SUMMER
2020

JOURNAL OF MANAGEMENT & ENGINEERING INTEGRATION

AIEMS
VOL 13
NO 1

ISSN: 1939-7984

JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Scope: The Journal of Management and Engineering Integration (JMEI) is a double-blind refereed journal dedicated to exploring the nexus of management and engineering issues of the day. JMEI publishes two issues per year, one in the Summer and another in Winter. The Journal's scope is to provide a forum where engineering and management professionals can share and exchange their ideas for the collaboration and integration of Management and Engineering research and publications. The journal will aim on targeting publications and research that emphasize the integrative nature of business, management, computers, and engineering within a global context.

Editorial Review Board

Reviewers

The Journal Editorial Team would like to thank the reviewers for their time and effort. The comments that we received were very constructive, detailed, and helped us to continue to produce a consistently top-quality journal. Your participation is very important in the success of providing a distinguished outlet for original valuable articles. Again, I would like to thank you all for your assistance in the review process. Below are the reviewers for the Summer 2020 issue.

Ron Barrett- Gonzalez, Ph.D.
Editor-in-Chief

Abdulaziz Abdulaziz	Jacob Cress	Robert Cattlet
Andrzej Gapinski	Jingyun Li	Roelof deVries
Ed Hernandez	Khalid Bachkar	Ronald Shehane
Ewa Rudnicka	Lauren Schumacher	Sandy Furterer
F. Michael Ayokanmbi	Mohammad T. Khasawneh	Scott Swain
Feng Zhou	Mohammed Ali	Sherisse Pierre
Frederick Augustine	Nabin Sapkota Narasimha	Sima Esfandiarpour
Gordon Arbogast	Nagaiah Palmer Frye	Stephen Frempong
Holger Mauch	Patric McNamee	Thilini Ariyachandra
Huitian Lu	Paul Nugent	Tristan Davison
Indra Gunawan	Paulus Wahjudi	Xiaochun Jiang
Isaac Heim	Rian Mehta	Xun Xu

Table of Contents

ANALYTICS, ACCOUNTING AND BIG DATA: ENHANCING ACCOUNTING EDUCATION

Fred K. Augustine, Jr.[1]
Joseph Woodside[2]
Monica Mendoza[3]
Valrie Chambers[4]
[1, 2] Department of Business Systems and Analytics, Stetson University
[3, 4] M.E. Rinker Sr. Institute of Tax and Accountancy, Stetson University
faugusti@stetson.edu; jmwoodsi@stetson.edu; monica.mendoza@stetson.edu; valrie.chambers@stetson.edu

Abstract

The role of Analytics and Big Data in business has been the subject of considerable research and discussion. Members of the accounting profession have recognized the importance of both and recognize that the emergence of these technologies require employees who possess a higher level of data analysis capabilities. Thus, industry is demanding that institutions of higher education place a greater emphasis on developing curricular elements, which will adequately address the needs of the ultimate consumers of these capabilities. This paper examines the data analysis skills relevant to accountants and how and where data analytics and big data have impacted undergraduate accounting education. This research examines the degree to which selected undergraduate Accounting programs have embraced Analytics and Big Data to the greatest extent and proposes a set of curricular elements the provide the appropriate level of learning within the current framework of undergraduate accounting education.

1. Introduction

As businesses change the way they operate, the demand for human resources also changes. One example of this is how the demand for "traditional business technology skill sets have fallen significantly" while demand for individuals who can deal with "solving specific business problems or exploiting specific business opportunities." (McClure, 2008) has increased. Thus, as the composition of the human resources demanded by industry changes, the requirements placed upon academic programs change in turn.

Economic events subsequent to the Crash of 2008 have caused business leaders and academics to re-examine the economics of their institutions and industries. This is especially true in terms of technology and resource utilization. For Academic leaders this had led to a need to reassess the relevance of information systems in the context of a university curriculum. These events, coming on the heels of transformational events such as the bursting of the dot.com bubble, the "solution" to the Y2K problem and the subsequent decline of the "tech stock" bull market, have led to a more resource scarce and efficiency centered environment for both business and academic institutions. As businesses evolve, educational institutions are forced to evolve their information systems programs. These programs, by their very nature are often at the forefront of such change. It is a given that the technologies that businesses use to fulfill their operational data infrastructure needs are and will continue to evolve at a rapid pace. But when this type of change is accompanied by changes of a disruptive nature (Bower and Christensen, 1995) at the level of the industry or even the economy, great pressure is placed on academic institutions to keep pace.

The demand for data and for data analyst (data scientists, business analysts, etc.) has been increasing at an ever-increasing rate. McKinsey Global Institute research shows that need for skilled Business Intelligence workers in the U.S. will exceed the available work force by 60 percent. (McKinsey and Company, 2018). According to the Most Promising Jobs Report 2018 released by LinkedIn, job openings for data scientists have increased by 45 percent year-on-year (Lewis, 2018). Becker's Health IT & CIO Report cites a Bloomberg analysis stating that "Data scientist represents the "hottest job" in the U.S., according to an analysis published in Bloomberg May 18. Job postings for data scientists rose 75 percent on Indeed.com since 2015. Job searches for data scientist roles have also increased, rising 65 percent during the past three years. Major companies also offer additional benefits and have streamlined the hiring processes to attract data scientists to their firms (Cohen, 2018)." This increasing demand is challenging industry and higher education to find data scientists, big data analysts and business intelligence workers. Because of the extraordinary events of the past decade the paradigm which has defined the relationship between higher education and industry has changed. Instead of merely requiring new employees with new skills that matched even newer technologies that organizations were employing to meet the competitive needs, we have seen a dramatic shift in the focus of these requirements. Much of this shift in focus came about as a result of the advent of analytics as a key element of competitive strategy (Davenport, 2006).

2. Background

Analytics or Business analytics can be defined as "the use of information technology, data, statistical analysis, quantitative methods, and mathematical or computer-based models to help managers gain improved insight into their business operations and to make better, fact-based decisions" (Evans, 2020) .The term was first popularized by Thomas Davenport in his seminal article "Competing on Analytics" where he describes the use of analytics techniques as not only being critical to an organization's ability to compete, but as a new form of competition. Davenport (2006) says that "Organizations are competing on analytics not just because they can—business today is awash in data and data crunchers—but also because they should be."

This "analytics movement" has been greatly enhanced by the emergence of "big data." The term "big data" refers to very large data sets that can be analyzed using statistical and computational methods in the hope of revealing trends and associations (cite). Big data has become an issue due to the convergence of several technological trends which have increased the volume, velocity and variety of data available to decision makers. Among these trends are the miniaturization of computing technology, the continuous evolution of storage technology and the increasing volumes of data available due to ubiquitous connectivity and social media. The data is an issue precisely because it is big. Computing systems today are generating 15 petabytes (define petabyte) of new information every day, which is eight times more than the combined information in all the libraries in the United States (cite). About 80% of the data generated every day is textual and unstructured data (Chiang, 2012). But, it is not just the volume, variety and velocity of the data sets to be managed that poses the challenge. It is the complexity faced by organizations in the management of this data. A report published by The Gartner Group (2011), maintains that: "The program management, technology and complexity of skills associated with the strategic use of business intelligence, analytics and project management increased dramatically as the scope of the initiative widens across multiple business processes(Chandler, 2011)."

This framework is depicted in Figure 1 and describes the problem faced by management, not only in the field of accounting but universally, as deriving from the increased complexity being forced upon organizations as they attempt to compete on analytics. Increased complexity, in turn, increases the needs of organizations, which in turn increases the desired skillset of employees. The most important of these extended skillsets are decision making, analytic abilities and the ability to keep and use information.

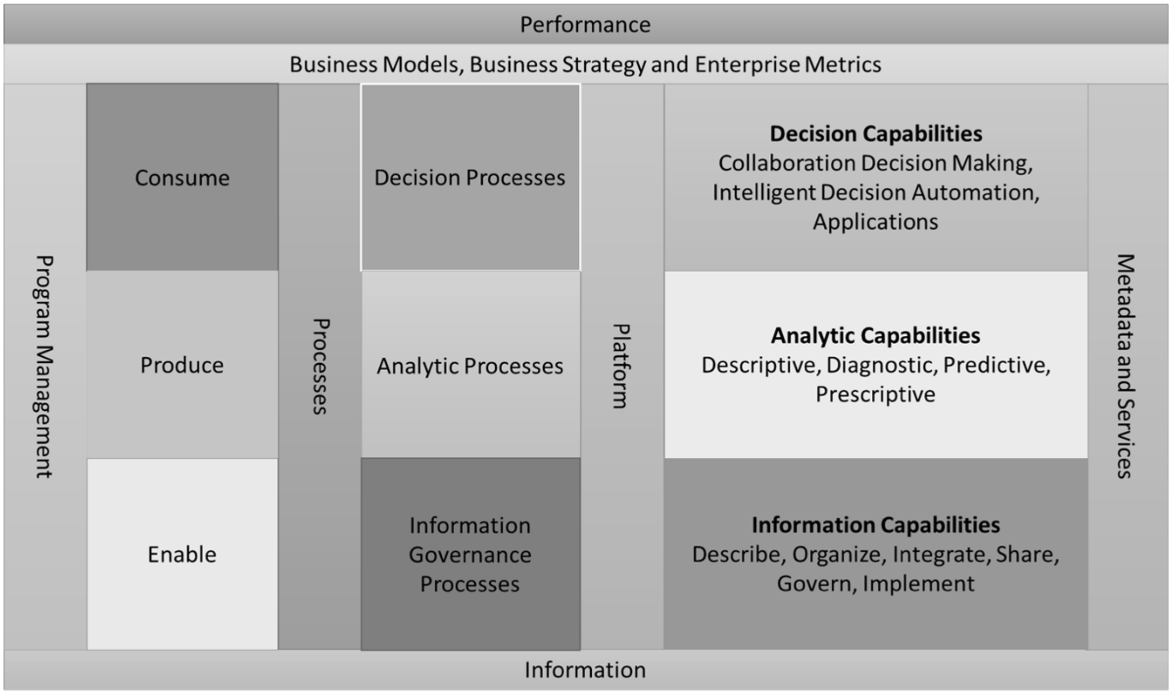


Figure 1. The Gartner Business Analytics Framework

An alternative view of the role of analytics in organizations is represented in Figure 2, Conway's Data Science Venn Diagram (Conway, 2013). This view is based on the diagram, developed by Drew Conway (CEO of Alluvium, a U.S. based Big Data and Analytics firm).

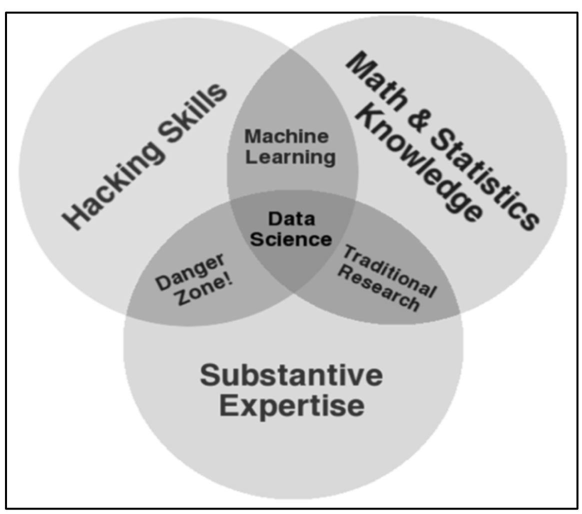


Figure 2. The Data Science Venn Diagram

This view proposes that hacking skills, math and statistics knowledge and substantive domain expertise possessed by employees define the analytics environment for an organization. The four areas of overlap define how data science can be applied in organizations. The area of overlap of all three defines what we know as data science. Note that Conway maintains that individuals possessing hacking skills and substantive (domain) expertise without requisite math and statistics knowledge represent a danger zone for organizations. Thus, by its nature, business analytics, while improving our ability to utilize data to gain insight and make decisions, also increases the complexity of our organization and its processes. This of course, increases the difficulties associated with decision-making and the management of organizational processes. It is because of this challenge that is critical that our educational institutions provide the best possible match between their curricula and the required analytics skill sets of those in business and industry.

3. Analytics, Accounting and Big Data

Accounting practitioners and academics have relatively recently recognized the impact of analytics and big data in organizations and the accounting profession. Griffin and Wright maintain that "The commentaries in this forum on big data confront one of our profession's most pressing challenges. How should we respond to Big Data? and further "Big Data and Analytics permeate almost all aspects of major companies' decision making and business strategies." (Griffin and Wright, 2015)

It has also been suggested that "Big Data and Business Analytics might eliminate potential discrepancies between U.S GAAP and international IFRS, thus accelerating movement toward a global accounting regime with fair value accounting as a key cornerstone" (Warren et al. 2015). And alternatively, "Presentation, aggregation and sampling are headed in the wrong direction … Aggregation and arbitrary allocations made on static, paper-based financial statements are artifacts of a bygone era of high transmission costs and slow data collection speeds" (Krahel and Titera, 2015). The clear message here is that Big Data and Analytics are exercising a disruptive effect on business organizations and consequently are poised to profoundly affect the accounting profession and the processes associated with accounting in the not too distant future.

