cybersecurity principles	○	○	○	○	○	○
Cybersecurity practices	○	○	○	○	○	○

Figure 4. PreGenCyber and PostGenCyber training on computer perceived proficiency self-ratings

Table 5-2 presents the results for the effects of GenCyber training on future career interest in computer science and cybersecurity related fields. The analysis provides interesting findings for computer science career interest ratings for GenCyber training and gender categories. As shown in Table 5-2, the analysis results of the level of perceived proficiency in computer science and cybersecurity related fields indicated that there was no statistical interaction effect between GenCyber training and gender, F(1,37)= 1.380, p(.248)>0.05. However, the result indicated that there were statistically significant difference of main effect for both GenCyber training, F(1,37)= 111.776, p <0.0001, ηp2 =.751, and the gender category, F(1,37)= 12.048, p (0.001)<0.05, ηp2 =0.246. Moreover, with respect to GenCyber training, the effect size of 0.751, which means that the effect of GenCyber training difference accounted for 75.1% of the group-differences, suggesting that the impact of GenCyber training on self-rating proficiency in computer science and related fields is relatively high. However, with respect to gender, the effect size of 0.246, which means that the effect of gender difference accounted for 24.6% of the group-differences, suggesting that the impact of gender on self-rating proficiency in computer science and related fields is relatively low.

Table 5-1. Descriptive Statistics of computer science and related field self-rated proficiency

GenCyber training	Gender	Mean	Std.Dev
PreGenCyber	Female	7.40	3.70
PreGenCyber	Male	13.88	6.83
PostGenCyber	Female	21.88	4.72
PostGenCyber	Male	20.15	5.42

Table 5-2. Two-Way Mixed ANOVA for computer science self–rated proficiency

Source of Variance	SS	df	MS	F	P(Sig)	η_p^2
GenCyber training	1495.385	1	1495.385	111.776	.000	.751
Error (GenCyber training)	495.000	37.000	13.378			
Gender	553.396	1	553.396	12.048	.001	.246
Error (Gender)	1699.450	37	45.931			
GenCyber training * Gender	18.462	1	18.462	1.380	.248	.036
Error (GenCyber training * Gender)	495.000	37.000	13.378			

Note: SS=Sum of Square, MS=Mean Square, Result of 2 (GenCyber training : PreGenCyber, PostGenCyber) × 2 (Gender: Female,

Figure. 5 shows the plot of the comparisons of estimated mean differences in self-rated proficiency in computer science and related fields by gender and GenCyber training. In general, the plot in Figure. 5 shows that for GenCyber training program improves the perceived computer science skills of the participants. This pattern is consistent with our hypothesis.

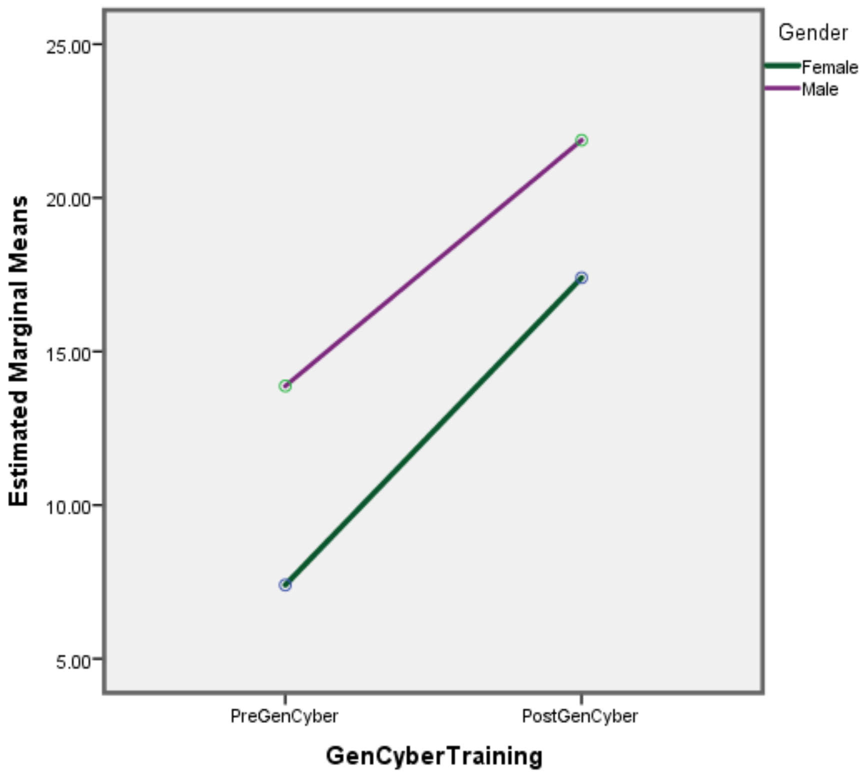


Figure 5. Estimated means for self-rated computer/cyber proficiency of GenCyber training for each gender category

DISCUSSION & CONCLUSION

Recently, most developing and developed countries have realized that STEM is tomorrow's most demand-driven career field for sustainable global economic growth. From the US perspective, various governmental and non-governmental agencies are working hand-in hand to prepare students for careers in STEM-related fields. As an important initiative, the GenCyber program was established as a framework to inspire and prepare young US citizens in an effort to fill the critical shortage of current and future experts in the constantly evolving field of cybersecurity. This research aims to evaluate the influence of the GenCyber training program on students' future career interest in STEM, cybersecurity, computer science, & related fields. In addition, recent studies indicate that although women represent roughly half of the entire US population, they represent less than a quarter (24%) of the nation's STEM workforce (Department of Commerce, August 2011). Given this disproportionately low number of females participating in STEM-related fields, it is the responsibility of society to encourage and provide the necessary assistance to citizens of all genders and ethnic origins to pursue STEM education. Therefore, this study also attempted to evaluate the GenCyber program's impact on minimizing the gender disparity between high school boys and girls in their future career interests.

The research identified the following two questions: first, would participating in the GenCyber summer program impact K-12 students' interest in future STEM careers? And second, would participating in the GenCyber summer program minimize students' gender bias toward future

STEM careers? As part of this study, forty high school rising sophomores to rising seniors were recruited to participate in the two-week residential National Cyber Warrior Academy (NCWA). While the ten first cybersecurity principles are conceptual ideas designed as the foundation of the GenCyber curriculum, students were given more practical, engaging, and intensive technical training in computer science and cybersecurity. Therefore, this study attempted to measure the effect of hands-on technical training on self-rated proficiency in the areas of computer science and cybersecurity. The experimental data was collected using pre- and post-training surveys. The analysis examined different categories via mixed factorial design.

A preliminary analysis including ethnicity as a between subject factor revealed no significant main or interaction effects, so it was omitted from further consideration. Consequently, this study was adjusted as five 2 by 2 repeated measures factorial design ANOVAs. Participants were asked to self-rate their interest in STEM, non-STEM, computer science and cybersecurity on a scale from 1 (less interested) to 6 (extremely interested). In addition, participants were asked to self-rate their computer science and cybersecurity skills before and after the GenCyber training program on a scale from 1 (less proficient) to 6 (extremely proficient).

Major findings were: a) GenCyber training program improves the future STEM-related career interest for males. However, the overall mean STEM career interest for females before and after the training shows that the training has a negative impact for females. This is not consistent with our prediction. b) The GenCyber training program improves the future computer science and cybersecurity related career interest for both females and males. Specifically, the findings indicated that GenCyber training improved career interest specifically in the area of computer science and cybersecurity for both genders. c) The training improved their self-rated proficiency in computer science and related fields for both females and males. d) Finally, the analyses indicated that there was no significant difference across GenCyber training and gender in influencing participants' future career interest in non-STEM careers. This tends to show that the GenCyber training didn't make a significant impact in deterring interest in non-STEM fields.

Overall, the results of this research suggest that GenCyber training could reasonably improve students' interest, skills, and proficiency in the field of computer science and cybersecurity, as well as their perceived efficacy in these areas of study. Given the increasing prevalence of cyber-threats in various governmental and non-governmental organizations, and the resultant economic, strategic, and security challenges to our society, the federal and private sectors seek large numbers of qualified cyber professionals with the requisite knowledge, skills, and abilities to protect the nation. Most importantly, this effort is particularly helpful to protect the nation from some of our most sophisticated adversaries and to safeguard our sensitive political and technological data, our financial and business systems, and other critical infrastructure. As a whole, expanding cybersecurity education by reaching out to primary and secondary school systems to simulate interest in computer science and cybersecurity-related fields, through initiatives such as the GenCyber program, will help fulfill the increasing demand for a greater cyber workforce. While the program has a positive impact on two significant factors in future career choices, tremendous effort is required to reach to the expected results in a one-camp-at-a-time approach. It is equally important to note that more significant changes are needed in both academic and societal culture in the US to attract more girls into STEM-related fields to minimize the gender gap. As part of future-work, we are planning to reach out to the participants and investigate if attending GenCyber has a long-lasting impact.

REFERENCES

Agogino, A. (2007, April). Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering. In APS April Meeting Abstracts (Vol. 1, p. 6001).

Bandura, A. (1986). Social foundations of thought and action: A social cognitive theory. Prentice-Hall, Inc.

Bloom, B. S. (1956). Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain. New York: David McKay Co Inc.

Boni, A. A., Weingart, L. R., Evenson, S. (2009). Innovation in an academic setting: Designing and leading a business through market-focused, interdisciplinary teams. *Academy of Management Learning & Education 8*(3), 407-417.