4. Analytics and Big Data in The Accounting Curriculum

Corporate leaders are increasingly adopting accounting analytics tools to drive business decisions, which support budgeting, forecasting, planning, and now increasingly, revenue generation. Organizations are utilizing data and information to generate new customers, upsell to existing customers, and improve supply chain management. Accountants are well-positioned within an organization to leverage information into actionable business and financial insights, however in order to achieve this capability, analytical skillsets must be enhanced to prepare for a successful career (Woodside, 2019; PWC, 2015; Shimamo, 2013; Vickrey, 2013). This research develops the structure for an interdisciplinary curriculum on Accounting Analytics which affords students the opportunities to develop domain knowledge along with application of data analytics. As industry experiences rapid technological change, university curricula are required to be updated in order to be effective. Curriculum contents are further advanced and established with input from industry organizations that employ graduates of the programs.

Despite the corporate requirements for advanced accounting analytics, scholars have argued that accounting research has stagnated, resulting in a limited number of topics, using similar methodologies and theories. A literature analysis found that most accounting educating articles are non-empirical, relate to similar topics, and omit issues that are important to practice resulting in minimal impact to practice with an ever-widening gap (cite). When empirical methods are used survey research is prominent, with few studies using experimental methods (Woodside, et.al., 2019). Others argue that there has been a lack of new ideas, insights, and interdisciplinary perspectives within accounting research (Rebele and St. Pierre, 2015, Woodside, et.al., 2019).

Until recently, the slow rate of change in the accounting major body of knowledge has resulted in programs utilizing similar courses across a standard curriculum, creating an opportunity for research to be focused on career paths for students (Woodside, et.al., 2019). While many graduates are initially employed in public accounting, most are employed in the private sector over the long term. Information systems skillsets have been identified as fundamental to accounting and are incorporated into all career areas, however most accounting programs focus on superficial coverage of information systems beyond introductory courses. One argument is that accounting faculty are not provided detailed information systems area content coverage and ways to integrate within the accounting curriculum (Rebele and St. Pierre, 2015).

Recent developments by the AACSB Accounting Standard A5 requires learning to develop knowledge and integration of information systems within accounting, with output of data creation, data sharing, data analytics, data mining, data reporting, and data storage within and across organizations. The AACSB recommends an integrated, interdisciplinary curriculum that includes statistics, data management, analytics, and big data tools (AACSB International, 2018).

PricewaterhouseCoopers has developed recommendations for curriculum changes which include the following skills for undergraduate programs (PWC, 2015).

1. Learning of legacy technologies (e.g. Microsoft Excel and Access)
2. Understanding of structured and unstructured databases (e.g. SQL, MongoDB, Hadoop)
3. Obtaining and cleaning data
4. Introduction to data visualization (e.g. Tableau, SpotFire, Qlikview)
5. Univariate and multivariate regression, machine learning, and predictive tools
6. Early coverage of programming languages such as Python, Java, or R.

It is clear that the accounting profession and accounting firms have a need for employees who have significant skills, knowledge and expertise in the area of data analytics. What is not clear is the degree to which undergraduate accounting programs are meeting this demand.

5. Discussion

To attempt to answer this question we conducted an analysis of sample data collected on 35 universities with top undergraduate accounting programs to determine the degree to which analytics and big data topics have made their way into the accounting curriculum at the course level. Our definition of “top undergraduate accounting programs” is based on an institution’s inclusion in one of these three listings: Professional Accounting Report, Accounting Degree Review and CollegeChoice.net. Using these individual program rankings, a composite ranking was developed and can be found in Table 2. Data on program content was collected via the examination of the institutional web sites of the constituent universities and drilling down within these sites to obtain specific information about their undergraduate accounting degree programs. Each program was examined to determine the number and specificity of courses which relate to analytics and/or big data topics. Since the definition of analytics includes information technology, statistics and data, courses which featured any one of these three topics were selected as having some specific level of analytics/big data content. The level of specificity is based on how specific the analytics or big data course was to the accounting discipline. The levels of specificity are presented in Table 1. They range from Level 1 (the least specific to accounting) which consists of Basic or Introductory Statistic courses to Level 6, courses on Analytics and/or Big Data for Accountants having content which is the most accounting specific.

Table 1. Levels of Course Specificity

Level	Course Category
1	Basic or Introductory Statistics
2	Introductory IS/MIS
3	Accounting Information Systems
4	Generic Decision Analysis/Advanced Statistics
5	Generic Analytics/Big Data
6	Analytics/Big Data for Accountants

All courses which were determined to have some level analytics/big data content were categorized based on where they appear in the accounting curriculum. The four categories used are general education or business foundation courses, business core courses, courses which form part of an accounting concentration and courses which form the curriculum for an accounting major. These results are summarized in Table 2, Analytics and Big Data Content in Top Undergraduate Accounting Programs by Specificity to Accounting.

A measure of the degree to which the programs have integrated analytics and Big Data content into their curricula was developed based on the classification of courses shown in Table 2. The table shows the number of courses in each category which are required in the undergraduate Accounting degree. This measure of Big Data/Analytics content considers the number of courses with this content in the undergraduate major and awards points depending on the level of specificity with regard to Big Data/Analytics content. Courses which are generic such as Basic Statistics or Information Systems, Accounting Information Systems or Decision Analysis courses are assigned a value of one point. Generic courses focusing on Big Data and/or Analytics are assigned a value of two points, and courses on Big Data or Analytics specifically for accounts are assigned a three-point value. These point values are summed and ordered with the highest point values corresponding to the highest Analytics rank for the undergraduate accounting program.

6. Analysis and Conclusions

An examination of the results of the analysis of the analytics and/or Big Data content of the top thirty-five undergraduate accounting programs shows that most of the programs feature courses which deal with these topics generically (i.e. Introductory Statistics or Information Systems). In fact, as of (date) only six of the programs (approximately 17%) include courses in Big Data or Analytics/Big Data for Accountants. Only two programs (approximately 6%) include a course that addresses Analytics or Big Data specifically for Accountants. Of the top thirty-five programs, only one featured five analytics or Big Data related courses. The great majority of these programs (approximately 66%) included either two or three analytics or Big Data related courses. In addition, a correlation analysis finds that the correlation between program rank and analytics content rank is only .228. This indicates that Big Data and Analytics do not seem to be prioritized in the top undergraduate accounting programs. Indeed, the average program ranking of the top five programs rated in terms of their Big Data/Analytics content is 16.8 while the Big Data/Analytics content ranking of the top five accounting programs is 6.4. It is apparent that, in spite of the high level of interest in Big Data and Analytics within the accounting profession, that this high level of interest has not translated into substantial, discrete coverage in the academic community, at least in terms of the concrete measures that result in curricular change. We expect that this gap between professional interest and curricular inclusion will narrow as the strategic benefits of the use of Big Data and its use in contemporary data analysis become more difficult to ignore for both academics and practitioners alike.

Table 2. Analytics and Big Data Content in Top Undergraduate Accounting Programs by Specificity to Accounting

Level→University↓	Basic or Introductory Statistics	Introductory IS/MIS	Accounting Information Systems	Generic Decision Analysis/ Advanced Statistics	Generic Analytics/ Big Data	Analytics/Big Data for Accountants	Points	Program Rank	Analytics Rank
Arizona State University	2	1			2		7	27	1
University of Wisconsin		1	1		2		6	5	2
University of Pennsylvania	2					1	5	9	3
Michigan State University	1	1	1		1		5	15	3
Florida State University	1	1		1	1		5	28	3
Brigham Young University	1	2	1				4	1	6
University of Notre Dame		1				1	4	3	6
University of Illinois	1	1		2			4	4	6
Univeristy of Mississippi	1	1	1	1			4	10	6
University of Georgia	1	1	1	1			4	16	6
North Carolina State University	1	1	1	1			4	35	6
University of Texas	1	1		1			3	2	12
Texas A&M Univesity	1	1	1				3	6	12
Ohio State University	1		1	1			3	8	12
Indiana University	1	1	1				3	11	12
University of Florida	1		1	1			3	12	12
University of Missouri	1		1	1			3	13	12
University of Washington	1	1	1				3	19	12
Wake Forest University	1	1	1				3	21	12
University of Tennessee	1	2					3	22	12
Virginia Tech	1	1	1				3	23	12
California State University -- Northridge	1	1	1				3	29	12
University of Maryland	1	1	1				3	30	12
Loyola University of Maryland	1	1	1				3	31	12
Bently University		1	1	1			3	33	12
University of Alabama	1		1				2	7	26
New York University	1	1					2	18	26
Pennsylvania State University	1	1					2	24	26
University of Minnesota	1	1					2	25	26
University of Massachusetts	1		1				2	26	26
City University of New York -- Baruch		1		1			2	31	26
City University of New York -- Brooklyn		1	1				2	33	26
University of Southern California	1						1	14	33
University of Virginia	1						1	17	33
Miami University			1				1	20	33

7. References

AACSB. 2018. Accounting Accreditation. Available at: http://www.aacsb.edu/accreditation/accounting.

Big Data Case Study: What Forensic Accountants Need to Know. (2015). Journal of Accountancy, 219(3), 33.

Bower, J. L. and Christensen, C. M. (1995). Disruptive Technologies: Catching the Wave, Harvard Business Review, Vol. 73, No. 1, 43-53.

Chiang, R. H. L., Goes, P., and Stohr, E. A. (2012). Business Intelligence and Analytics Education, and Program Development: A Unique Opportunity for the Information Systems Discipline, ACM Transactions on Management Information Systems, Vol. 3, No. 3, Article 12, October 2012.

Chandler, N., Hostmann, B., Rayner, N. and Herschel, G. (2011). Gartner's Business Analytic Framework. Gartner, Inc.

Capriotti, R. J. (2014). BIG DATA: Bringing Big Changes to Accounting. Pennsylvania CPA Journal, 85(2), 36-38.

Cohen, K. (2018). Bloomberg: Today's 'hottest job' is data scientist, Becker's Health IT & CIO Report, May 23, 2018.

Conway, D. (2013). The Data Science Venn Diagram. Retrieved from http://drewconway.co/zia/2013/3/26/the-data-science-venn-diagram.

Davenport, T. H., (2006). Competing on Analytics, Harvard Business Review, 84, 98-107.

Evans, J. R. (2020). Business Analytics: Methods, Models and Decisions, 3rd Edition, pg. 4, Pearson Education, Inc.

Griffin, P.A. and Wright, A.M. (2015). Commentaries on Big Data's Importance for Accounting and Auditing, Accounting Horizons, 29(2), pp.377-380.

Krahel, J. and Titera, W., (2015). Consequences of Big Data and Formalization on Accounting and Auditing Standards, Accounting Horizons, 29(2), pp. 431-438.

Lewis, G, (2018). The 20 Most Promising Jobs of 2018, LinkedIn Talent Blog, https://business.linkedin.com/talent-solutions/blog/trends-and-research/2018/the-20-most-promising-jobs-of-2018.

McClure, R. and Sircar, S., (2008). Quantitative Literacy for Undergraduate Business Students in the 21st Century. Journal of Education for Business, July/August, 369-374.

McKenzie Analytics, (2018). Analytics Comes of Age, McKenzie and Company.

Murphy, M. L. and Tysiac, K., 2015. Data analytics helps auditors gain deep insight, Journal of Accountancy, April 13, 2015.

PWC, (2015). Data Driven—What Students Need to Succeed in a Rapidly Changing Business World, February 2015, http://pwc.to/2uyf7EO.

Rebele, J. and Pierre, K. 2015. Stagnation in accounting education research. Journal of Accounting Education, 33 128–137.

Shimamo, D, (2013). CPAs and Business Intelligence: How CPAs Can Drive Business Intelligence. AICPA.

Vasarhelyi, M.A., Kogan, A, and Tuttle, B.M. (2015), Big Data in Accounting: An Overview, Accounting Horizons, 29(2), pp. 381-396.

Warren, J.D., Moffitt, K.C. and Byrnes, P. (2015). How Big Data Will Change Accounting, Accounting Horizons, 29(2), pp. 397-408.

Vickrey, W. (2013). CFOs and Analytics: Should CFOs Oversee Big Data and Accounting Analytics? Business Analytics and Financial Reporting.

Woodside, J., Augustine, F., Chambers, V. and Mendoza, M., (2019). Integrative Learning and Interdisciplinary Information Systems Curriculum Development in Accounting Analytics. EDSIG Conference on Information Systems & Computing Education.

MODELING CUSTOMER BEHAVIOR USING AGENT-BASED SIMULATION FOR INTERNET OF THINGS ENVIRONMENT

Mohammed Basingab[1]
[1]Industrial Engineering Department, King Abdulaziz University
mbasengab@kau.edu.sa

Abstract

The Internet of Things (IoT) technology represents a paradigm where things are viewed as objects that are connected and traceable through the internet network. As IoT technology is progressing nowadays, the number of connected objects to the internet is increasing. Therefore, modeling the IoT system becomes complex but at the same time extremely important which needs simulation technique with a high degree of flexibility. For that reason, this paper proposes a conceptual study to investigate the applicability of Agent-Based Simulation (ABS) as a method to model and determine the usage of IoT technology. The communication and interaction between customers are simulated using ABS to demonstrate the market behavioral intention toward IoT in the state of Florida in the USA.

1. Introduction

The concept of IoT was initially proposed by Kevin Ashton in 1999 (Zhang, 2011). Ashton defined IoT technology as identifiable connected objects (devices) with radio-frequency identification (RFID). However, till now the IoT does not have a unique definition (Madakam, Ramaswamy, & Tripathi, 2015). The IoT definition is still in the enhancing process (Li, Xu, & Zhao, 2015). According to Lee & Lee (2015), IoT technology can be defined as a universal network of physical things that can integrate and communicate with each other without human intervention. IoT, also known as the internet of everything, nowadays is recognized as one of the most important future technologies (Basingab, 2017). Consequently, IoT applications began to receive quick attention from a diversity of industries and a wide range of businesses (Lee & Lee, 2015).

Although simulation and modeling are vital aspects that might efficiently support the Internet of Things (IoT) system improvement, an integrated method providing both of them is lacking (Kaminski, Murphy, & Marchetti, 2016). For this reason, this paper suggests using Agent-Base Simulation (ABS) to investigate the factors that affect the customers purchasing behavior toward IoT technology.

ABS model is a bottom-up technique of modeling that considers a network of independent agents that is mainly applicable to complex systems. ABS approaches model entity of a system as agents. Each agent has its own set of behaviors and characteristics. An agent behavior can direct the interactions not only of another agent but also of other agents in the environment. The system level operation arises from the agent's interactions, which allows modelers to demonstrate the effects of the alterations under study (Kaminski et al., 2016; C. M. Macal & North, 2005). The contributions of this study include:

1. Proposing the applicability of Agent-Based Simulation (ABS) to examine Internet of Things (IoT) system.

2. Investigating the IoT adoption in Florida State market.
3. Examining the impact of different internal and external factors that affect the agent's purchasing behaviors.

2. Related Work

Originally, the social scientists exposed the potential uses of Agent-Based Simulation (ABS) when they began to form study groups that examined the applicability of ABS modeling for simulating individuals' behaviors in the 1990s (Brailsford, 2014; C. Macal & North, 2009). According to North and Macal (2007), applications of ABS span a wide range of areas. Applications of ABS range from modelling customer (agent) behavior in stock markets (Arthur, 1997) and supply chain management(C Macal, Sallach, & North, 2004) to expecting the spread of epidemics (Bagni, Berchi, & Cariello, 2002) and the risk of bio-warfare (Carley et al., 2006), from simulating the adaptive immune system (Folcik, An, & Orosz, 2007) to understanding agent behavior (North et al., 2010), from realizing the ancient civilizations fall (Kohler, Gumerman, & Reynolds, 2005) to simulating the involvement of armies on the battlefield (Moffat, Smith, & Witty, 2006), and many others (North & Macal, 2007). Table 1 shows some of the potential research issues for ABS (Garcia, 2005).

Table 1. Potential research issues for ABS (Garcia, 2005)

Diffusions of Innovations	• Effects of network externalities • Word-of-mouth networks • Modeling tipping points • Social networks and viral marketing
Organizations	• Innovation networks and collaboration • Coevolution of competitive strategies • R&D emergence of innovations • Portfolio management • Innovation strategies and external environmental influences
Knowledge/ Information Flows	• Supply chain networks • Innovation/R&D collaboration (inter- and intraorganizational) • Technology transfer (inter- and intraorganizational; to and from customers–lead users) • Strategy planning (organizational)

In this study, the agents are defined as elements that use improved automaton to accommodate decision making based on collected data from our preliminary survey.

3. Case Study

ABS was built to model the customer purchasing behavior, based on collected data from our preliminary survey, of IoT predictive maintenance software in Florida state in the USA. The predictive maintenance software based on IoT helps clients to constantly analyze real-time data collected from various IoT sensors to predict when the maintenance should be scheduled and therefore minimize the failure rate. In this case study, the main focus is to examine the adoption rate of the IoT software for three types of potential users (agents): banks, information technology (IT) companies, and manufacturers. The customer segments in this study were selected according to the accessibility and geolocalization of agents (Table 2).

Table 2. Types of agents

#	Types of Agents	Number of agents
Agent 1	Banks	200
Agent 2	IT companies	100
Agent 3	Manufacturers	200

Each type of agent has its unique variables and parameters. ABM was built using agent statechart as an intuitive way of simulating agents' purchasing behavior. AnyLogic software was used to build ABS model. The simulation software was chosen based on its ability to capture the complexity of any system and provide a high degree of flexibility.

4. Description of The Simulation Model

Statechart was used to model the customer purchasing behaviors for each customer (agent). The structure of the statechart for each agent is shown in Figure 1. Potential state includes the potential users who are interested to buy the IoT software. User state contains the users who are currently using the IoT software. Wants_To_Buy state includes customers who have made their decision to purchase the software but they have not used it yet. The transition between one state to another is triggered by two types of factors (external and internal). In the ABS model, the external factors that affect agent behavior directly are: 1) Online Advertisements and 2) Direct Sale while the internal factor that affects agent decision indirectly is the agent communication through word of mouth (WOM). In ABS, agent communication was modeled by considering internal and external communication. The interaction between the same type of agent (agent type 1 interacts with agent type 1) is considered as the internal communication while the external communication is the interaction between different types of agents (agent type1 interacts with agent type 2 or 3).

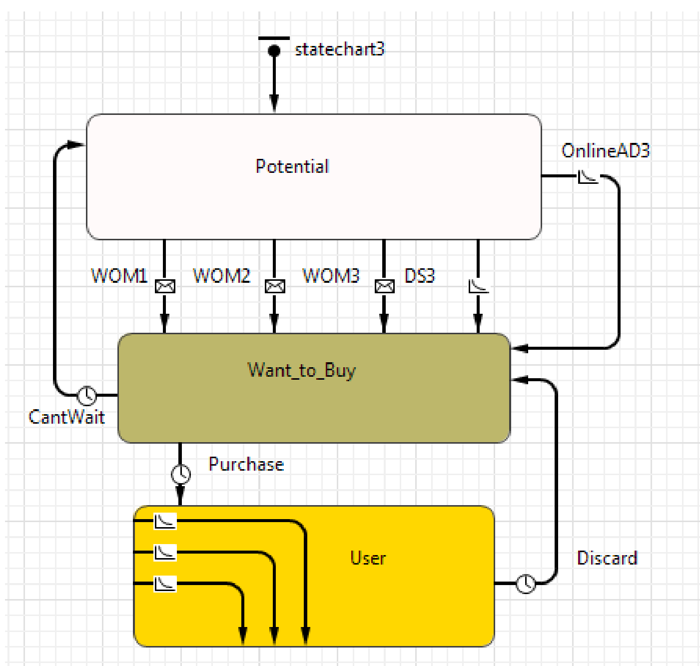


Figure 1. Structure of agent statechart

To model the agent communication, internal transition in the *User* state and message passing logic were inserted in ABS model. The massage passing logic was used to simulate the WOM. Any type of agent can send messages to the same agent or to a group of different types of agents. For example, Agent 2 can send the message "Buy_2To_1" (Figure 2) to affect the purchasing behavior of agent 1. In the internal transition, the contact rate between all the agents was defined. After passing the messages, the transition from *Potential* state to *Wants_To_Buy* state is triggered by a specified adoption fraction rate (Figure 3).

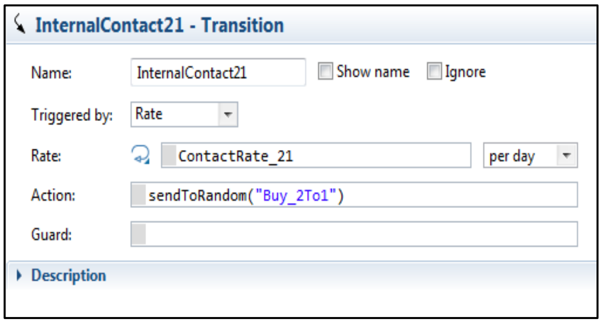


Figure 2. Massage passing from agent type 2 to 1

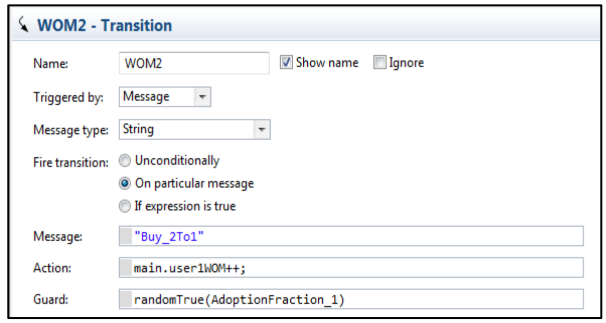


Figure 3. Adoption fraction (agent type 1)

In ABS, the following assumptions are made:

- The demographic characteristics are not considered in the model
- No competitors in Florida state market
- All agents are interacting with each other
- The direct sales and adoption fraction effect are 15% and 30% respectively
- Online Advertisement and Word of mouth (WOM) effectiveness are 17% and 18% respectively (Based on McKinsey report which was issued in April 2010 (Bughin, Doogan, & Vetvik, 2010).

5. Data Analysis

ABS modeling needs data to represent the dynamic setting of the real world. This paper investigates the IoT adoption as an emerging technology. The data in this study was collected by searching for related information in well-known industry reports and approaching experts from numerous industries to investigate their awareness of IoT technology and big data analytics. Based on 50 data points collected, our survey result provides an important statistical and probabilistic pattern on how different sales channels such as direct sales and online advertisements can affect the agents purchasing decision. Also, the survey helps in finding how likely a potential user/business would suggest certain products to another user/business within the same industry or not. Table 3 shows the ABS input parameter. Part of this data was taken from a previous publication of the same author (Basingab, 2019).

Table 3. Simulation input parameter

Parameter	Data	Type	Source
The effect of online AD on agent 1	Triangular (0.13, 0.15, 0.17)	Rate / Year	Data Analysis
The effect of direct sale on agent 1	Triangular (0.12, 0.15, 0.17)	Rate / Year	Assumption
The effect of online AD on agent 2	Triangular (0.15, 0.17, 0.18)	Rate / Year	Data Analysis
The effect of direct sale on agent 2	Triangular (0.12, 0.15, 0.17)	Rate / Year	Assumption
The effect of online AD on agent 3	Triangular (0.08, 0.11, 0.12)	Rate / Year	Data Analysis

The effect of direct sale on agent 3	Triangular (0.12, 0.15, 0.17)	Rate / Year	Assumption
Interaction between agent 1 and 1	0.16	Rate / Day	Data Analysis
Interaction between agent 1 and 2	0.011	Rate / Day	Data Analysis
Interaction between agent 1 and 3	0	Rate / Day	Data Analysis
Interaction between agent 2 and 1	0.02	Rate / Day	Data Analysis
Interaction between agent 2 and 2	0.16	Rate / Day	Data Analysis
Interaction between agent 2 and 3	0	Rate / Day	Data Analysis
Interaction between agent 3 and 1	0.08	Rate / Day	Data Analysis
Interaction between agent 3 and 2	0	Rate / Day	Data Analysis
Interaction between agent 3 and 3	0.103	Rate / Day	Data Analysis
Adoption Fraction for All Agents	0.3	Percentage	Assumption

- **Simulation Preliminary Results**

ABS model is considered as a pilot testing that helps to understand how a system responds due to various inputs. It provides significant insights about the effectiveness of particular parameters such as online advertisement, direct sales, or WOM. The firm could adjust its strategy according to the insights provided by ABS. Figure 4 shows the adoption rate for various customer segments after running ABS model for 2 years (the simulation parameters values can be adjusted during the simulation run and the adoption rate changes can be realized instantly).

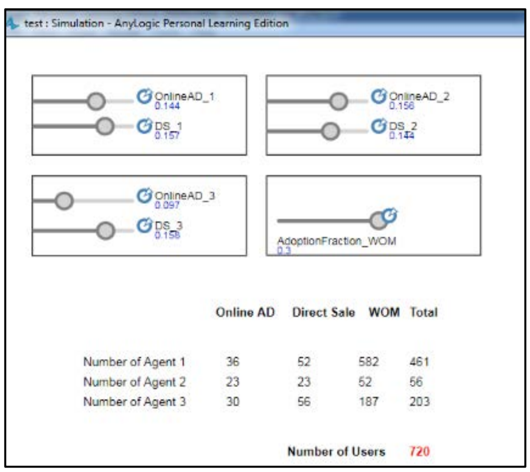


Figure 4. The effect of different sales channels on adoption rate during the simulation run

Figure 5 displays the visualization of ABS animation for the three agents' types of Florida state. Each type of agent has a unique shape and its state is represented by different colors. Theses colors are changing during ABS run which represents the changing of the agent's state. For example, if the color changes from yellow to green, this indicates that state of agent 2 changes from *Potential* state to *User* state. Figure 6 shows the history of the data contribution after running ABS model for 2 years. Figures [7-10] show four different scenarios to realize the effect of the different sale channels on the adoption rate of IoT software.