Certified Ethical Hackers, EC-Council (2016). Retrieved from https://www.eccouncil.org/programs/certified-ethical-hacker-ceh/#

Chen, X. (2013). STEM Attrition: College Students' Paths Into and Out of STEM Fields (NCES 2014-001). National Center for Education Statistics, Institute of Education Sciences, U.S.Department of Education. Washington, DC.

Christensen, R., Knezek, G., & Tyler-Wood, T. (2015). Gender differences in high school student dispositions toward science, technology, engineering, and mathematics careers. *Journal of Computers in Mathematics and Science Teaching, 34*(4), 395-408.

Clark Blickenstaff*, J. (2005). Women and science careers: leaky pipeline or gender filter?. *Gender and education, 17*(4), 369-386.

Coley, C., & Coley, T. (2010). Retention and student success: Staying on track with early intervention strategies. Sungard Higher Education.

Crissey, S. (2009). Educational attainment in the United States: 2007 (No. P20-560). US Census Bureau. Retrieved from http://www.census.gov/prod/2009pubs/p20-560.pdf

Dark M., & McNair B. (2015). *GenCyber Evaluation Report,*

Department of Commerce, (2011, August). Women in STEM: A Gender Gap to Innovation.

Erceg-Hurn, D. M., & Mirosevich, V. M. (2008). Modern robust statistical methods: an easy way to maximize the accuracy and power of your research. *American Psychologist 63*(7), 591.

Executive Office Of The President Office Of Management And Budget Washington, D.C. 20503 (2016, July 12) Retrieved from https://www.whitehouse.gov/sites/default/files/omb/memoranda/2016/m-16-15.pdf

Faul F., Erdfelder E., Lang A. G., & Buchner A. G (2007, May 1). *Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior research methods, 39*(2),175-91.

Fouad, N. A. (2007). Work and vocational psychology: Theory, research, and applications. *Annu. Rev. Psychol., 58*, 543-564.

Frome, P. M., & Eccles, J. S. (1998). Parents' influence on children's achievement-related perceptions. *Journal of personality and social psychology, 74*(2), 435.

Furnham, A., Reeves, E., & Budhani, S. (2002). Parents think their sons are brighter than their daughters: Sex differences in parental self-estimations and estimations of their children's multiple intelligences. *The Journal of genetic psychology, 163*(1), 24-39.

GenCyber Spring Meeting, San Francisco, May 14-15, 2016.

Gibbons, M. M. (2004). Prospective first-generation college students: Meeting their needs through social cognitive career theory. *Professional School Counseling, 8*, 1, 91-97.

Hyde, J. S., Lindberg, S. M., Linn, M. C., Ellis, A. B., & Williams, C. C. (2008). Gender similarities characterize math performance. *Science, 321*(5888), 494-495.

Lehming, R. F., Alt, M. N., Chen, X., Hall, L., Burton, L., Burrelli, J. S., & Moris, F. A. (2010). Science and Engineering Indicators 2010. NSB 10-01.National Science Foundation.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of vocational behavior, 45*(1), 79-122.

Lent, R. W., & Brown, S. D. (2006). On conceptualizing and assessing social cognitive constructs in career research: A measurement guide. *Journal of career assessment*, *14*(1), 12-35.

Li, Q. (1999). Teachers' beliefs and gender differences in mathematics: A review. *Educational Research*, *41*(1), 63-76.

López, X., Valenzuela, J., Nussbaum, M., & Tsai, C. C. (2015, December 15). Some recommendations for the reporting of quantitative studies. *Computers & Education, 91*(C),106-10.

Marzano, R. J. (2004). Building background knowledge for academic achievement: Research on what works in schools. Alexandria, VA: ASCD.

Mau, W. C. (2013). STEM career aspirations of high school students: A social-cognitive perspective, paper presented at the annual meeting of Association for Institutional Research, Long Beach, CA.

Mau, W. C. J. (2016). Characteristics of US Students That Pursued a STEM Major and Factors That Predicted Their Persistence in Degree Completion.

Moreno, R., & Mayer, R. (2007). Interactive multimodal learning environments: Special issue on interactive learning environments: Contemporary issues and trends. *Educational Psychology Review, 19*(3), 309-326. doi:10.1007/s10648-007-9047-2

National Academy of Sciences, Beyond Bias and Barriers: Fulfilling the Potential of Women in Academic Science and Engineering (National Academies Press, Washington, DC, 2006).

National Research Council. (2011). Successful K-12 STEM Education: Identifying Effective Approaches in Science, Technology, Engineering, and Mathematics

National Science Foundation (2012). Science and Engineering Indicators 2012. Arlington VA: National Science Foundation (NSB 12-01). Retrieved from http://www.nsf.gov/statistics/seind12/pdf/seind12.pdf

National Science Foundation (2013). Women, Minorities, and Persons with Disabilities in Science and Engineering, Arlington, VA: National Science Board (NSF 13-304). Retrieved from http://www.nsf.gov/statistic s/wmpd/2013/start.cfm

Program Director Guide (2016), Generation of Cyber Stars (GenCyber)

Sadler, P. M., Sonnert, G., Hazari, Z., & Tai, R. (2012). Stability and volatility of STEM career interest in high school: A gender study. *Science Education, 96*(3), 411-427.

U.S. Department of Education. (2006). A Test of Leadership: Charting the Future of U.S. Higher Education. Washington, DC.

Wang, M. T., Eccles, J. S., & Kenny, S. (2013). Not lack of ability but more choice individual and gender differences in choice of careers in science, technology, engineering, and mathematics. *Psychological Science*, *24*(5), 770-775.

This page has been intentionally left blank.

ERROR RATE IMPACTS ON DECISION EFFICACY: ACTIVITY-BASED COSTING SYSTEMS IN SMALL BUSINESS

Vincent J. Shea, St. John's University

Bobby E. Waldrup, Loyola University Maryland

Helen Xu, Loyola University Maryland

Steven Williamson, University of North Florida

ABSTRACT

Activity-based costing (ABC) systems research has extensively studied and theorized the benefits of implementing ABC in comparison to traditional costing systems. ABC systems can improve firm profitability by the use of sophisticated cost drivers, improved cost pool measurement, and through the ability of capturing the cause/effect relationship of product costing and firm pricing decisions. However, the accounting literature often lacks empirical evidence through firm level data. This paper fills this gap by testing the customer profitability differences between complex and simple ABC system using two-years of disaggregated, product cost information in the small to medium sized distributors in the fine paper service industry. The outcome shows that increases in measurement error for activity cost drivers and cost pools using a simplistic ABC system appear to demonstrate robustness in decision usefulness for these firms. Therefore, evidence is provided that a complex system may not outweigh the time and cost required to implement a successful system in smaller firms.

Keywords: Activity-based costing, measurement error, decision efficacy

INTRODUCTION

ABC Sophistication

The tenets of activity-based costing (ABC) as a vehicle for attaching indirect costs to cost objects (such as products or customers) have been well-researched in the cost management literature as a premier method for producing accurate product costs in both the manufacturing and service industries (Babad & Balachandran, 1993; Englund & Gerdin, 2008; Maiga & Jacobs, 2008; Soin, Seal, & Cullen, 2002; Wegmann, 2009). ABC has now become one of the most prolifically researched and employed methodology of cost allocation over a traditional allocation method [see for instance Cooper (1988a; 1988b; 1989a; 1989b) for an example of this discussion].

Much of the prior literature in product costing has centered on the two issues of comparing ABC systems to traditional costing systems, and to comparing the efficacy of greater sophistication in the design and use of ABC systems. In comparing the effectiveness of ABC systems to traditional costing ones, Brierley (2008) broadly delineates these systems as being either sophisticated (ABC) or unsophisticated (traditional). While the entire literature does not make this specific delineation, it is a widely accepted distinction (Al-Omiri & Drury, 2007; Bjørnenak, 1997; Brown, Booth, & Giacobbe, 2004; Joshi, 1998; Schoute, 2009) . Later papers in this research stream have criticized this distinction as being either too narrow (Al-Omiri & Drury, 2007; Drury & Tayles, 2006) or being an incompatible comparison (Dugdale & Jones, 1997).

In comparing the efficacy of the greater sophistication afforded by the use of ABC systems, several pieces of research have considered alternative or escalating forms of sophistication (also sporadically referred to as *complexity*) in the design or use of those systems (Abernethy, Lillis, Brownell, & Carter, 2001; Al-Omiri & Drury, 2007; Brierley, 2008; Drury & Tayles, 2005).

The main objective of any costing system is to provide both relevant and timely information to managers (Babad & Balachandran, 1993), though it is dependent upon the accuracy of the resulting allocations (Labro & Vanhoucke, 2007). Several studies have indicated that even modest distortions of product costs can be linked to inaccurate decision making (Drury & Tayles, 1994). Errors in product costs have been found to result from time-driven estimates (Cardinaels & Labro, 2008), the level of heterogeneity (Gupta, 1993), and the interaction among these various errors (Labro & Vanhoucke, 2007).

The level of decision usefulness for users of ABC systems depends upon their ability to both understand and contextualize the output for actual cost based decisions (Briers, Chow, Hwang, & Luckett, 1999; Drake, Haka, & Ravenscroft, 1999; Gupta & King, 1997; Waller, Shapiro, & Sevcik, 1999). Several experiments have found mediating factors on decision usefulness such as prior cost accounting knowledge (Cardinaels, 2008), asymmetric information (Drake & Haka, 2008), and market feedback effects (Gupta & King, 1997).

The purpose of this paper is to link the analytical findings of cost system efficacy with the real world needs for decision useful information for managers using these systems. One gap in the literature is a lack of empirical studies based upon actual firm data. This study attempts to fill this gap by attempting to test the robustness of theoretical and experimental postulates using actual firm data. While the results suggest an advantage of using a simple ABC model over a traditional allocation, subsequent reductions in decision usefulness appear to be relatively small for large variation in measurement error. These results are relevant for designing optimal levels of costing sophistication for decision makers.

Small v. Large Business Models

Successful implementation of ABC systems in small businesses can yield a wealth of benefits, both tactical and strategic. The use of ABC can help identify which products are being sold at a profit and which ones at a loss. This information is useful to management in developing marketing and pricing strategies. According to Baxendale (2001), products and services that produce high profits should be pushed more than those being produced at a lower profit or loss. Additionally, management can use the information on the unprofitable products to pursue longer-term goals of making those products profitable through continuous focused process improvements. It is important to note that unprofitable products should not just be eliminated, as this will shift the associated fixed costs to the other products (Emerson, 2016). It is better to either improve or replace the product. Hall and McPeak suggest that the use of ABC "alleviates managers' concerns regarding the accuracy of cost allocations, the cause-effect relationship between allocations and resources consumed, the timeliness of cost/profit information, and the capability to update systems" (2011, p. 12). With this information, management can make better decisions regarding finances, operations and strategy. This includes, but is not limited to, decisions about product mix, budgets, pricing, special orders, product development, outsourcing, marketing and process improvements. Ultimately, this will increase the company's competitiveness in the market (Rundora, Ziemerink, & Oberholzer, 2013). Jänkälä and Silvola (2012) note a very important characteristic of the benefits received by ABC- their lagging effects. According to their study, "the effects of ABC may not be visible in financial performance immediately after adoption, and it may take even several years before any improvements in financial performance are achieved" (Jänkälä & Silvola, 2012, p. 517). The wealth of information that ABC provides is well worth the justification of the one-time cost of implementation (Bharara & Lee, 1996).

Although prior studies suggest that smaller businesses actually have an advantage in implementing ABC systems in that they are more flexible due to their small size and simple organizational structure (Jänkälä & Silvola, 2012), when implementing ABC, it is important to realize that small businesses are not simply scaled down versions of larger companies, but that they are unique and require different methods of implementation (Needy, Nachtmann, Roztocki, Warner, & Bidanda, 2003, p. 6). One important characteristic of many small businesses is a high proportion of fixed versus variable costs (Needy et al., 2003). "The high ratio of fixed to variable cost combined with variation in sales and cash flows restricts small manufacturers to limited financial freedom" (Needy et al., 2003, p. 7). Implementation of ABC systems allows small businesses to carefully consider the implications of the high proportion of fixed versus variable costs and to better align the performance drivers into the model in order to avoid the unintended consequences (Emerson, 2016). Another unique characteristic of many small businesses is their skewed customer distribution when a few primary customers generate a significant amount of sales (Needy et al., 2003). Accurate product costing through ABC will help prevent business owners from giving into customers trying to take advantage of this situation by demanding low prices.

An additional major difference in the implementation of ABC is the complexity of the system (Bharara & Lee, 1996). When designing a costing system, managers should avoid using "an inappropriate number of activities (usually too many) and unnecessarily complex systems" (Needy et al., 2003, p. 7). In addition, data availability and the need for updating should be taken into consideration (Woutersa & Stechera, 2017). Roztocki, Porter, Thomas, and Needy agree that "standard implementation of ABC is too expensive and complex" for small businesses (2004, p. 19). They recommend using a flowchart to identify main activities and drivers. Needy et al. (2003) note that the simplicity of the system is not important just financially, but also for managers to easily understand ABC developments and results. It is better to start with a small number of activities and cost drivers and to later improve the system by either introducing more or splitting up existing activities, if necessary. As with any project or change in a company, management support and commitment is an essential aspect to the success of an implementation of an ABC system. Roztocki et al. also note that successful implementation "requires organizational changes, employee acceptance, investment in software and hardware, (and) equipment for data collection" (2004, p. 26). Hall and McPeak list factors that influence the success of implementation as "organizational readiness, financial impact, workflow productivity and overall business environment" (2011, p. 17).

Due to an increasingly globalized and competitive business environment, it is no longer suitable for small businesses to overlook the concept of activity-based costing. According to Jänkälä and Silvola, "the use of ABC is related to the managerial needs of the organizational life cycle stages rather than firm size only" (2012, p. 500). If the managerial and organizational needs require ABC, small business owners should make the long-term investment, providing a base for the company's future development and an avenue for the generation of financial benefits over time through improved tactical and strategic decision-making. Bharara and Lee suggest that the "most important factor for competitiveness, profitability and success of a company, big and small, is the control over their processes" (1996, p. 1128). Implementation of an ABC system provides accurate product costing in order to attain this control. Empirical evidence also demonstrates that the management accounting system with activity based costing implementation results in a better performance, even for enterprises operating in an uncertain and dynamic environment (Elhamma, 2015).

RESEARCH METHODOLOGY

Data Collection

The researchers conducted a case study with a group of U.S. based volunteer distributors in the fine paper industry to determine the feasibility of upgrading their costing systems from a traditional allocation methodology towards activity based costing standards. The companies fall into the Standard Industrial Classification (SIC code) 5111 Printing and Writing Paper. The distributors averaged sales volumes of five to seven hundred million dollars annually, and on average warehoused approximately 2,500 separate inventory items.

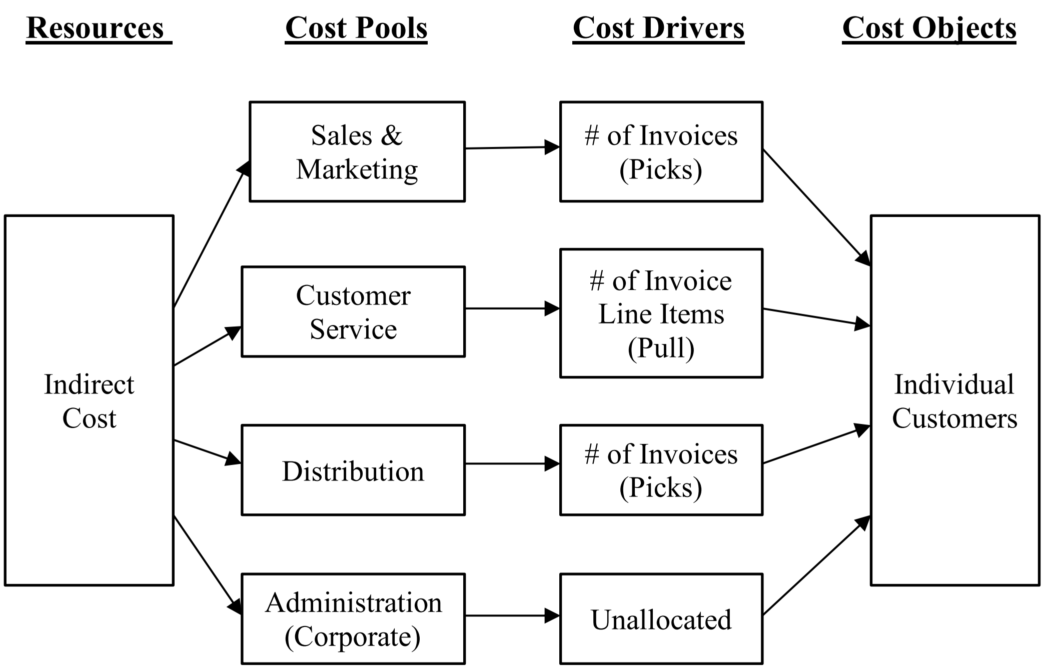
Product level cost data were collected on 100% of the sales volumes of the firms in an attempt to accurately assign costs on a customer level with the end goal to assess individual customer profitability. The data were collected across a two year period ending with fiscal year 2009 for each firm, and the cost-to-serve was attached to 100% of the distributor customers. With an average 14,000 customers each, the distributors' customers drove, on average, sales of $38,000 and generated allocated overhead of $4,000. These figures have been rounded and blended across years and between volunteer distributors. They are presented for the purpose of contextualization only. Figure 1 illustrates the contribution income statement approach used to calculate an individual customer's profitability when expenses are grouped into cost pools and then allocated to customers as cost objects.

Figure 1. Customer Profitability Computation

	Customer X	Customer Y	Customer Z	etc…	Total
Sales	xxx	xxx	xxx	xxx	xxx
Less: Commission	xxx	xxx	xxx	xxx	xxx
Net Sales	xxx	xxx	xxx	xxx	xxx
Less: COGS	xxx	xxx	xxx	xxx	xxx
Gross Margin	xxx	xxx	xxx	xxx	xxx
Less: Allocated Overhead	xxx	xxx	xxx	xxx	xxx
Sales & Market Costs	xxx	xxx	xxx	xxx	xxx
Customer Service Costs	xxx	xxx	xxx	xxx	xxx
Distribution Costs	xxx	xxx	xxx	xxx	xxx
Administration Costs	xxx	xxx	xxx	xxx	xxx
Net Income	xxx	xxx	xxx	xxx	xxx

Figure 2 illustrates graphically the typical ABC allocation approach used in this study. This indicates what cost pools were specified in the study, and what drivers were assigned in which to distribute those pools the the study's cost objects (individual customers).

Figure 2. Activity-Based Costing Allocation Approach



Prior to this study, the subject firms used a traditional, firm-wide overhead allocation process when determining product profitability, and secondarily, customer profitability. One example of this methodology was to simply allocate overhead costs to a given customer based upon a percentage of sales attachment process. For example, if customer A generated xx% of sales, then xx% of corporate-wide overhead (including sales & marketing, customer service, distribution, and administration) would be allocated to customer A. Table 1 provides descriptive statistics representing customer profitability, in dollars.