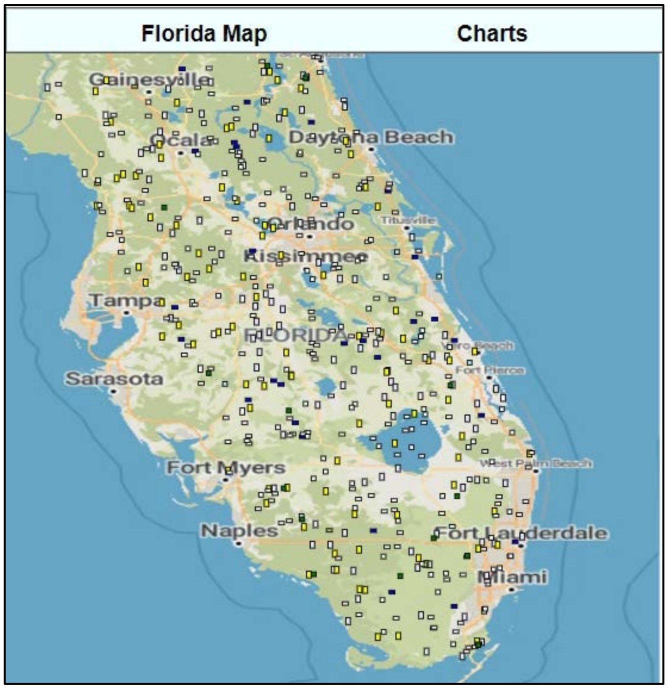


Figure 5. ABS Animation for different agents in Florida state, USA

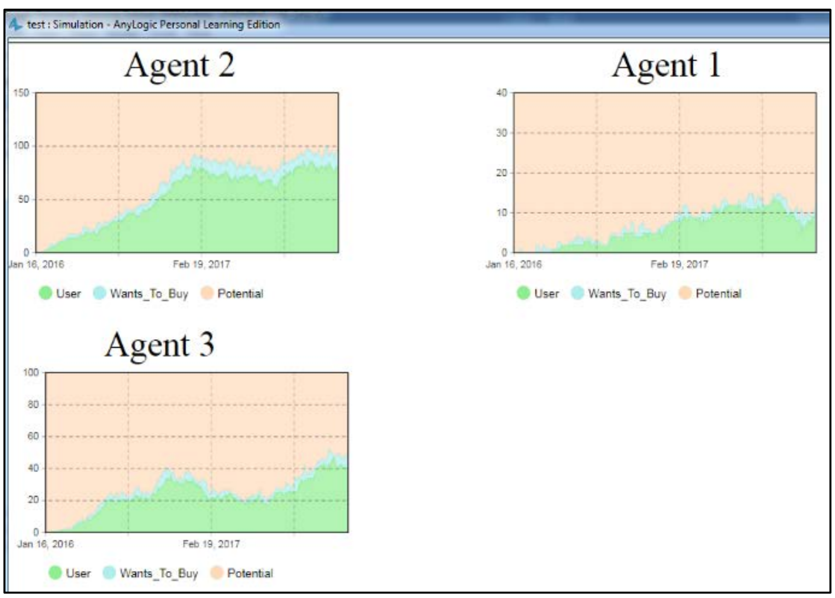
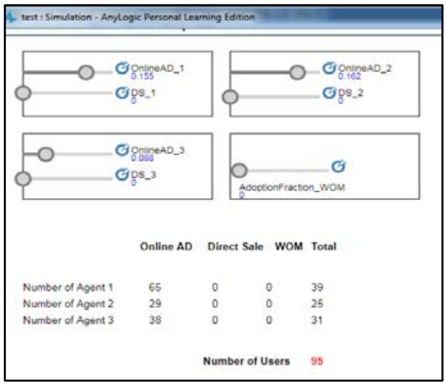


Figure 6. History of the data contribution



Figures 7. The effect of online AD only (Scenario 1)

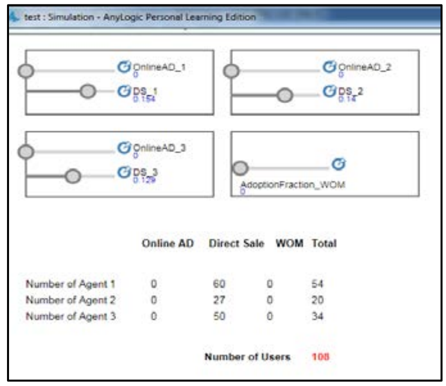


Figure 8. The effect of direst sale only (Scenario 2)

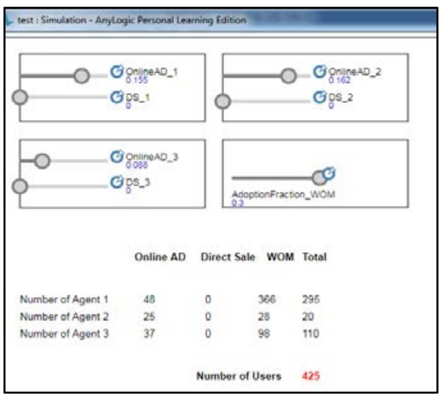


Figure 9. The effect of online AD+WOM (Scenario 3)

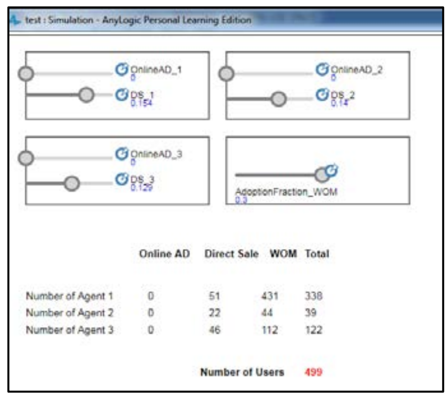


Figure 10. The effect of direst sale + WOM (Scenario 4)

- **Optimization**

The purpose of the optimization experiment is to find the optimal combination of parameters that satisfies the main objective function and system constraints. An optimization experiment using the capabilities provided in AnyLogic software was conducted to maximize the net income. Each optimization run contains repetitive simulations with different parameters. As it was mentioned before, the main purpose of ABS model is to examine the adoption of IoT predictive maintenance software by simulating customer (agent) behavior. Our optimization experiment determines how to increase the net income by manipulating the simulation parameters such as the online advertisement and WOM.

The inputs of the optimization experiment are the parameters analyzed from the survey results and expert opinions while the outputs represent the maximum net income with the optimal parameters values under predefined constraints. This section provides a straightforward example of using the optimization approach for the proposed ABS model. Based on the predefined input simulation parameters, some assumptions related to the sales revenue and the marketing strategy (online advertisement and direct sale) cost are assumed to realize the optimization logic (Table 4). The ABS results are shown in Figure11.

Table 4. Revenue and costs assumptions

Parameter	Data	Type
Revenue per user	Triangular (90, 100, 120)	$
Cost of online advertisement per user	Triangular (20, 30, 40)	$
Cost of Online direct Sale per user	Triangular (10, 25, 50)	$
Fixed Cost per user	Triangular (40, 50, 80)	$

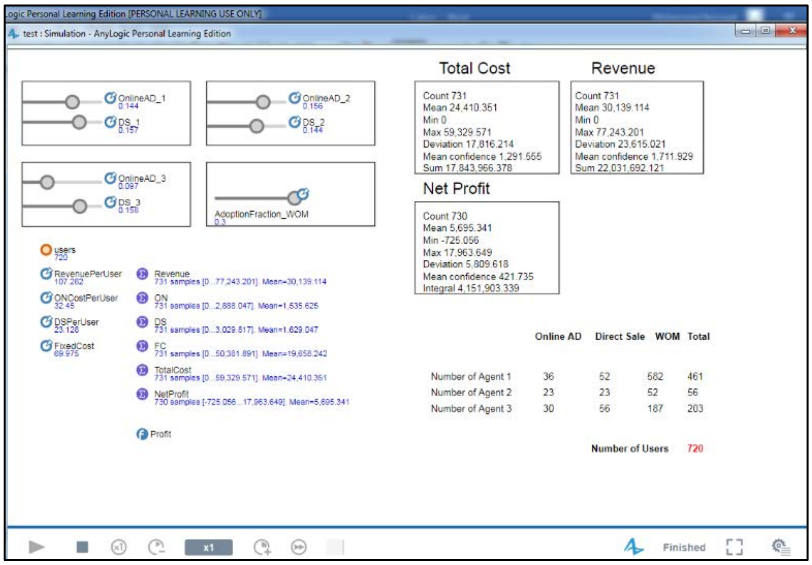


Figure 11. The observed results of ABS Model

After perceiving the results from the ABS model, an optimization experiment was conducted (100 iterations with 3 replications per each iteration) to obtain the best-case scenario based on the objective function and constraints shown in Figure 12 and Figure 13 respectively. The objective function is to maximize the net profit under the following constraints:

1. The total investment cost is less than $ 85,000
2. The maximum number of customers is 1200 users
3. The minimum waiting time is from 12 days
4. The maximum waiting time is from 15 days
5. The minimum delivery time is from 15 days
6. The maximum delivery time is from 25 days

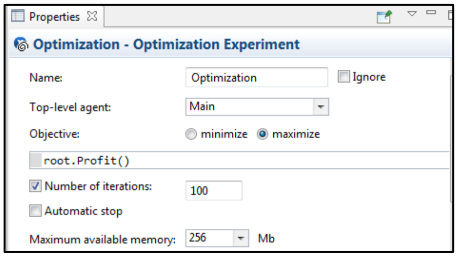


Figure 12. The objective function

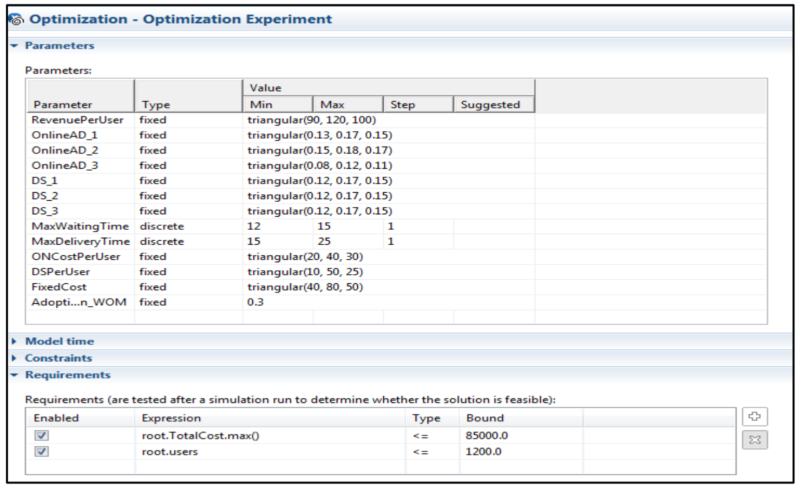


Figure 13. Model constraints

Figure 14 shows the optimization results. In this experiment, the software manipulates the parameters to find their optimal values using various combinations which will lead to maximizing the net profit. In our case, the net profit was improved by around 12%.

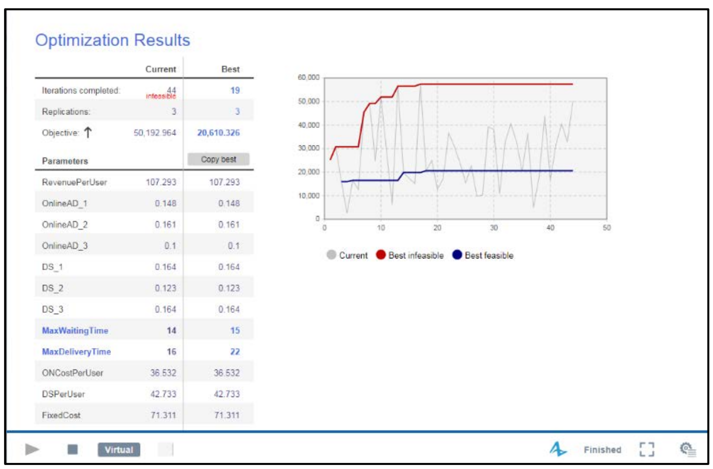


Figure 14. Simulation results after optimization

6. Conclusion and Future Work

ABS concept allows modelers to capture customer behaviors more than any statistical or analytical models. Agent preferences, learning, and history are examples of such behaviors. Capturing these behaviors allows a high-quality prediction in customer markets. In the proposed hypothetical market model, Agent-Based Simulation (ABS) helped to simulate potential customers behaviors in the state of Florida in the USA which might provide valuable insights. Simulating agents in IoT paradigm has some challenges including: IoT technology security concerns, overcome connectivity issues, keep IoT hardware updated and wait for the governmental regulation.

Since modeling and simulating Internet of Things (IoT) system is still at an early stage, this paper proposed a conceptual approach by providing a preliminary design. The future work includes: 1) ABS validation. 2) Increase the size of the data sample and the level of details. 3) Deeper investigation on the information related to online advertisement and WOM effectiveness. 4) Capturing agent behaviors continuous in time by including system dynamics fragments inside each agent. 5) Simulating the differences in users' behaviors based on their locations by considering the exact physical locations of each agent in Florida state. 6) Sensors will not be enough to predict IoT maintenance schedule, but manufacturers of IoT component data sheet be included to determine when such devices need replacement.

7. References

Arthur, W. (1997). Introduction,(pp. 1-14) in Arthur WB, Durlauf SN, Lane DA (eds.) The economy as an evolving complex system II: Addison-Wesley, Reading.

Bagni, R., Berchi, R., & Cariello, P. (2002). A comparison of simulation models applied to epidemics. Journal of Artificial Societies and Social Simulation, 5(3).

Basingab, M. (2017). Internet of Things Business Modeling and Analysis Using Agent-based Simulation. University of Central Florida.

Basingab, M. (2019). Investigating The Adoption of Internet of Things Technology Using Agent-Based Simulation. Paper presented at the International Conference on Industrial Engineering and Operations Management, Bangkok, Thailand.

Brailsford, S. (2014, 7-10 Dec. 2014). Modeling human behavior - an (id)entity crisis? Paper presented at the Simulation Conference (WSC), 2014 Winter.