Table 1.
Descriptive Statistics of Customer Profitability

Mean	\$ 1,370.20
Median	\$ (66.91)
Quartile 1	\$ (443.11)
Quartile 3	\$ 224.78
Standard Deviation	\$ 16,070.48
Observations	N = 13,563

As can be seen from Table 1, while the mean customer drives profitability of about \$1,370 each, there is a large standard deviation. The median customer generated a net loss to the companies. This reaffirms the literature indicating that 20% of customers drive 80% of profitability (i.e. the 80/20 rule of profitability).

As an outcome of the ABC study, three activity cost pools representing distribution, customer service, and marketing were constructed (a forth pool called Administration was collected but could not be allocated on any rational cost driver other than percentage of sales). Each pool was assigned a cost driver based on the output of the initial case study using typical methods including interviews, researcher observation, and statistical analysis. Once overhead costs were allocated to products and added to direct costs, the profitability of customers was calculated using the following model:

$$\begin{aligned} &Customer\ Profitability \\ &\qquad = Sales - Commission - COGS - Allocated\ Distribution\ Costs \\ &\qquad - Allocated\ Customer\ Service\ Costs - Allocated\ Marketing\ Costs \\ &\qquad - Adminstrative\ Costs \end{aligned}$$

Where sales is the total sales for an individual customer. COGS is the total cost of goods sold for an individual customer. “Distribution” is the distribution costs associated for the same customer. Distribution costs can include costs to ship product from the firm to the customer, costs associated directly from the manufacturer to the customer, and costs of the customer picking up the goods at the firm’s warehouse including picking, packing, and shipping. Customer service costs include all telephone, in person, or onsite customer interaction. Marketing costs are associated with the sales staff and all communication to complete the sale.

Initial Data Analysis

As an initial exercise, the subject firms' customers were ranked ordinally by profitability as calculated utilizing the traditional overhead allocation based upon sales volume. Second, utilizing the aforementioned ABC customer profitability model, the firms' customers were again ranked ordinally using this more sophisticated calculation of customer profitability.

Cost pool construction

Cost pools were constructed entirely through the researchers' interviews with corporate process owners and their related expertise in the area of costs pools. Financial statement traditional expenses were constructed by process owners (corporate CFO's for instance), and through the interview process these income statement expenses were then stratified into four cost pools. For instance, "Distribution" costs were aggregated to include all costs of receiving, off-loading, warehousing, picking, packing, and shipping costs.

Cost driver construction

Cost drivers were similarly first identified through a cross-comparison of the literature and interviews with corporate management. As an example, *Customer Service* costs were determined to be most predictive through the measurement of inventory *line item pulls*. A simple weighting model was constructed assuming that, based on corporate experience, smaller customers require more service time and effort than larger customers per pull, given that larger customers have greater ordering sophistication, dedicated salespersons, and greater predictability in order patterns. Customers were quartiled based upon size (as measured by sales), and customer activities (line item pulls) were weighted with a simple 2, 1.5, 1.5, or 1. For example, a customer's activities in the lowest size quartile were weighted twice as burdensome as a customer categorized in the highest size quartile for the same activity level to reflect a more realistic view of the differential time-and-motion between the two types of customers for the same type of activity.

Customer profitability estimation

Customers' profitability was estimated twice. The first ranking was computed just as currently done by the subject firms in the traditional costing model using traditional financial statement expenses. The customers were then allocated costs of each cost pool as illustrated in figures 1 & 2, and were again ordinally ranked from highest to lowest.

To determine if there is a significant difference in customer profitability ranking between the traditional overhead allocation and the ABC customer profitability model, a Wilcoxon Signed Rank Test was used to compare between the two samples. With a sample of 13,563 observations we observe a student's t of 75.6711 (p value = <.0001) and signed rank of 3,489.5 (p value = <.0001). These results (untabulated) lend evidence that there is significant difference in the apparent profitability of the customers as measured by the differential rankings.

Research Questions

In the tradition of Datar and Gupta (1994) this study sets out to determine the level of specification error in both cost pools and activity drivers that is necessary to affect the level of decision usefulness. For purposes of this exercise we assume that significant changes in customer profitability as measured by the Wilcoxon Signed Rank Test are indicative significant changes in decision usefulness. Two research questions are investigated in this study.

RQ 1: What level of activity cost pool specification error is required to affect the decision usefulness of customer profitability ranking?

RQ 2: What level of activity driver specification error is required to affect the decision usefulness of customer profitability ranking?

RESULTS

Activity Cost Pool Error

In designing the ABC to allocate overhead to customers, the three cost pools of distribution, customer service, and sales and marketing represented 77%, 6%, 17%, respectively of overhead that can be allocated using cost drivers. A sensitivity analysis was conducted to estimate the amount of specification error that could be introduced into the formation of the activity cost pools without affecting the decision usefulness of the customer profitability information. Error was introduced into the model at levels of 1%, 2%, 5%, 10%, and 20% of misallocation between pools. An example of introducing specification error into the pools involved the researchers randomly selecting one cost pool, (e.g. distribution), and reducing the cost pool's overhead dollar amount by 1% and increasing the other two pools (e.g. customer service and sales and marketing) by a corresponding dollar amount. After this error was introduced customers were again ranked on profitability and compared to the original ABC rankings. As shown in table 2, using a two-tailed test, there is no significant change in profitability rankings until the level of error approaches 10%, and does not actually become significant until exceeding that error level and approaching 20%.

Table 2
Cost Pool Allocation Error
Wilcoxon Signed Rank Test

	Sign	P-Value
1%	-38.5	0.1295
2%	-117	0.2108
5%	-231	0.5428
10%	-1,527.5	0.0892
20%	-6,860.5	0.0117*

*Significant at the .05 level.

Activity Driver Error

Utilizing the original cost pool allocations the activity driver weightings were then introduced with escalating levels of specification error. Rather than using percentage changes as conducted in the activity pool portion of the experiment we increased (decreased) the level of complexity of the drivers and their weighting. Specifically the three levels of complexity are:

(1) Simple – a single minor change in the calculated activity weighting such as a shift from a variably weighted, time driven variable to an equally weighted variable which affects only one cost pool,
(2) Moderate – two relatively minor changes in the calculated activated weightings, and
(3) Complex – three or more changes in the calculated activated weightings sufficient to affect the allocation of all three cost pools.

As shown in Table 3, a simple specification error in activity driver weightings did not produce a considerable effect on the customer profitability rankings. However, in both the moderate and complex categories a significant effect was found in these rankings.

Table 3
Cost Driver Allocation Error
Wilcoxon Signed Rank Test

	Sign	**P-Value**
Simple	-4,978.5	0.5075
Moderate	1,758	0.008*
Complex	94,671	<.0001*

*Significant at the .05 level.

This result implies that model misspecification error may be more senstitive to cost driver allocation errors than those introduced into the cost pools.

Discussion and Analysis

The methodology employed was meant to take real small to mid-sized organizational data and determine the extent that ABC model misspecification error would cause changes in decision efficacy. Small organizations tend to avoid using ABC modelling due to their relative lack of financial sophistication and talent pool. The results of this paper lend evidence to two outcomes: (1) ABC analysis does appear to differentially and positively affect decision usefulness over traditional costing when using data from small to mid-sized organizations. (2) Even if a user in this size organization does misspecify either cost pools or cost activity drivers, the resulting decision usefulness of using ABC modeling still has a robustness that exceeds traditional costing until misspecification errors become egregious.

CONCLUSION

This study addressed the two research questions to determine the extent of error that must be introduced into an ABC system in order to differentiate the effect of decision usefulness. As shown in Table 2, there is little evidence that errors imposed into cost pools differentially affect decision usefulness until those errors approach a ten percent deviation from the optimal cost model. Similarly, Table 3 illustrates that specification errors in cost drivers do not influence decision making until at least two or more cost pools are affected.
Interpretation of these results may be limited due to the number of firms tested in this study, as well as size, industry and nature of the firms. However, we believe these results mark a good first step at empirically validating theoretical and experimental research in this field. Future work in this area would benefit from extended firm and industry data.

REFERENCES

Abernethy, M. A., Lillis, A. M., Brownell, P., & Carter, P. (2001). Product diversity and costing design: Field study evidence. *Management Accounting Research, 12*(3), 261-280.

Al-Omiri, M., & Drury, C. (2007). A survey of factors influencing the choice of product costing systems in UK organizations. *Management Accounting Research, 18*(4), 399-424.

Babad, Y. M., & Balachandran, B. V. (1993). Cost driver optimization in activity-based costing. *The Accounting Review, 68*(3), 563-575.

Baxendale, S. J. (2001). Activity-based costing for the small business: A primer. *Business Horizons, 44*(1), 61-68.

Bharara, A. A., & Lee, C. Y. (1996). Implementation of an activity-based costing system in a small manufacturing company. *International Journal of Production Research, 34*(4), 1109-1130.

Bjørnenak, T. (1997). Diffusion and accounting: the case of ABC in Norway. *Management Accounting Research, 8*(1), 3-17.

Brierley, J. A. (2008). Toward an understanding of the sophistication of product costing systems. *Journal of Management Accounting Research, 20* (supplement), 61-78.

Brierley, J. A. (2011). A comparison of the product costing practices of large and small- to medium-sized enterprises: A survey of British manufacturing firms. *International Journal of Management, 28*(4), 184-193.

Briers, M. L., Chow, C. W., Hwang, N. R., & Luckett, P. F. (1999). The effects of alternative types of feedback on product-related decision performance: A research note. *Journal of Management Accounting Researc*h, *11*, 75-92.

Brown, D. A., Booth, P., & Giacobbe, F. (2004). Technological and organizational influences on the adoption of activity-based costing in Australia. *Accounting and Finance, 44*(3), 329-356.

Cardinaels, E. (2008). The interplay between cost accounting knowledge and presentation formats in cost-based decision-making. *Accounting, Organizations & Society, 33*(6), 582-602.

Cardinaels, E., & Labro, E. (2008). On the Determinants of Measurement Error in Time-Driven Costing. *The Accounting Review, 83*(3), 735-756.

Cooper, R. (1988a). The rise of activity-based costing - Part one: What is an activity-based cost system? *Journal of Cost Management for the Manufacturing Industry, (Summer)*, 45-54.

Cooper, R. (1988b). The rise of activity-based costing - Part two: When do I need an activity-based cost system? *Journal of Cost Management for the Manufacturing Industry, (Fall)*, 41-58.

Cooper, R. (1989a). The rise of activity-based costing - Part three: How many cost drivers do you need, and how do you select them? *Journal of Cost Management for the manufacturing Industry, (Winter)*, 34-46.

Cooper, R. (1989b). The rise of activity-based costing - Part four: What do activity-based cost systems look like? *Journal of Cost Management for the Manufacturing Industry, (Spring),* 38-49.

Datar, S., & Gupta, M. (1994). Aggregation, specification and measurement errors in product costing. *The Accounting Review, 69*(4), 567-591.

Drake, A. R., & Haka, S. F. (2008). Does ABC Information Exacerbate Hold-Up Problems in Buyer-Supplier Negotiations? *The Accounting Review, 83*(1), 29-60.

Drake, A. R., Haka, S. F., & Ravenscroft, S. P. (1999). Cost system and incentive structure effects on innovation, efficiency and profitability in teams. *The Accounting Review, 74*(3), 23-345.

Drury, C., & Tayles, M. (1994). Product costing in UK manufacturing organizations. *European The Accounting Review, 3*(3), 443-469.

Drury, C., & Tayles, M. (2005). Explicating the design of overhead absorption procedures in UK organisations. *British Accounting Review, 37*(1), 47-84.

Drury, C., & Tayles, M. (2006). Profitability analysis in UK organizations: An exploratory study. *British Accounting Review, 38*(4), 405-425.

Dugdale, D., & Jones, C. T. (1997). How many companies use ABC for stock valuation? A comment on lnncs and Mitchell's questionnaire findings. *Management Accounting Reseach, 8*(2), 233-240.

Emerson, M. (2016, April). Rethinking Activity Based Costing, *Electrical Wholesaling*.

Elhamma, A. (2015). The Relationship between Activity Based Costing,Perceived Environmental Uncertainty and Global Performance. *International Journal of Management, Accounting and Economics, 2*(1), 73-90.

Englund, H., & Gerdin, J. (2008). Transferring knowledge across sub-genres of the ABC implementation literature. *Management Accounting Research, 19*(2), 149-162.

Gupta, M. (1993). Heterogeneity Issues in Aggregated Costing Systems. *Journal of Management Accounting Research, 5*, 180-212.

Gupta, M., & King, R. R. (1997). An experimental investigation of the effect of cost information and feedback on product cost decisions. *Contemporary Accounting Research, 14*(1), 99-127.

Hall, O. P., & McPeak, C. J. (2011). Are SMEs ready for ABC? *Journal Of Accounting & Finance, 11*(4), 11-22.

Jänkälä, S., & Silvola, H. (2012). Lagging effects of the use of activity-based costing on the financial performance of small firms. *Journal of Small Business Management, 50*(3), 498-523.

Joshi, P. L. (1998). An exploratory study of activity based costing practices and benefits in large size manufacturing companies in India. *Accounting and Business Review, 5*(1), 65-93.

Labro, E., & Vanhoucke, M. (2007). A Simulation Analysis of Interactions among Errors in Costing Systems. *The Accounting Review, 82*(4), 939-962.

Maiga, A. S., & Jacobs, F. A. (2008). Extent of ABC Use and Its Consequence. *Contemporary Accounting Research, 25*(2), 533-566.

Needy, K., Nachtmann, H., Roztocki, N., Warner, R., & Bidanda, B. (2003). Implementing activity-based costing systems in small manufacturing firms: A field study. *Engineering Management Journal, 15*(1), 3-10.

Roztocki, N., Porter, J., Thomas, R. M., & Needy, K. (2004). A procedure for smooth implementation of activity-based costing in small companies. *Engineering Management Journal, 16*(4), 19-27.

Rundora, R., Ziemerink, T., & Oberholzer, M. (2013). Activity-based costing in small manufacturing firms: South African study. *Journal Of Applied Business Research, 29*(2), 485-498.

Schoute, M. (2009). The relationship between cost system complexity, purposes of use, and cost system effectiveness. *British Accounting Review, 41*(4), 208-226.

Soin, K., Seal, W., & Cullen, J. (2002). ABC and organizational change: an institutional perspective. *Management Accounting Research, 13*(2), 249-271.

Waller, W. S., Shapiro, B., & Sevcik, G. (1999). Do cost-based pricing biases persist in laboratory markets? *Accounting, Organizations and Society, 24*(8), 717-739.

Wegmann, G. (2009). The Activity-Based Costing Method: Development and Applications. *ICFAI Journal of Accounting Research, 8*(1), 7-22.

Woutersa, M., & Stechera, J. (2017). Development of real-time product cost measurement: A case study in a medium-sized manufacturing company. *International Journal of Production Economics, 183*, Part A, 235-244

This page has been intentionally left blank.

ARTIFICIAL INTELLIGENCE TERMINOLOGY CAN BE MISLEADING: A FRAMEWORK FOR RATIONALIZING THE DISCOURSE

Kenneth R. Walsh, University of New Orleans

Sathiadev Mahesh, University of New Orleans

ABSTRACT

Steven Hawkins and Elon Musk have both commented on the potential terror that could occur as machines develop true artificial intelligence powers. However, Wasserman shows that much machine learning is little different from age old statistical analysis, although supercharged by the latest computer technology. Terms such as artificial intelligence, machine learning, and deep learning can evoke emotions in the general public and in the political arena, inconsistent with the true state of the art. In this paper we debate whether a machine can truly learn and conclude that the more useful question is about the manner in which business practice and the legal environment permit a machine to operate autonomously within the decision context. In answering the latter question we develop a stage model of machine learning systems based on the decision level of the system governed autonomously by machine algorithms. The model provides a useful framework for discussion, understanding, and governance of machine learning systems and reduces the hyperbole that can follow loaded terms such as learning and intelligence.

Keywords: artificial intelligence, machine learning, scientific framework

INTRODUCTION

Machine learning (ML) systems play a role in many facets of our daily lives. Not only do they direct us to preferred sites through search engines and match faces in our photos, but they also playing a role in health care diagnostics, insurance pricing, law enforcement, and employment decisions. Unlike the in-your-face approach of social news feeds which have aroused serious discussion at many levels of the community and government, most business applications of ML have replaced human decision makers in making many decisions that have serious consequences, with little press coverage and even less public attention. The consequences of machine learning algorithms can be lifesaving or life threatening, depending on the situation. Carr (2014) in his book, "The Glass Cage", describes the problems of highly automated environments which strip the human decision maker of any real work, reducing human learning and human alertness, creating cascading problems when the human is suddenly charged with making a decision. Diakopoulos (2016) notes that many news articles are written by machine learning algorithms and are often well written, but sometimes have single word errors that completely change the meaning of the article.

At the same time, the methods used in particular decisions may be opaque, giving little understanding to users about how a decision or classification was made (Burrell, 2016). Yarkoni and Westfall (2017) make the argument that in the field of experimental psychology, the emphasis on explanation, has reduced the predictive power of some machine learning techniques. A "black box" could fit data perfectly, while a well-explained model would only partially explain the data. While the causes of those predictions may lead to better explanation in the future, since knowledge is almost always built on top of current knowledge, at the present time, an explained model may provide a less useful, and less commercially valuable, predictive technique. Businesses focus on the commercial value of the ML model, and a good black box predictor which provides the best stock trade or optimal inventory policy is more suitable to the business than a well explained model with less short-term predictive power. Burrell (2016) identified three types of opacity stemming from the secrecy of the organization, the complexity of the algorithm, and interpretability by human decision makers. The secrecy of the organization is the desire to protect proprietary corporate methods, either for profit or political advantage. Statistics is a challenging field and most managers have only a limited knowledge of basic, often century old-methods. Only a few experts even understand complex models used in ML and even fewer understand the computer code used to process the data with the algorithm. Many ML tools yield black box models which are opaque to even the sophisticated user. The lack of interpretability by human decision makers renders the human decision makers useless and irrelevant in the process, leading to a surrender to the decision of the system. With such levels of opacity, it is more difficult for those affected by such systems to understand or dispute decision made by the systems. Garfinkel, Mathews, Shapiro, and Smith (2017) and Knight (2017) argue for the need for algorithmic transparency and accountability. The Defense Advanced Research Projects Agency has recognized the problem and initiated a program to study the discipline of explainable artificial intelligence (Gunning, 2016).