Bughin, J., Doogan, J., & Vetvik, O. J. (2010). A new way to measure word-of-mouth marketing. McKinsey Quarterly, 2, 113-116.

Carley, K. M., Fridsma, D. B., Casman, E., Yahja, A., Altman, N., Chen, L.-C., . . . Nave, D. (2006). BioWar: scalable agent-based model of bioattacks. IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans, 36(2), 252-265.

Folcik, V. A., An, G. C., & Orosz, C. G. (2007). The Basic Immune Simulator: an agent-based model to study the interactions between innate and adaptive immunity. Theoretical Biology and Medical Modelling, 4(1), 39.

Garcia, R. (2005). Uses of agent-based modeling in innovation/new product development research. Journal of Product Innovation Management, 22(5), 380-398.

Kaminski, N. J., Murphy, M., & Marchetti, N. (2016). Agent-based modeling of an IoT network. Paper presented at the 2016 IEEE international symposium on systems engineering (ISSE).

Kohler, T. A., Gumerman, G. J., & Reynolds, R. G. (2005). Simulating ancient societies. Scientific American, 293(1), 76-84.

Lee, I., & Lee, K. (2015). The Internet of Things (IoT): Applications, investments, and challenges for enterprises. Business Horizons, 58(4), 431.

Li, S., Xu, L. D., & Zhao, S. (2015). The internet of things: a survey. Information Systems Frontiers, 17(2), 243-259. doi:http://dx.doi.org/10.1007/s10796-014-9492-7

Macal, C., & North, M. (2009, 13-16 Dec. 2009). Agent-based modeling and simulation. Paper presented at the Simulation Conference (WSC), Proceedings of the 2009 Winter.

Macal, C., Sallach, D., & North, M. (2004). Emergent structures from trust relationships in supply chains. Paper presented at the Proc. Agent 2004: Conf. on Social Dynamics.

Macal, C. M., & North, M. J. (2005). Tutorial on agent-based modeling and simulation. Paper presented at the Proceedings of the Winter Simulation Conference, 2005.

Madakam, S., Ramaswamy, R., & Tripathi, S. (2015). Internet of Things (IoT): A literature review. Journal of Computer and Communications, 3(05), 164.

Moffat, J., Smith, J., & Witty, S. (2006). Emergent behaviour: theory and experimentation using the MANA model. Advances in decision sciences, 2006.

North, M. J., & Macal, C. M. (2007). Tutorial on agent-based modeling and simulation. Retrieved from

North, M. J., Macal, C. M., Aubin, J. S., Thimmapuram, P., Bragen, M., Hahn, J., Hampton, D. (2010). Multiscale agent-based consumer market modeling. Complexity, 15(5), 37-47.

Zhang, Y. (2011). Technology Framework of the Internet of Things and its Application. Paper presented at the 2011 International Conference on Electrical and Control Engineering.

ADDRESSING CORRUGATED BOARD WARP WITH A 2^3 FACTORIAL DESIGN

Robert S. Keyser[1]
Lois A. Jordan[1]
[1]Kennesaw State University
rkeyser@kennesaw.edu; ljorda61@kennesaw.edu

Abstract

The corrugator is a machine that produces combined corrugated sheet stock from roll stock. A 23 factorial experiment was used to determine which factors and levels provide the optimal result – flat, rather than warped, sheet stock in a full-line box plant located in the Southeast. The factors include Edge Tension (T), Preheater Wrap Pressure (P), and Run Speed (S). The response variable is board warp, measured in 1/64ths of an inch. The two levels are indicated by (+) or (–) signs to indicate high and low levels, respectively. In a 23 factorial design, there are 8 treatment combinations and n = 20 samples of each combination of factors and levels were collected and analyzed using one replication for immediate use in a real-time production environment. The conclusion yielded the following best settings to produce flat 200#C sheet stock: loose tension, 25% pressure on the preheater wrap, and a run speed of 380 ft/min.

1. Introduction

Corrugated boxes, often mistakenly referred to as cardboard boxes, are omnipresent in our society. From a user's standpoint, corrugated boxes are used to package and protect their contents. However, from a producer's standpoint, the production of corrugated boxes is a way of life, and various elements of a corrugated box are of utmost importance to ensure that the functionality, usability, and reliability of the finished product satisfy the user's purpose.

In the corrugated box industry, trees are harvested and go through a multitude of processes at a paper mill to make paper – either linerboard (i.e., the inner and outer plies of paper in a corrugated sheet) or medium (i.e., the fluted paper in a corrugated sheet). Paper is turned into roll stock of varying widths and paper grades. Some of the roll stock may be treated with chemical additives to increase the strength properties of the paper or for moisture protection. Roll stock is then shipped to customers such as box plants or sheet-feeder plants by freight line or railway, where it is mounted on a corrugator at various stations at the beginning of the process, treated with starch for an adhesive bond, and then converted into combined board at the end of the process. Sheet stock, cut to specific sizes as per customer requirements, may have scores, or creases, in the combined board added by the slitter section of the corrugator during this process.

A corrugator is a machine that produces combined corrugated sheet stock from roll stock. A corrugator can be found in a sheet-feeder plant or a full-line box plant. Whereas a sheet-feeder plant only produces combined board (or sheet stock) from roll stock, a full-line plant produces combined board from roll stock but also has the machinery and capability of converting sheet stock into a variety of value-added products, such as corrugated boxes, inner-packing, partitions, and die-cuts. A sheet plant differs from a full-line plant in that a sheet plant does not produce combined board; rather, it only converts combined board that it purchases from sheet feeders into a variety of products similar to those of a full-line plant.

At a full-line box plant in the Southeast, a 23 factorial experiment was used to determine which factors and levels provide the optimal result – flat, rather than warped, sheet stock that will eventually run through converting equipment such as a press or die-cutter. Flat board typically runs much better through machinery than warped board. Warped board comes in many styles. For example, “up warp” means that the front and back ends of sheet stock are curved “up.” “Down warp” means that the front and back ends of sheet stock are curved “down.” “S-warp” means that the front edge of sheet stock could be curved “up” while the back edge of sheet stock is curved “down” or vice-versa. S-warped board could also give the appearance of being “twisted” such that opposite corners, diagonal from one another, are curved in opposite directions – one corner curved “up” and the opposite, diagonal corner curved “down.” Techniques such as “breaking” the board (i.e., bending the flaps in the opposite direction of the warp in order to flatten the board) can be employed to make warped board run through machinery, but this is a tedious and time-consuming manual endeavor, particularly on long production runs. However, if the order has to run and ship, this technique usually works although the run time might be significantly longer than scheduled. The results of this experiment were implemented immediately on the corrugator due to a very long run of this particular board grade.

2. Literature Review

A large body of literature exists pertaining to various strength properties of corrugated board using different methodologies; however, no published research was discovered addressing the fabrication of corrugated board on the corrugator, the machine that converts roll stock into combined board, using designed experiments.

Much published work concerns the compressive strength of corrugated boxes. For example, Baker, Horvath & White (2016) employs simulation modeling to understand the interface between corrugated boxes and the pallets they are stored on to estimate box compression strength. Berry, Ambawa, Defraeye, Coetzee, & Oparaa (2019) simulates refrigerated freight containers (RFCs), which include ventilated box designs transported in refrigerated units in high relative humidity climatic conditions to determine box compression strength. Frank (2014) examines the literature focusing on the process of box compression and the utility of box compression testing while exploring various parameters that affect box compression strength.

Park et al. (2017) investigates the physical interactions among packages stacked on pallets, known as ‘load bridging,’ via the relationships of package size, flute type and pallet stiffness to load bridging and unit-load deflection (i.e., when a box begins to buckle under pressure).

Biancolini & Brutti (2003) uses simulation to determine the effects of edge crush and compressive strength of corrugated boxes and compares the experimental simulated data with numerical results using finite element methods (FEM). Similarly, Hua, Shen, Zhao, & Xie (2017) employs experimental and numerical simulation methods to investigate the edge effect of corrugated boxes and their influence on edgewise compressive strength. Finite element analysis is also used to develop a conceptual sustainable packaging model that integrates the technical design of corrugated boxes, supply chain systems, and environmental factors (Dominic, Ostlund, Buffington, & Masoud, 2015). Nonlinear finite element analysis is used to analyze fluted corrugated sheets for the effects of shape and size of flutes and their performance in corrugated boxes via a static pressure test and a drop test (Zhang, Qiu, Song, & Sun, 2014).

3. Design Methodology

At this particular box plant, it was hypothesized through brainstorming that the three primary factors contributing to the production of flat or warped board at the corrugator were Edge Tension (T), Preheater Wrap Pressure (P), and Run Speed (S), and each factor was then tested at two levels, which are indicated by "+" and "-." For Edge Tension, which indicates the tension level at both edges across the web of paper roll stock as it runs through the corrugator, the two levels were Tight (+) and Loose (-). For Preheater Wrap Pressure, which indicates the percentage of rotational pressure applied as roll stock paper wraps around the preheater cylinders so that it can bond well with the starch adhesive to form combined board, the two levels were 75% (+) and 25% (-). For Run Speed, which indicates the corrugator run speed in feet per minute, the two levels were 480 ft/min (+) and 380 ft/min (-). The product of this experiment was 200#C sheet stock, which consists of two 42 lb. liners and a 26 lb. medium. The paper weights are basis weights and translate into pounds per 1000 sq. ft. The response variable is the board warp, measured in 1/64ths of an inch. A value of zero indicated no warp, or perfectly flat sheet stock.

A 23 factorial design yields 8 treatment combinations, or runs, for each replication. Since this designed experiment was performed in a real-time production environment, only one replication was used so that results could be quickly analyzed and employed. The sample size was n = 20 samples for each combination of factors and levels. The average and Range for each run is shown in Table 1.

Table 1. Raw Data

	Edge Tension (T)			
	Tight (+)		Loose (-)	
	Preheater Wrap (P)		**Preheater Wrap (P)**	
	75% (+)	25% (-)	75% (+)	25% (-)
Speed (S) 480 ft/min (+)	x-bar = 27 R = 22	x-bar = 16 R = 8	x-bar = 25 R = 21	x-bar = 14 R = 17
Speed (S) 380 ft/min (-)	x-bar = 26 R = 18	x-bar = 26 R = 9	x-bar = 14 R = 5	x-bar = 10 R = 0

A full 2^3 factorial with Minitab software was initially used to analyze the raw data yielding the X-bar ANOVA results followed by the Range ANOVA results below.

Full Factorial Fit: X-bar Result versus Edge Tension, Preheater Wr, Speed (S)

```
Estimated Effects and Coefficients for X-bar Result (coded units)

Term                                 Effect    Coef
Constant                                     19.750
Edge Tension (T)                      8.000   4.000
Preheater Wrap (P)                    6.500   3.250
Speed (S)                             1.500   0.750
Edge Tension (T)*Preheater Wrap (P)  -1.000  -0.500
Edge Tension (T)*Speed (S)           -6.000  -3.000
Preheater Wrap (P)*Speed (S)          4.500   2.250
Edge Tension (T)*Preheater Wrap (P)*  1.000   0.500
  Speed (S)
```

```
S = *   PRESS = *

Analysis of Variance for X-bar Result (coded units)

Source                                          DF   Seq SS   Adj SS   Adj MS
Main Effects                                     3  217.000  217.000   72.333
  Edge Tension (T)                               1  128.000  128.000  128.000
  Preheater Wrap (P)                             1   84.500   84.500   84.500
  Speed (S)                                      1    4.500    4.500    4.500
2-Way Interactions                               3  114.500  114.500   38.167
  Edge Tension (T)*Preheater Wrap (P)            1    2.000    2.000    2.000
  Edge Tension (T)*Speed (S)                     1   72.000   72.000   72.000
  Preheater Wrap (P)*Speed (S)                   1   40.500   40.500   40.500
3-Way Interactions                               1    2.000    2.000    2.000
  Edge Tension (T)*Preheater Wrap (P)*Speed (S)  1    2.000    2.000    2.000
Residual Error                                   0        *        *        *
Total                                            7  333.500

Source                                          F  P
Main Effects                                    *  *
  Edge Tension (T)                              *  *
  Preheater Wrap (P)                            *  *
  Speed (S)                                     *  *
2-Way Interactions                              *  *
  Edge Tension (T)*Preheater Wrap (P)           *  *
  Edge Tension (T)*Speed (S)                    *  *
  Preheater Wrap (P)*Speed (S)                  *  *
3-Way Interactions                              *  *
  Edge Tension (T)*Preheater Wrap (P)*Speed (S) *  *
Residual Error
Total
```

Full Factorial Fit: Range Result versus Edge Tension, Preheater Wr, Speed (S)

```
Estimated Effects and Coefficients for Range Result (coded units)

Term                                  Effect    Coef
Constant                                      12.500
Edge Tension (T)                       3.500   1.750
Preheater Wrap (P)                     8.000   4.000
Speed (S)                              9.000   4.500
Edge Tension (T)*Preheater Wrap (P)    3.500   1.750
Edge Tension (T)*Speed (S)            -7.500  -3.750
Preheater Wrap (P)*Speed (S)           1.000   0.500
Edge Tension (T)*Preheater Wrap (P)*   1.500   0.750
  Speed (S)

S = *   PRESS = *

Analysis of Variance for Range Result (coded units)

Source                                          DF   Seq SS   Adj SS   Adj MS
Main Effects                                     3  314.500  314.500  104.833
  Edge Tension (T)                               1   24.500   24.500   24.500
  Preheater Wrap (P)                             1  128.000  128.000  128.000
  Speed (S)                                      1  162.000  162.000  162.000
2-Way Interactions                               3  139.000  139.000   46.333
  Edge Tension (T)*Preheater Wrap (P)            1   24.500   24.500   24.500
  Edge Tension (T)*Speed (S)                     1  112.500  112.500  112.500
  Preheater Wrap (P)*Speed (S)                   1    2.000    2.000    2.000
3-Way Interactions                               1    4.500    4.500    4.500
  Edge Tension (T)*Preheater Wrap (P)*Speed (S)  1    4.500    4.500    4.500
Residual Error                                   0        *        *        *
```

```
Total                                           7  458.000

Source                                         F  P
Main Effects                                   *  *
  Edge Tension (T)                             *  *
  Preheater Wrap (P)                           *  *
  Speed (S)                                    *  *
2-Way Interactions                             *  *
  Edge Tension (T)*Preheater Wrap (P)          *  *
  Edge Tension (T)*Speed (S)                   *  *
  Preheater Wrap (P)*Speed (S)                 *  *
3-Way Interactions                             *  *
  Edge Tension (T)*Preheater Wrap (P)*Speed (S)  *  *
Residual Error
Total
```

The missing values in the output, such as the standard deviation, PRESS statistic, F-values, p-values, and Residual Error values are indicative of a saturated model including all main effects and all interaction terms and leaving zero degrees of freedom for calculating the residual error.