If machines are learning, in the manner of humans, then to what extent can people understand what they have learned? Terms such as artificial intelligence and machine learning may give the impression that these creatures are autonomous learners, however, many such systems fall short of this hyperbole. This paper opens with a debate as to whether or not machines can learn followed by a classification of machine learning levels, differentiated by the extent to which the computer systems are operating autonomously. Diakopoulos (2016) identified five broad categories of information that should be disclosed about machine learning systems. These areas will be used to inform our categorization. By classifying machine learning systems that have different components and levels of learning, more specificity can be given as to what aspects are technically opaque and what information could be provided to reduce their opacity when accountability is needed. The implication of the categories follows.

Machine Learning is not Learning

"Statistics is the science of learning from data. Machine Learning (ML) is the science of learning from data" (Wasserman, 2013, p. 2). Wasserman clearly pointed out the similarities between statistical analysis and machine learning and concluded they were largely the same. However, few would describe statistical analysis as learning. It would be more accurate to describe statistical analysis as a mathematical technique that humans use to identify patterns in data. A data-driven decision maker uses these patterns in decision making. We would therefore argue that machine learning is not about a machine learning either, but rather merely a different implementation of statistical analysis. This leads to two important implications. First, many predictions of what machine learning can achieve are overblown and ill-informed and much that has been written about the implications of machine learning in various industries may not be realistic. Second, public policy is often shaped by word choice, and since learning represents one of the higher achievements of human beings, the use of the word "learning" has caused an over-reaction among policy makers on the ability of machines to completely replace human workers. An interesting corollary to this argument is whether we could take steps to actually get machines to learn as humans do.

An argument can be made that the user interface of machine learning gives the impression of learning because of how the output is used in practice. If a user of Google types a search term and receives a useful result, or verbally states a question to a home automation device, and receives a useful answer, they tend to be amazed by how well the system works. They may over-estimate the "smartness" of the machine to have learned so much and come up with the answer. Consider a different scenario where human statisticians work with scientists or business leaders to use data collected by either a targeted survey instrument or by surveillance equipment, follow the rules of accepted statistical analysis, and interpret and present the results. Often, a satisfied client will likely attribute the good work to the statistician rather than the computer software and hardware used to obtain the results. This attribution of skill to the human statistician happens primarily because the underlying computer system is not well understood by the user and the human is the interface through which the results are made available. Many user-focused applications using ML have a user-friendly interface, often designed with much thought given to user-stories, the script describing user interaction with the system. Often, this user interface (UI) is designed to be user-friendly to even an unsophisticated user. The major difference between the first and second scenarios is that in the first the machine directly communicates with the user, while in the second a human interprets and delivers the result to the users.

To take the Google search example further, the consumer may go to google the next day and re-enter the same query. If the results are better than the day before, the consumer may observe that the system is getting better. This improvement may have happened without the intervention of any human and so it may give the appearance that the machine has learned. However, the machine has just run the same algorithms or statistical analysis on a new dataset with more data. Updating a dataset and rerunning an analysis is a useful task for computers, but we would argue that automated data gathering is not enough to be called learning. Consider the scenario of the human statistician who returns the next day with one more day's data in the data set and provides even better results than the previous day. One could argue that the biggest leap in the new ML systems is that they have finessed the toolset to continually add new data and re-analyze larger data sets in a convenient manner, eliminating the delays and manual effort to add and re-analyze data. From this perspective, ML is just an automated tool within a normal statistical update processes.

A simple example of when statistical analysis and machine learning yield the same result would be a simple classification problem. Whereas some may call it supervised learning and others may call it classification, the results are mathematically equivalent. Wasserman (2013) describes how researchers in both statistics and machine learning are working on an extension of supervised learning call semi-supervised inference. When using ML to classify data, for example, in a classification of photos by background location, data needs to be initially classified, often by human classifiers. This manually classified data provides the training set for the ML tool. In supervised learning, a large training set is used to train the ML tool, and the creation of this training set is a major expense. In semi-supervised inference, the set of manually classified data is analyzed along with a much larger unclassified dataset, using techniques of likelihood maximization to obtain better results than with only the manually classified data set. The tools are part of the repertoire of a sophisticated statistician; the difference is that ML automates the process of data gathering and analysis, and also provides the results in a decision impelling format, and may even complete the decision process automatically. An interesting extension, but just a more in-depth example of how analysis performed on a static data set, can be updated with a new data in a dynamic dataset, presenting the appearance of learning over time.

Much machine learning research is done on improving the algorithm and tackling certain problems that can arise. Often machine learning approaches use a large number of independent variables. Since so many input variables are used, the approach may uncover independent variables that are surprising. This is again the consequence of larger dataset and not real learning. A downside to using a large number of independent variables can be over fitting. Over fitting occurs when a model very accurately represents the training set but does not predict new datasets well because it is too highly influenced by idiosyncrasies of the training data set. Regularization techniques are used to prevent such over fitting. For example, the dropout regularization technique is a method of dropping random nodes when training a neural network to prevent large scale neural networks from over fitting (Srivastava, Hinton, Krizhevsky, Sutskever, & Salakhutdinov, 2014).

The term deep learning may conjure up the idea that something is being learned. Many stunning improvements in natural language recognition and translation have been attributed to deep learning. However deep learning simply refers to a more complex version of machine learning or neural networks, often implemented as several hidden layers in a neural network design. The benefits have accrued from better hardware, massive datasets, and algorithmic improvements in the weights and processing within the neural network (Monroe, 2017). "Deep learning allows computational models that are composed of multiple processing layers to learn representations of data with multiple levels of abstraction" (LeCun, Bengio, & Hinton, 2015, p. 436). By this definition, deep learning is an extension of the concept of the neural network that has been afforded business and science through increases in computing power. Deep learning neural networks can have many layers and many nodes. A principle advantage of handling many nodes and layers is for machine learning systems to be developed without the often time consuming and largely manual stage of feature engineering. With many forms of ML, real world data is not in the right format for the input to the system as features. Feature engineering creates features out of real world data. Deep learning neural networks can use a large number of input nodes to access real world data directly while using many layers to effectively extract features on its own. As such, deep learning may not strictly be considered learning, but it does automate a time consuming human tasks in many ML projects. Chainer is an open source framework for deep learning models (Tokui, Oono, Hido, & Clayton, 2015). Blocks and Fuel are software applications that help deal with complexity of deep neural networks having many layers and nodes as well as large data sets (Van Merriënboer, 2015).

Machine Learning is Learning

What is learning? Consider the case of a hunter-gatherer tribe twelve millennia ago with no knowledge of farming. They find that a particular variety of grass provides seeds which are edible. The learning process in this case involves trial and error, where the tribe tries out many different seeds and finds one that is edible and available for easy foraging. They then determine visible patterns in this type of grass and use them to identify the grass which yields this type of seed. Later, maybe in order to avoid having to forage for this plant, or by accident because they carry the seed home frequently, they find that the plant grows in their vicinity and can be cultivated to enhance yields. There are many learning steps in the process of moving from a hunter-gatherer group to a farming community. A trial and error process leads to understanding how to detect one type of grass which yields wheat from another that does not provide a healthy yield. Communication between members of the tribe speeds up the learning process.

One major aspect of learning is the learning of a society rather than the learning of an individual. When viewing the historical arc of research, we often jump from one discovery to another, often highlighting the "winners" in the intellectual arms race. In reality however, the words of Isaac Newton (1675, p. 1) hold true, "If I have seen further, it is by standing on the shoulders of giants." Learning typically develops in a community where the successes and more often the failures of others are used to work toward a new concept. Machines can learn almost instantaneously from other machines, since the most successful neural network's weights, identified clusters, or detected rules from a decision forest can be immediately replicated across multiple machines. Just as viral concepts rise and flow rapidly through internetworked social media, multiple networked machines can quickly learn from one another. Hence, ML can be learning if the machines learn from one another rather than working in isolation.

Consider adaptive exponential smoothing (McClelland, 1971), a computer-based time series forecasting approach developed in the early 1970s. Exponential smoothing developed in the 1950s used a smoothing parameter and the value of this parameter was selected to provide the best fit to the data. However, if the parameter selected at one point in time was no longer optimal, the forecast error grew and the predictions became unreliable. A correction mechanism using computer code to re-calculate the optimal value of a smoothing parameter when errors exceeded a threshold made the forecasts much better. Can this be called a learning system? It is trivial by the standards of today's machine learning tools, but it clearly encapsulates the learning process. The approach uses a simple model, determines the error, and if the cumulative error exceeds a threshold, it launches a corrective action. We may not term this intelligent in a machine, and of course the knowledge is limited to one type of time series forecasting, but we have the essence of a learning system built into the program. Machine learning is a form of learning if it has a built-in feedback mechanism which monitors the outcome of its actions and takes appropriate corrective actions to its module when the outcomes are deemed undesirable.

Intelligence is not merely in computational complexity but also lies in the interface. If we make an entry in a cell formula in a spreadsheet and the intelligent formula advisor signals an error, it is clearly more intelligent than a formula entry mechanism which blindly accepts the bad formula and merely fails to calculate a result. As the formula advisor in a workbook becomes more capable and provides a wider range of advice, it is perceived as smarter. However, if it repeats the same error statement repeatedly, users term it a "dumb machine", and seek human help. How do humans exhibit intelligence in this scenario? An intelligent human advisor, will not monotonously repeat the same error statement multiple times. After repeating the error statement, maybe a couple of times for emphasis, the error statement will be rephrased. Failure at this step in getting the user to enter the correct formula will open up a more detailed analysis of the error and better targeted advice to the user to help fix the problem. This is the adaptability expected of a human instructor who learns about the problem faced by the user after interaction. Machine learning is a form of learning when the system does not repeat itself interminably, and monitors a log of its actions to ensure that alternate approaches are tried out. If the entire logic of the action sequence is pre-programmed, then it does not constitute learning. The system should have basic rules of behavior about what constitutes rational behavior and what constitutes stupidity.

An aspect of learning is to get better over time, by the repeated exercise of cognition. One part of this learning is increased hand to eye coordination gained by repetitive action. While the motor abilities of robots have been continually improved, they still face challenges opening a door, climbing a flight of stairs, or grasping a soft item without crushing it. Another part of this learning is the development of mental maps of situations in the brain of the decision maker. These mental maps help the decision maker make sense of data that deviates from the norm, such as a scene which does not fit the known profile of a route. While human drivers often use landmarks to assist in navigation, automated navigation tools often use a strictly algorithmic, turn by turn signaling approach which requires constant attention to, and complete dependence on the system's instructions. Human drivers may either surrender to the system or undertake the intense cognitive effort to maintain a mental map of their location, in addition to following the turn-by-turn instructions. An approach to make the system more intelligent incorporates landmarks in the directions and creates and maintains maps of these landmarks in a manner similar to human drivers (Zhu & Karimi, 2015). Machines can learn if they can create and maintain multiple higher-level maps of the data. Regression software merely calculates the vector values of the data and does not understand the variables, and this applies to the most sophisticated statistical tools in ML toolsets. However, if the variable names are mapped to real-word objects in a map of images, their physical locations, and assigned names, and other ML tools connect the data to the objects, as well as to other prior data analysis, we can have real learning.

The human body is highly adaptable and the muscles become more adept at physical activities by training. The body and brain often re-allocate resources to meet the needs of the environment. Repeated mathematical training makes the student better at math while physical training on a wood-working tool makes the human a better carpenter. The way robots are designed today, they are incapable of transferring their capabilities to different parts to get better at what they do. A CPU which may be as good at math on the day it is fabricated in a chip plant as it will be years later, may also deteriorate or fail completely, but it will not improve. Hardcoded software has the same problem; it does not learn. The only part of the system getting better is that of pattern recognition systems which improve as they get more data, and are corrected when they make errors. However, they too do not learn anything new unless they place their results in a mental map that displays the connections between the data.

Rather than recognizing patterns by mere trial and error training of pixelated images, multi-modal neural networks generate pattern similarity scores and assign text to describe the image and its context. The combination these two modes, text descriptors of the object and its context and image patterns yield better recognition of objects than mere pattern similarity. In essence, this process not only recognizes the pattern in an image but also assigns text descriptors to the objects in the image and uses both to make a recognition. Stores of these mental maps have enhanced the performance of robots in many areas (Di Nuovo, De la Cruz, & Marocco, 2013).

When ML is applied to large data sets, the approach is not very different from traditional statistical analysis. Data is collected, cleaned, and then processed through dimensionality reduction, clustering, and regression to find patterns in the data which can be used to explain and predict human behavior. Of course, the availability of ML tools makes it easy to set up data acquisition and cleaning processes to repeatedly access and clean the data, and in fact to set up the entire statistical sequence to generate classifications of targeted customers or predictions of human behavior. For example, Microsoft's Azure ML toolset has models for Regression, Logistic Regression, Boosted Decision Trees, Random Decision Forest Algorithm, Support Vector Models, and Neural Networks, all readily available in drag and drop format. The availability of larger datasets ensures more frequent use of split data to test the model, and the availability of automated data processing tools supports the testing of data on multiple models to select and use the model providing the most effective classification. In addition, tools to monitor performance, can be linked back to the model to modify it when the outcomes are unsatisfactory.

What about the new tools available through the cloud which have learned to detect objects from videos, speech, or even emotions? These tools provide knowledge gained from a vast pool of training data that can be applied to the data available to a researcher. This is different from the statistical tools described earlier, which encoded knowledge developed to recognize patters in data, but did not include the knowledge gained from using these tools. To summarize, when we use factor analysis, or logistic regression, or even an ANN tool, we do not use any of the knowledge gained from prior use of the tool. SPSS's regression has been used millions of times, and it has not become any better because of prior use.

However, when a cloud-based voice recognition or object recognition tool is used, it becomes better and the dominant tools in the marketplace become superior to human classifiers. This improved ability of cloud based apps such as Google's Tensor Flows, Apple's Core ML or Amazon's Polly, or Microsoft's emotion detection API offer continually improving ability to identify objects in images, detects specific types of video and add labels to identify content as well as to detect changes in scenes and content, to detect speech and provide either a transcript from audio or generate speech from text, and to detect emotions from images and videos. This is where the real machine learning is taking place, i.e. where machines are becoming more capable than humans at certain tasks and available at next to no marginal cost. When individuals and businesses use this learning, and connect it to their in-house systems, we have machines with greater capability than many humans.

LEVELS OF MACHINE LEARNING

One way to better understand the extent to which machines are learning is to classify them by what steps in the process are automatically generated by the computer versus what part of the process involves decisions made by the machine learning analyst (MLA). For example, in a simple system, an MLA may choose the design of the neural network and define the number of layers and nodes at each layer as well as the aggregation system. In a more automated and hands-off system, the ML system will split the data, run different models and decided on whether or not to use a network and select the parameters of the network. Knowing what the computer is choosing autonomously helps understand how the term learning is being used and what strengths and weaknesses might occur in practice.

We propose a set of levels of increasing machine autonomy to classify machine learning systems. In general the levels are increasing orders of autonomy and would indicate a system with greater ability to learn.

A Level 1 ML system would be the most basic type of ML system. The MLA would choose a training dataset and type of model and run an algorithm to determine the model parameters. The algorithm may have been coded by the MLA or may have been provided in an ML library. At this level, ML is being used in much the same ways as it would be in traditional statistical analysis. The predictions based on new data may be presented to the end user either by the MLA or directly embedded in the computer system. In traditional statistical analysis, often the analyst presents the finding personally or in a written report to decision maker. In the ML community, the results may be embedded in a computer application.

A Level 2 ML system is a Level 1 system where the training dataset is automatically updated and the model parameters are recalculated. This may be one of the most common forms of systems colloquially referred to as learning. In such a system, the computer collects new observations which improve the training dataset. The model parameters are recalculated with each update to the training dataset or at a periodic interval. The methodology can allow a system to be deployed in practice even without adequate training data if sufficient data is expected to be forthcoming. This level can appear to an end user as learning because the system may initially make weak decisions and improve those decisions over time.

A Level 3 system is one in which the computer autonomously chooses the type of model. Many ML models exist ranging from simple regression to decision trees and multilayer neural networks. In a Level 3 system, the ML compares all models at its disposal and choose the model best suited to the data. In such a system, the MLA is responsible for providing the range of available models. A Level 3 system may or may not have a training dataset that is updated over time such as in Level 2. If such a dataset were updated, then a Level 3 system could choose different models over time as well as the parameters for such models.

A Level 4 system chooses its own input regularization method. Regularization is used to keep model from overfitting the training data and a regularization method is often selected by the MLA. In a Level 4 system, the computer chooses the regularization method and parameters from those at its disposal.

A Level 5 system conducts its own feature engineering based on raw real-world data. Deep learning systems usually fall into this category whereby raw data can be fed directly to the systems inputs.

A Level 6 system chooses its own feature set. Typically ML systems are given training data by the MLA. A Level 6 system may or may not have an initial dataset. The Level 6 system will search the data at its disposal on internal networks to the organization or the public Internet to find feature sets that are useful to the desired outcome. Over time, features may be added or dropped by the system.

In a Level 7 system, results of machine learning prediction are implemented without human oversight. For example, many ML systems provide recommendations to humans for decision making. However, one distinguishing characteristic of ML systems from previous statistical techniques is technical suitability for making embedded systems that can automatically act upon their results. In simple information retrieval tasks, no human stands between the computer and end user in displaying search results, however, in medical diagnosis, this is still an important step. If an artificial intelligence system were connected to an intervening drug delivery system administering pain medication, it may in fact do a better job than medical personnel while it may also have consequences for errors. Note that this characterization of a Level 7 system could be combined with other levels and does not strictly follow Level 6.

A Level 8 system chooses the outcome objective. For example, in most systems, the MLA choses and objective such find the modest relevant document of the shortest path. In a Level 8 system, the computer automatically choose an objective. In an organization, it might have the power to hire or fire people in a certain job description. A Level 8 system might choose not only how a self-driving car will get to its destination, but also what the destination should be. Table 1 summarizes the levels identified.

Table 1. Levels of Machine Learning

Level	Computer Learning	MLA Oversight
1	Choose model parameters	Choose Model, Choose training dataset
2	Update training dataset and update parameters	Choose Model, Choose training dataset
3	Choose Model	Provide Available Models, Choose training dataset
4	Choose input regularization	
5	Automated or no feature engineering	Provide raw data source
6	Find new training data	
7	Results acted upon without human intervention	
8	Set Objective of ML System	

AUDITING MACHINE LEARNING

Auditing of machine learning is critical when important decisions are made by a relatively autonomous computer system. Machine learning systems can be seen as both socially consequential and opaque (Burrell, 2016). The systems are certainly consequential as they can make decisions on employment, credit worthiness, and health, all vital to individual well-being. They are also consequential even in less threatening systems such as web advertising where it is still unclear to what extent a society can be influenced by machine learning controlled social media. The system are also opaque in that the user of the system and the target of such systems may have little insight into the black box that created the decision. They must simply accept or reject the results, but being at a loss for the argument behind the results, they may be in a position where they must accept the results. Further, if such results are challenged on a legal or ethical basis, it may be difficult to find or prove the flaw in the system. Content management systems may be one area where the consumer may demand algorithm transparency (Mittelstadt, 2016). In such a system, the consumer relies on system to filter the wide range of information sources available for news and opinion. The system is intentionally biased for the purpose of providing the person with the type of information they are interested in, but at the same time the consumer may be looking for an unbiased view of those issues of interest.