Having zero degrees of freedom for the error term in ANOVA causes many calculations to fail. For example, the Adj MS (adjusted mean square) for Residual Error, also known as the mean square of the error (MSE), cannot be calculated because it is impossible to divide anything by zero degrees of freedom. Because MSE cannot be calculated, neither the F nor the p-value can be calculated since the F-value uses MSE in its calculation and the p-value uses the F-value in its calculation.

Because of this problem with the error term degrees of freedom, both the X-bar and R results were analyzed separately using Yates' algorithm, and those factors that appeared to have the largest effect on the results were identified. If significant terms can be identified with this method, the other non-significant terms can be dropped and the ANOVA re-run as a Reduced model to increase the error degrees of freedom.

Table 2 shows results and net differences between the high and low X-bar values.

Table 2. X-bar results using Yates' algorithm

				X-bar				
Run	**T**	**P**	**S**	**T*P**	**P*S**	**T*S**	**T*P*S**	**Result**
1	+	+	+	+	+	+	+	27
2	+	-	+	-	-	+	-	16
3	-	+	+	-	+	-	-	25
4	-	-	+	+	-	-	+	14
5	+	+	-	+	-	-	-	26
6	+	-	-	-	+	-	+	26
7	-	+	-	-	-	+	+	14
8	-	-	-	+	+	+	-	10
(+)	95	92	82	77	88	67	81	
(-)	63	66	76	81	70	91	77	
Difference	32	26	6	-4	18	-24	4	

Note: *n* = 20 samples for each combination of variables

We observed from the X-bar results that Edge Tension (T), Preheater Wrap Pressure (P), and the Edge Tension (T) x Run Speed (S) interaction terms had the largest effect on the results.

Next, Table 3 shows results and net differences between the high and low Range values.

Table 3. Range results using Yates' algorithm

				Range				
Run	**T**	**P**	**S**	**T*P**	**P*S**	**T*S**	**T*P*S**	**Result**
1	+	+	+	+	+	+	+	22
2	+	-	+	-	-	+	-	8
3	-	+	+	-	+	-	-	21
4	-	-	+	+	-	-	+	17
5	+	+	-	+	-	-	-	18
6	+	-	-	-	+	-	+	9
7	-	+	-	-	-	+	+	5
8	-	-	-	+	+	+	-	0
(+)	57	66	68	57	52	35	53	
(-)	43	34	32	43	48	65	47	
Difference	14	32	36	14	4	-30	6	

Note: *n* = 20 samples for each combination of variables

We observed from the Range results that Preheater Wrap Pressure (P), Run Speed (S), and the Edge Tension (T) x Run Speed (S) interaction terms had the largest effect on the results. Since the only interaction term which had a large effect on both the x-bar and Range results was the Edge Tension (T) x Run Speed (S) interaction, the other interaction terms were dropped from the Full model to yield a Reduced model. The X-bar results in a Reduced model, followed by the Range results, are provided below. The factors included all three main effects and only the Edge Tension (T) x Run Speed (S) two-factor interaction term.

Reduced Factorial Fit: X-bar Result versus Edge Tension, Preheater Wr, Speed (S)

```
Estimated Effects and Coefficients for X-bar Result (coded units)

Term                         Effect    Coef  SE Coef      T      P
Constant                             19.750    1.362  14.50  0.001
Edge Tension (T)              8.000   4.000    1.362   2.94  0.061
Preheater Wrap (P)            6.500   3.250    1.362   2.39  0.097
Speed (S)                     1.500   0.750    1.362   0.55  0.620
Edge Tension (T)*Speed (S)   -6.000  -3.000    1.362  -2.20  0.115


S = 3.85141     PRESS = 316.444
R-Sq = 86.66%   R-Sq(pred) = 5.11%   R-Sq(adj) = 68.87%


Analysis of Variance for X-bar Result (coded units)

Source                         DF   Seq SS   Adj SS   Adj MS     F       P
Main Effects                    3  217.000  217.000   72.333  4.88  0.113
  Edge Tension (T)              1  128.000  128.000  128.000  8.63  0.061
  Preheater Wrap (P)            1   84.500   84.500   84.500  5.70  0.097
  Speed (S)                     1    4.500    4.500    4.500  0.30  0.620
2-Way Interactions              1   72.000   72.000   72.000  4.85  0.115
```

Edge Tension (T)*Speed (S)	1	72.000	72.000	72.000	4.85	0.115
Residual Error	3	44.500	44.500	14.833		
Total	7	333.500				

Reduced Factorial Fit: Range Result versus Edge Tension, Preheater Wr, Speed (S)

Estimated Effects and Coefficients for Range Result (coded units)

Term	Effect	Coef	SE Coef	T	P
Constant		12.500	1.137	11.00	0.002
Edge Tension (T)	3.500	1.750	1.137	1.54	0.221
Preheater Wrap (P)	8.000	4.000	1.137	3.52	0.039
Speed (S)	9.000	4.500	1.137	3.96	0.029
Edge Tension (T)*Speed (S)	-7.500	-3.750	1.137	-3.30	0.046

S = 3.21455 PRESS = 220.444
R-Sq = 93.23% R-Sq(pred) = 51.87% R-Sq(adj) = 84.21%

Analysis of Variance for Range Result (coded units)

Source	DF	Seq SS	Adj SS	Adj MS	F	P
Main Effects	3	314.50	314.50	104.83	10.15	0.044
Edge Tension (T)	1	24.50	24.50	24.50	2.37	0.221
Preheater Wrap (P)	1	128.00	128.00	128.00	12.39	0.039
Speed (S)	1	162.00	162.00	162.00	15.68	0.029
2-Way Interactions	1	112.50	112.50	112.50	10.89	0.046
Edge Tension (T)*Speed (S)	1	112.50	112.50	112.50	10.89	0.046
Residual Error	3	31.00	31.00	10.33		
Total	7	458.00				

4. Analysis

From the Reduced model, we observed that no term is significant at the alpha = 0.05 level in the X-bar results, and that Preheater Wrap Pressure (P), Run Speed (S), and Edge Tension (T) x Run Speed (S) interaction are significant terms at the alpha = 0.05 level in the Range results.

The results are validated in the following Main Effects Plot for X-bar results. Figure 1 shows that X-bar results for main effects Edge Tension (T) and Preheater Wrap Pressure (P) as being somewhat significant, but not significant at the alpha = 0.05 level. The mostly flat line for Run Speed (S) indicates that this term is not significant.

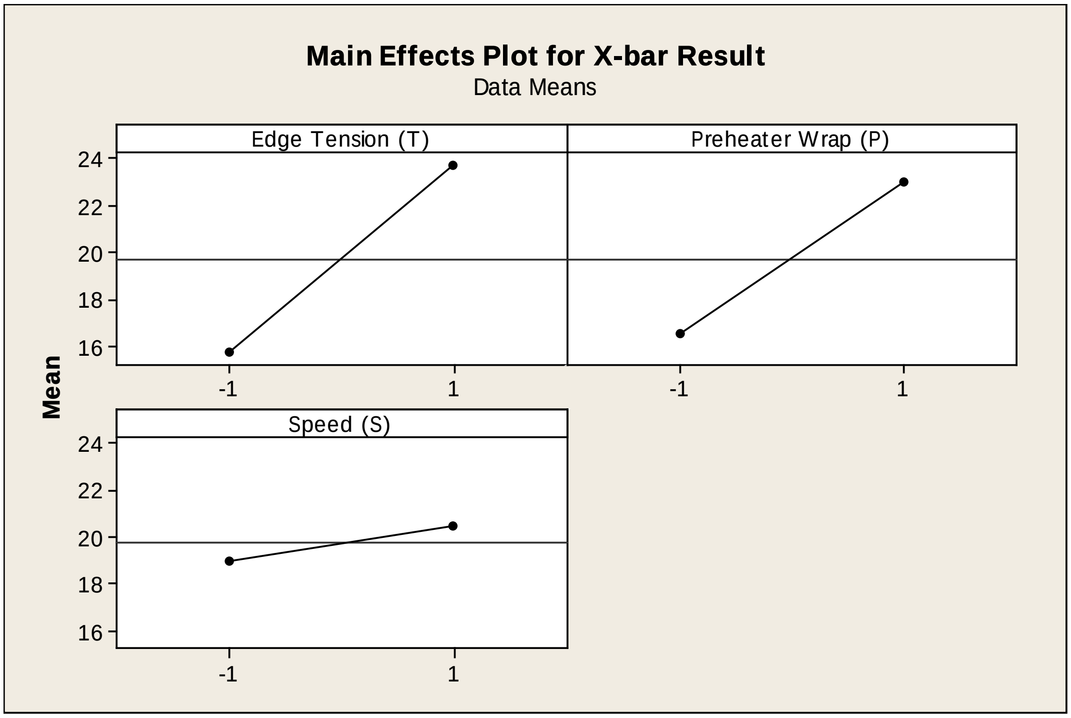


Figure 1. Main Effects Plot for X-bar Result in the Reduced Model

The two-factor interaction plots of x-bar results shown in Figure 2 compare favorably to what we see in Table 2; particularly, that Edge Tension (T) x Run Speed (S) interaction and Preheater Wrap Pressure (P) x Run Speed (S) interaction both have an effect on the result, but neither interaction term is significant at alpha = 0.05.

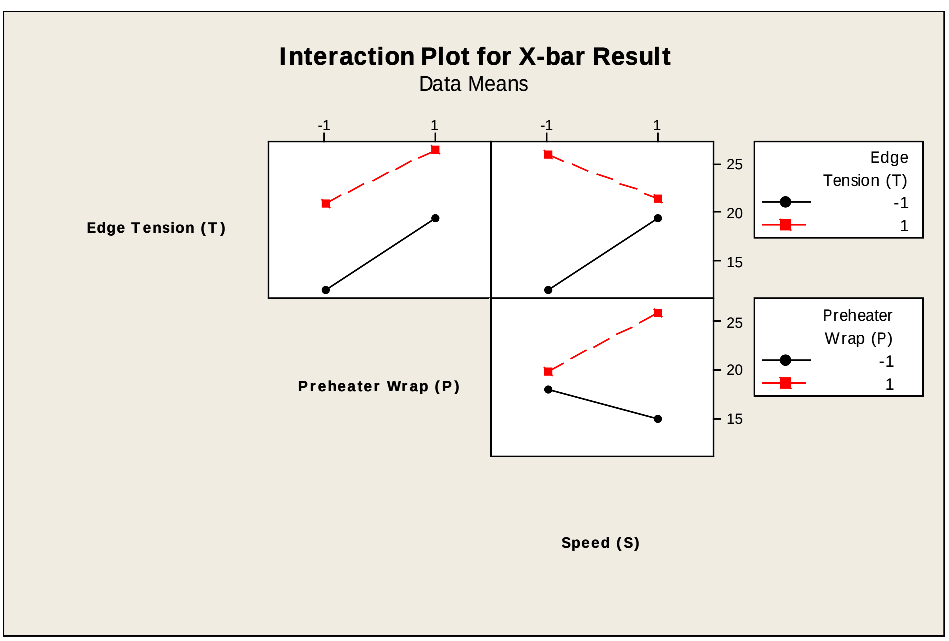


Figure 2. Interaction Plot for X-bar Result in the Reduced Model

However, we observed that Preheater Wrap Pressure (P), Run Speed (S), and Edge Tension (T) x Run Speed (S) interaction are all significant terms at the alpha = 0.05 level in the Range results for the Reduced model.

The results are validated in the following Main Effects Plot for the Range results. Figure 3 shows that Range results for main effects Preheater Wrap Pressure (P) and Run Speed (S) are significant at the alpha = 0.05 level. The mostly flat line for Edge Tension (T) indicates that this term is not significant.