This combination of social consequence with the significance of those consequences expected to increase over time and the opacity of the decision models lead to a great need for a systematic audit procedure. In some contexts, legal issues may demand such audits, while in other contexts consumers may demand it. Such a systematic audit procedure can analyze a system and shed light on operations of different stages of processing. An audit check can be developed based on the level of autonomy of the ML system.

Auditing using such a classification scheme, however may have limited use if the audit is sought by a third party seeking to audit the organization. Since the business organization using the ML will not reveal the details of the model used for the protection of proprietary algorithms, business processes, and proprietary data, the organization may not share enough internal information for systems to be categorized and for the audit procedure appropriate for the category to be applied. As seen earlier, opacity in machine learning system can be divided into three levels, organizational opacity, technical opacity, and mismatches between human decision making and machine (Burrell, 2016). The classification most directly addresses technical opacity and gives a start at considering differences in human decision making and machine. Although, it does not give a solution to organizational opacity, it does provide a framework for what questions can and should be asked as well as a possible framework for what questions organization may be compelled to answer.

IMPLICATIONS FOR PUBLIC POLICY

If public policy experts are not aware of the similarities between machine learning and statistical analysis, they are likely to misunderstand the problem and propose solution that can be circumvented. Policy makers should be more focused on the improvements in the science of making inference from datasets. If the public or a company has access to an array of datasets, modern analysis techniques can be used to conduct a fine grained classification of individuals including buying power, purchase likelihood, emotional triggers, heath status, financial status, political affiliation or criminal participation.

Governmental decision making based on machine learning is likely to expand and can be regarded as having a positive potential if care is taken in the implementation (Coglianese and Lehr, 2017). Three aspects of machine learning can be seen as factors for worry about implementation, the complexity of the algorithm, the black box nature of the decision, and the automation of the decision process (Coglianese & Lehr, 2017). First, because of the self-learning property of ML, the computer derives algorithms that are not prescribed by computer analysts. Second, because the derived algorithm is defined with such a wide variety of variables and combinations it acts like a black box whereby the human decision maker would find it hard to tell why a decision was made. Finally, due to the fact that ML systems are often design to be embedded in other computer systems, their results can be acted upon with no human intervention. Debate on these proposed challenges of ML systems can enhanced by using the proposed framework. In such a framework we can separate sources of data available to the ML system and types of algorithms applied. For a given algorithm we could debate what parameters may be useful as audit points. And in the case of Level 7 systems where the output of the ML systems is applied autonomously, the benefits and risks can be weighed and mitigating systems could be proposed.

For example, if advanced machine learning techniques were applied in a health care institution then much may be predicted about its patients' future health conditions. Once such future health conditions are found, is it ethical to disclose or not disclose the findings? Further, it is the nature of machine learning analysis that many inputs and many outputs are simultaneously considered in what could be described as exploratory approach. If a machine learning study were conducted for the purpose of capacity planning for a hospital, it may be approved will little concern for ethical questions. However, the study may simultaneously predict which patients are likely to return because of new health concerns.

Insurance companies in many fields from health care to auto insurance use machine learning to classify customers into profitability categories. Such classification could lead to fine grained insurance pricing. The internal dataset of hospitals and insurance companies may be able to make significant predictions, possibly more accurate predictions are combined with publicly available data such as Facebook or public government records.

POSSIBILITIES FOR INCREASED LEARNING

The more the techniques of ML are classified and described by what they do autonomously, the less they look like the general public's idea of learning. However, future developments and the use of multiple layers of ML systems could change this.

Choosing Relevant Problems

Today, ML analysts choose the problem for which to apply ML techniques, but if ML systems begin to make this choice autonomously, new systems with goals not explicitly enumerated by humans could emerge. The assumption behind all research papers we have found to date is that a human decided upon the problem to be solved and selected or designed ML systems to solve this problem. Even if the system appeared to be somewhat of a black box with hidden complexity, ultimately, the human picked a goal and assessed the output. Particularly in the area of non-labeled classification, one could imagine a computer system that autonomously creates categories and possibly had a way to act upon such results.

Choosing Relevant Data

When humans craft problems for machine learning systems to solve, the humans choose relevant data sets. The machine learning algorithm is applied and a mathematical prediction model is calculated. One area that could bring machine learning closer to human learning would be an architecture where the machine chooses the relevant data sets. Zhang et al. (2016) show how their transformation, hybrid orthogonal projection and estimation tool, improves the performance of neural network training. In their article they describe the machine learning approach as having two stages, extraction and data modeling. Extraction, sometimes called feature engineering, is the process largely done by the human analyst which select which data, interactions, and transformation should be made on the raw data to best prepare it for machine learning. The importance of this process its labor intensity show why much human guidance is given to the machine learning algorithm. Zhang et al. point out that better neural network approaches, such as their own, allow the neural network to function effectively will little or no feature engineering. Raw data set can be fed to the neural network. As this science progresses, it does remove the human decision maker from an important step in the analysis process and may be yet another small step to independent learning.

CONCLUSION

The eight level ML scale was designed to focus debate on where and how ML systems derive their effectiveness and deliver either benefit or risk. ML system are so diverse that it not useful to say in general what they do and this classification can inform ethical and legal debates over what constitute and issue. It also highlights that while computers may be considered by some as running the world because of the vast number of decisions they make, in all cases there are humans with goals involved at some point in the process and a computer that can control humans is not likely to be developed in our life time.

The eight level categorization scale for ML systems was designed based empirical evidence of ML design. It is robust in that the introduction of newer deep learning methods do not change the scales but just fit within the framework. However, one can imagine that new systems may introduce new technique that could be consider as new categories. More work should be done investigating the latest techniques to document whether they are improved methods of techniques without altering the classification of the technique or in fact change the classification. Deep learning networks are a good example of where the neural network techniques have evolved into new areas. Further, one should be on the lookout where new techniques offer truly new categories. In any of these cases, though, the framework is useful tool for having a consistent debate about the benefits and implications of such new techniques.

The model is also useful at classification of new ML systems. It would be useful to know as new systems are implemented if they are refinements on an existing class of ML system or if they warrant a new class.

REFERENCES

Burrell, J., (2016, Januay - June). How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society*, 1–12.

Carr, N. (2014). *The glass cage: automation and us.* New York, NY, US: W W Norton & Co.

Coglianese, C., & Lehr, D. (2017). Regulating by robot: Administrative decision making in the machine-learning era. *Georgetown Law Journal, 105*, 1147.

Di Nuovo, A., De la Cruz, V., & Marocco, D., (2013). Special issue on artificial mental imagery in cognitive systems and robotics. *Adaptive Behavior*, *21*(4) 217 – 221.

Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of ACM*, *59*(2), 56-62.

Garfinkel, S., Matthews, J., Shapiro, S. S., & Smith, J. M. (2017). Toward algorithmic transparency and accountability. *Communications of The ACM*, *60*(9), 5.

Gunning, D. (2016). Explainable artificial intelligence (XAI). Defense Advanced Research Projects Agency, Retrieved from https://www.darpa.mil/program/explainable-artificial-intelligence.

Knight, W. (2017). The dark secret at the heart of AI. *MIT Technology Review, 120*(3), 54-65.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

McClelland, C. A. (1971). The management and analysis of international event data: A computerized system for monitoring and projecting event flows. WEIS Technical Report, USC, Los Angeles, CA.

Mittelstadt, B. (2016). Automation, algorithms, and politics: Auditing for transparency in content personalization systems. *International Journal of Communication, 10*, 12.

Monroe, D. (2017). Deep learning takes on translation. *Communications of the ACM*, *60*(6), 12-14.

Newton, I. (1675). Isaac Newton letter to Robert Hooke. Feb. 5, Retrieved from https://discover.hsp.org/Record/dc-9792/.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research, 15*, 1929-1958.

Tokui, S., Oono, K., Hido, S., & Clayton, J. (2015). Chainer: a next-generation open source framework for deep learning. Proceedings of workshop on machine learning systems in the twenty-ninth annual conference on neural information processing systems, 5.

Van Merriënboer, B., Bahdanau, D., Dumoulin, V., Serdyuk, D., Warde-Farley, D., Chorowski, J., & Bengio, Y. (2015). Blocks and fuel: Frameworks for deep learning. Cornell University Library, Retrieved from https://arxiv.org/abs/1506.00619.

Wasserman, L. (2013). Rise of the machines. Retrieved from http://www.stat.cmu.edu/~larry/Wasserman.pdf

Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives On Psychological Science, 12*(6), 1100-1122.

Zhang, S., Jiang, H., & Dai, L. (2016). Hybrid orthogonal projection and estimation (HOPE): A new framework to learn neural networks. *Journal of Machine Learning Research*, *17*, 1-33.

Zhu, R., & Karimi, H. A. (2015). Automatic selection of landmarks for navigation guidance. *Transactions In GIS, 19*(2), 247-261.

Blank inside back cover.

QUARTERLY REVIEW OF BUSINESS DISCIPLINES

VOLUME 5 NUMBER 1 MAY 2018

This issue is now available online at www.iabd.org.

A JOURNAL OF INTERNATIONAL ACADEMY OF BUSINESS DISCIPLINES
SPONSORED BY UNIVERSITY OF NORTH FLORIDA
ISSN 2334-0169 (print)
ISSN 2329-5163 (online)