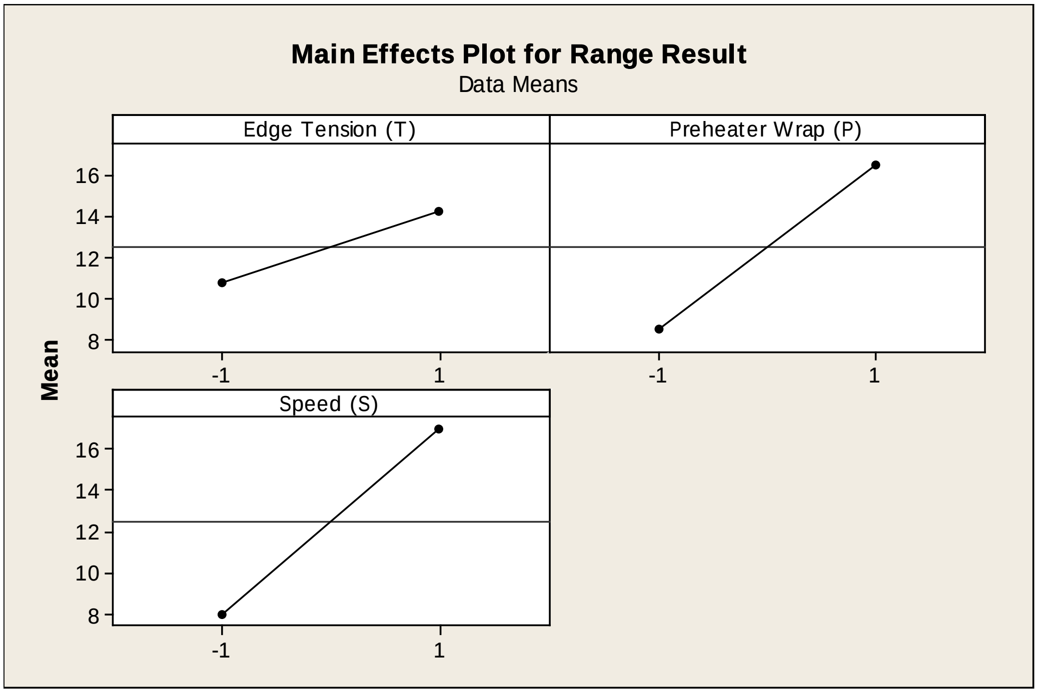


Figure 3. Main Effects Plot for Range Result in the Reduced Model

The two-factor interaction plots of Range results shown in Figure 4 compare favorably to what we see in Table 3; particularly, that Edge Tension (T) x Preheater Wrap Pressure (P) interaction and Edge Tension (T) x Run Speed (S) interaction both have an effect on the result. However, only the Edge Tension (T) x Run Speed (S) interaction term is significant at alpha = 0.05.

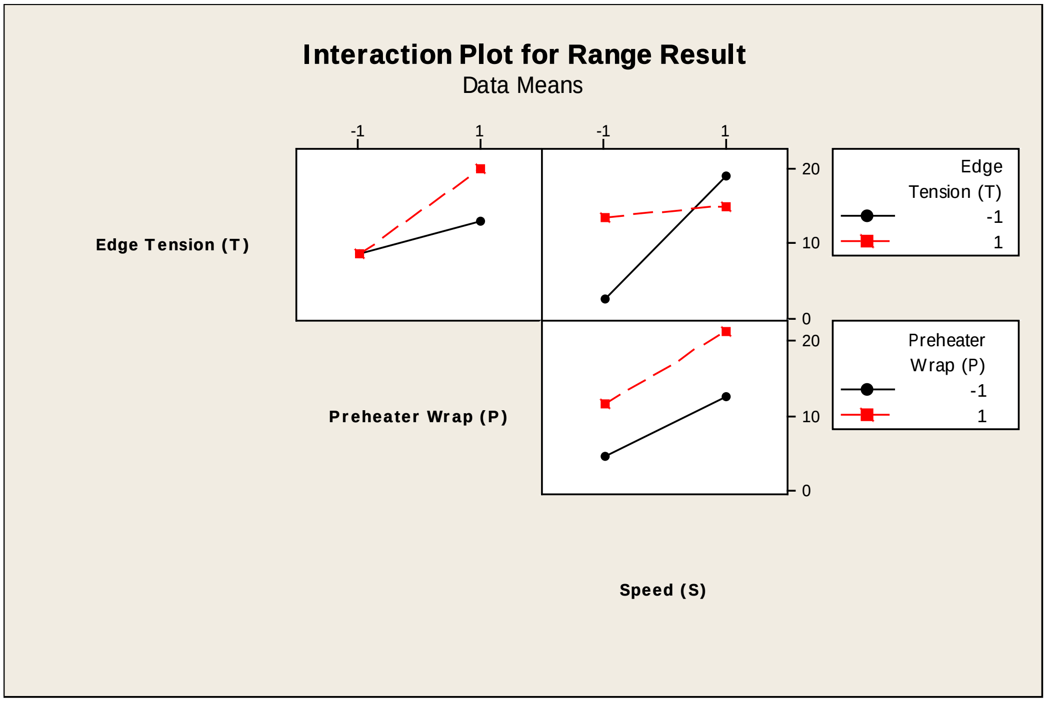


Figure 4. Interaction Plot for Range Result in the Reduced Model

We now take a closer look at the Edge Tension (T) x Run Speed (S) interaction term since it is common to both X-bar and Range results. Table 5 shows X-bar board warp results for the T x S interaction term.

Table 5. X-bar Results for T x S Interaction Term

		X-bar	
		T * S Interaction	
		Edge Tension (T)	
		Tight (+)	Loose (-)
Speed (S) (+)	480	(1) + (2) 27 + 16 = 43	(3) + (4) 25 + 14 = 39
Speed (S) (-)	380	(5) + (6) 26 + 26 = 52	(7) + (8) 14 + 10 = 24

We observed by the net difference in board warp results that changing from Tight to Loose Edge Tension (T) doesn't matter much for a Run Speed (S) of 480 ft/min, but it has a much larger effect at a Run Speed (S) of 380 ft/min. Similarly, changing from one Run Speed (S) to another has a much larger effect with a Loose Edge Tension (T) than with a Tight Edge Tension (T). Preheater Wrap Pressure (P) doesn't have an effect here.

Table 6 shows Range board warp results for the Edge Tension (T) x Run Speed (S) interaction term.

Table 6. Range Results for the T x S Interaction Term

		Range	
		T * S Interaction	
		Edge Tension (T)	
		Tight (+)	Loose (-)
Speed (S) (+)	480	(1) + (2) 22 + 8 = 30	(3) + (4) 21 + 17 = 38
Speed (S) (-)	380	(5) + (6) 18 + 9 = 27	(7) + (8) 5 + 0 = 5

We again observed by the net difference in board warp results that changing from Tight to Loose Edge Tension (T) doesn't matter much for a Run Speed (S) of 480 ft/min, but it has a much larger effect at a Run Speed (S) of 380 ft/min. Similarly, changing from one Run Speed (S) to another has a much larger effect with a Loose Edge Tension (T) than with a Tight Edge Tension (T). Preheater Wrap Pressure (P) also doesn't have an effect here.

5. Conclusions

Table 7 shows concluding X-bar and Range results from the Reduced model.

Table 7. Best Settings for Significant Main Effects and Interaction Terms

X-bar/Range	Best Settings
Edge Tension/Edge Tension Preheater Wrap/Preheater Wrap Run Speed/Run/Speed	25%/25%
Tension*Run Speed/Tension*Run Speed	Loose, 380/Loose, 380

The Reduced model shows that all three main effects: Edge Tension (T) (p-value = 0.221), Preheater Wrap Pressure (P) (p-value = 0.039), and Run Speed (S) (p-value = 0.029) and the Edge Tension (T) x Run Speed (S) two-factor interaction term (p-value = 0.046) are significant variables. From Table 1, when looking at Edge Tension, the X-bar and Range results are smaller (i.e., closer to 0 for flat sheet stock) for Loose tension vs. Tight tension. Similarly, when examining Preheater Wrap, the X-bar and Range results are smaller (i.e., closer to 0) for 25% vs. 75%. Additionally, when examining Run Speed, the X-bar and Range results are smaller (i.e., closer to 0) for 380 ft/min vs. 480 ft/min.

The main effects conclusions are supported by the results for the Edge Tension x Run Speed two-factor interaction term in Table 6. When looking at Edge Tension, the results are smaller at a Run Speed of 380 ft/min vs. 480 ft/min. When looking at Run Speed, the results are smaller for Loose tension vs. Tight tension. The conclusion is that the best settings should be Edge Tension = Loose (-), Preheater Wrap Pressure = 25% (-), and Run Speed = 380 ft/min (-) in order to achieve the desired outcome of producing flat 200#C corrugated sheet stock off the corrugator.

6. Research Limitations

The analysis and conclusion for these 23 factorial experiments came as the result of only one replication of 8 runs; however, this was desired because the designed experiment was performed during a production day and the results were to be implemented immediately. The experiment worked in the sense that the conclusion of best settings achieved the desired results – producing flat corrugated sheet stock in a live production environment that would be run on converting equipment in the plant the same day.

7. Areas of Future Research

Future research could include multiple replications in the same production environment to determine whether these results are reproducible. This could be performed during a scheduled downtime period in the plant so as not to interfere with current production needs. Additionally, experiments could be run at different times of the year (i.e., each quarter, day shift vs. night shift) to determine whether the best settings differ throughout the year due to climatic conditions.

8. References

Baker, M., Horvath, L., & White, M. (2016). Effect of pallet deckboard stiffness on corrugated box compression strength. Packaging Technology and Science, 29, 263-274/ DOI:10.1002/pts.2201

Berry. T., Ambawa. A., Defraeye, T., Coetzee, C., & Oparaa, U. L. (2019). Moisture adsorption in palletised corrugated fibreboard cartons under shipping conditions: A CFD modelling approach. Food & Bioproducts Processing, 114, 1-32.

Biancolini, M. E., & Brutti, C. (2003). Numerical and experimental investigation of the strength of corrugated board packages. Packaging Technology & Science, 16, 47-60. DOI:10.1002/pts.609

Dominic, C. A., Ostlund, S., Buffington, J., & Masoud, M. M. (2015). Towards a conceptual sustainable packaging development model: A corrugated box case study. Packaging Technology & Science, 28, 397-413. DOI:10.1002/pts.2113

Frank, B. (2014). Corrugated box compression – A literature survey. Packaging Technology & Science, 27, 105-128. DOI:10:1002/pts.2019

Hua, G. J., Shen, Y., Zhao, D., & Xie, Y. (2017). Experimental and numerical analysis of the edge effect for corrugated and honeycomb fiberboard. Strength of Materials, 49, 188-197.

Park, J., Horvath, L., White, M. S., Phanthanousy, S., Araman, P., & Bush, R. J. (2017). The influence of package size and flute type of corrugated boxes on load bridging in unit loads. Packaging Technology & Science, 30, 33-43. DOI:10.1002/pts.2279

Zhang, Z., Qiu, T., Song, R., & Sun, Y. (2014). Nonlinear finite element analysis of the fluted corrugated sheet in the corrugated cardboard. Advances in Materials Science and Engineering, 1-8. Retrieved from http://dx.doi.org/10.1155/2014/654012

APPLICATIONS OF ADDITIVE MANUFACTURING IN CONSTRUCTION INDUSTRY: A LITERATURE REVIEW

Abdelhakim A. Al Turk[1]
Gamal S. Weheba[1]
[1]Department of Industrial, Systems and Manufacturing Engineering , Wichita State University
gamal.weheba@wichita.edu

Abstract

Additive manufacturing is the process where three–dimensional parts are constructed layer-by-layer based on their computer-aided design (CAD) models. AM has gained a strong reputation in the aerospace and medical industries, and its use is now considered revolutionary. In the past few years, the utilization of AM technology has been expanded and is now leading to breakthroughs in the construction industry. This paper provides a state-of-the-art review of the use, advantages, and limitations of AM in the construction industry. The review sheds light on the different AM techniques used and highlights contemporary projects and achievements in the construction industry.

1. Introduction

Additive manufacturing (AM) has been defined by the American Society for Testing and Materials (ASTM) as "the process of joining materials to make objects from 3D model data, usually layer upon layer." Even though materials and ways vary, the basic process divides the CAD model into layers and constructs them as prototypes or end-use products. Originally, the technology was confined to creating prototypes to support new product design activities. Currently, all AM techniques have been successfully adopted for the manufacturing of functional end-use products. Among these techniques are Stereolithography, Binder Jetting, Material Jetting, Powder Bed Fusion, Material Extrusion, Sheet Lamination, and Directed Energy Deposition. These techniques can create solid objects in layers by using different materials and methods for bonding them together. A more detailed description of these techniques can be found in the ASTM F2792 Standards.

In the construction industry, the constant growth of AM led to a wide range of applications. Design modeling, automated constriction of full structures or structural elements, restoration of historical buildings, and repair of existing structures are examples of current applications of AM in this industry. Early adoptions of the AM technology appeared to target restoration projects. Rather than focusing on solid 3D-printed parts, some construction firms utilized the technology to construct concrete molds. These plastic molds were inlaid with wire mesh to reinforce the cast. A good example of this application is the process developed by Edg (edgnyc.com), an architecture and engineering firm in NY. This paper is focused on large-scale AM systems that have been developed and utilized in construction projects. The following sections present a review of Contour Crafting, Binder Jetting, and Hybrid Systems currently utilized in the industry.

2. Contour Crafting

Contour Crafting (CC) is a layered fabrication technology introduced by Khoshnevis in 1986. In his article, Khoshnevis (2003) described CC as a process that uses a computer to exploit the superior surface-forming capability of a bladed trowel for creating smooth and accurate surfaces with intricate features. A distinguishing feature of CC is its capability to fabricate large-scale objects. The process is based on combining an extrusion system with the filling process to build concrete objects as shown in Figure 1. As the material is extruded through an extrusion nozzle, the top and side trowels work on the material to create a smooth outer edge of the rim. The filling process consists of pouring material to fill the area constructed by the extrusion nozzle.

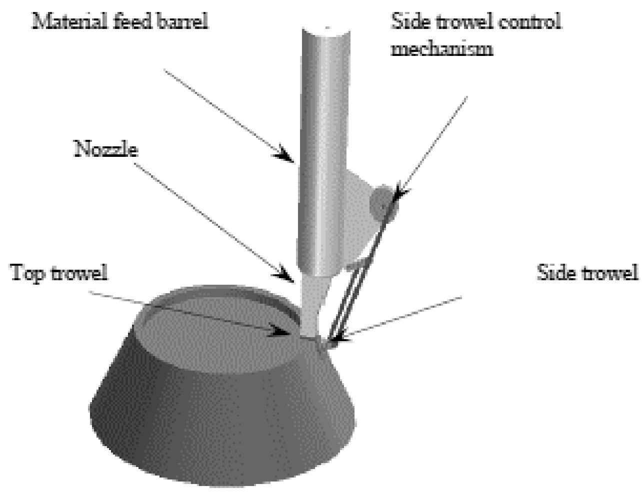


Figure 1. Schematic of the Contour Crafting Process

CC is a layered manufacturing technique that offers a potential solution for difficulties faced in the construction industry. Higher fabrication speed, improved surface quality along safety benefits are some of the advantages offered with this technology as was noted by Khoshnevis et al (2006). The use of two trowels greatly influences the surface finish and quality of the final objects. Typically, top and side trowels are perpendicular to each other. The latter is capable of pivoting to facilitate the creation of the rim surfaces with non-orthogonal shapes. Trowel geometry, height, and rotation are key factors in achieving accurate and smooth surfaces. The pouring mechanism is still needed to fill the area bounded by the rim walls.

In 2014 WinSun, a Chinese Engineering Design Company achieved success in building concrete houses shown in Figure 2 using CC technology. During this project, 10 concrete houses were built in Shanghai, China, within one day. The area of each house was 195 m2 at the cost of $4800. The construction was performed using four giant printers. Each printer is about 10 meters wide and 6.7 meters high. The basic components of each house were built offsite and then assembled on site.

Figure 2. Concrete Houses by WinSun

In 2015, the same company constructed the tallest 3D printed structure. This project consisted of constructing a five-story apartment building. The total area of this project was about 1100 m2, making it the tallest 3D printed structure in China to date. According to WinSun, using this technology reduced the construction time by 50%-70%, labor cost by 50%-80%, and material costs by 30%-60%. By the end of 2015, WinSun achieved another millstone by constructing the world's first 3D printed villa. The total area of this villa is approximately 1,100 m2 complete with internal and external decoration. WinSun holds 98 national patents for construction materials.

In 2016 Huashang Tengda, another Chinese company, constructed a two-story villa, using similar printing technology in 45 days. The total area of this project was 400 m2. The project started after all the plumbing pipes and steel reinforcements were installed. The project utilized a standard class C30 concrete containing coarse aggregates. According to the Huashang company, the building is safe and can withstand earthquakes of magnitude 8.0 on the Richter scale.

In 2016, some researchers from the University of Federico II, Italy, configured the WASP 3D printer, capable of printing 3 m long concrete beams using CC printing technology. Using this system, a 3 m tall cave, called the Y- Box Pavilion was structured in Thailand. The construction cost was around $28,000. The components of this project were constructed offsite and then assembled on site.

Figure 3. Apis Core World's biggest 3d printed building

In 2019, Apis Core, a US-based Company, completed the World's Largest 3D printed building shown in Figure 3. This company printed wall structures of a two-story building for the Dubai Municipality. The walls are 9.5 meters tall and the area of this project is 640 m2. This project entered the Guinness Book of World Records as the World's largest 3D printed structure to date. Also, all the project was constructed on-site without any off-site assembly work. According to Apis Core, this project was completed under critical environmental conditions, since the work area was not covered.

3. Binder Jetting Technology

As reported by Gibson et al (2015), this technology was developed at MIT in the early 1990s. Originally, it was called 3D Printing (3DP) in which binder material is printed onto a powder bed to form one layer of the object. Once the first layer is constructed, a fresh layer of powder is spread over the bed, smoothed by a roller, and the binder is applied. These steps are repeated until the object is constructed. A schematic presentation of the process is shown in Figure 4.

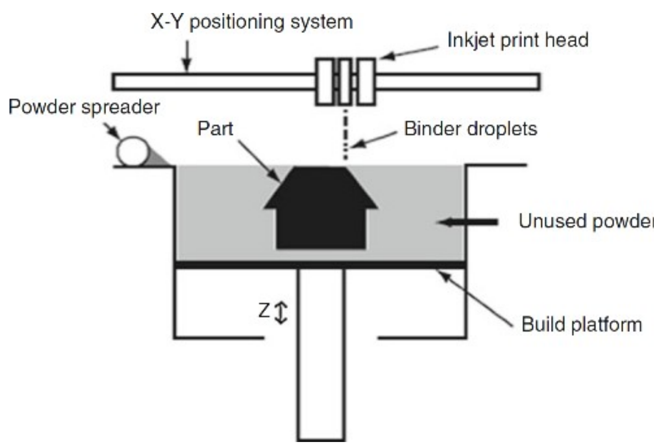


Figure 4. Schematic of the Binder Jetting Process

Depending on the material and the binder used, heat may be applied to assure adhesion within and between layers. The extra powder has to be removed using vacuum or compressed air during post-processing steps. Typically, constructed objects require some infiltration to increase their mechanical properties.

Early applications of the binder jetting process in construction were reported by Duxson et al (2006), by utilizing a geopolymer instead of the standard ceramic powder material. This is a mix of slag and fly ash, a type of alkali aluminosilicate, with the same characteristics as cement. They used a Z- Corp 450 3DP system with the standard Zb 63 binder. Under these conditions, the geopolymer material was shown to have some advantages over the standard cement including good resistance to sulfate and acid attack, high compressive strength, and low shrinkage. They proposed a new post-processing technique and tested five different mixes of geopolymers. They recommended using a mix of 100% slag and 0% fly ash and concluded that the new post-processing technique significantly increased the strength of the constructed samples.

In 2007, Monolite Ltd, a British company, introduced its d-shape process. This is a 3DP-based process capable of printing architectural components using micro-concrete powder with an inorganic binder in layers of 5-10 mm thickness. The powder is made of granules of different materials ranging from 0.2 mm to 5 mm size.

The process is claimed to construct components with marble-like properties. The company offers four different configurations with build volumes of up to 12 m x 12 m x 10 m. Printers are equipped with mixers and feeding systems to deliver the material to the powder spreader. A vacuum system is used to remove unbounded powder during post-processing stages. The constructed components are typically ground and polished to the final shape. The process was successfully utilized in constructing a 1.6 m high architectural pieces called "Radiolaria" depicted in Figure 5.

Recently, Xia et al (2019) reported on applications of 3DP technology for concrete printing using geopolymers. They considered the effect of different powder parameters and binder droplet penetration behavior on the printability of the geopolymers. They also investigated the effect of fly ash content on the compressive strength of the printed specimens. The results indicated that geopolymers can be used in the commercially available binder jetting systems.

Figure 5. The Radiolaria: Example of full-scale build using the d-shape process

4. Hybrid Systems

Hybrid systems combine the flexibility of additive systems and the accuracy of traditional subtractive processes. The additive component of these systems covers the concept of building parts layer by layer. Whereas, the subtractive component covers computer numerical control (CNC) machining operations including drilling, milling, and grinding. This combination allows the construction of tools and models used for concrete casting. The definition of hybrid systems in the ASTM F2792 Standards is restricted to laser metal deposition combined with CNC machining. However, Zhu et al (2013) described a hybrid system as a "combination of an additive and subtractive processes, sequential or integrated, including planning for fixturing and orientation in the quest of a final, usable part." Under this wider definition, hybrid AM systems would include a much larger number of AM systems. Examples include Sheet Lamination (e.g., Ultrasonic AM) and some Material Jetting systems like 3D Plotters.

Thermwood, a US-based company, is a leader in large scale additive manufacturing of thermoplastic composite molds, tooling, patterns, and parts. Its line of Large Scale Additive Manufacturing (LSAM, pronounced L-Sam) machines uses a two-step process. The part is 3D printed layer by layer, to slightly larger than the final size, then it is trimmed to its exact final net size and shape using a CNC router. Thermwood's LSAM Print3D software utility, which operates within Mastercam, is used to generate the CNC print program needed to print the part using the print gantry. The CAD model is then used to generate a trim program that is used to finish the part using the trim gantry.

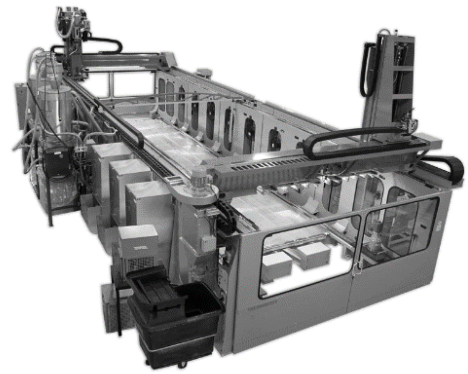


Figure 6. The LSAM 1040

Although suitable for producing a wide variety of components, Thermwood is focusing on producing industrial tooling, masters, patterns, molds, and production fixtures for a variety of industries. In 2017 Thermwood demonstrated its ability to print and trim a concrete mold out of carbon fiber-filled ABS. The company offers three models of the LSAM with the same working principle but different sizes. The build size of the 1040 model (shown in Figure 6) is 3.0 m x 12.2 m, with a maximum height of 1.5 m. Their standard 40 mm melt core (extruder) has a maximum output of between 86.2 kg and 95.3 kg per hour, depending on the polymer used.

5. Cited Advantages

As reported by Camacho et al (2018), an essential advantage of 3D printing is its ability to produce complex geometries that were hard or even impossible to construct with conventional construction methods. Large-scale 3D printing of end-use structures allows architects to build complex passageways, undercuts, and other interior and exterior designs. They are allowed to rethink their sketches and forms without affecting the productivity of the construction process. Architects do not have to worry about the constructability of each part anymore and can now focus on their design and functionality. The Radiolaria project, in Figure 5, is a perfect illustration of geometric freedom made possible by AM.

In traditional construction, the use of formworks is a must. Formworks are made of pieces of wood assembled to give the finished concrete its desired shape. Camacho et al (2018), pointed out that 3D printing allows for the reduction or elimination of formworks. They also noted the potential benefits of using the technology for in-situ repairs of existing structures. They indicated that AM could be used to construct temporary support structures inside the damaged building to allow for inspection and even restoration, thus decreasing the risk of work-related injuries.

Noguchi et al (2018) indicated that applications of AM in the construction industry will lead to cost reduction. 3D printing of concrete structures was reported to reduce the project cost by decreasing the amount of labor, material, formworks, and the project lead-time.

More importantly, as was noted by Khoshnevis (2003), the construction industry is a significant source of waste around the world. Conventional construction activities often consume a sizable amount of natural resources like water, wood, sand, and fine and coarse aggregates while polluting the environment with hazardous emissions and byproducts. He estimated that seven tons of waste is generated each time a conventional house is built using traditional construction methods. AM will help minimize this waste and replace traditional construction methods by the environmentally friendly, emission-free, electrical 3D printing systems.

6. Cited Challenges

As in all potential applications of AM, a major challenge facing concrete printing is the availability of materials. Le et al (2012), indicated that the concrete used for 3D printing needs to have special characteristics concerning its hardening, compressive strength, flowability, and stability. Concrete must bond together to form each layer; it must have enough buildability features to enable it to lay down properly and remain in position. For extrusion-based processes, the concrete must be stiff enough to support further layers without failing with an adequate degree of extrudability. Also, there are conflicting requirements between the flowability of the mix, setting speed, workability, buildability, and the resulting compressive strength. These represent challenges for the widespread applications within the industry.

As was noted by Tay et al (2017) projects constructed using CC and similar systems contain no steel reinforcement, and hence have low tensile strength. These structures must be designed with hollow voids to add the steel later. This may limit the freedom of architectural design. They also indicated that the need for highly skilled labor, capable of operating printers and integrating robotics, is another challenge for AM in construction. They anticipated that re-training of current labor and incorporating newer trainees will become mandatory as 3D printing use expands over the next few years.

Also, as stated by Abderrahim et al (2003), the first challenge faced in the construction industry is workers' safety. Studies have indicated that primary hazard sources are collisions with machines. Zhou et al (2013) pointed to the need for eliminating human activities during the 3D printing process to minimize the risk of injury. They proposed installing a real-time camera to monitor and recorder activities around the printer to improve safety and reduce accidents.

7. Conclusion

Successful applications of AM in Aerospace, Automotive, and the Medical industries have drawn the attention of other industries including the construction industry. The availability of large-scale systems is encouraging architects and designers to try their innovative ideas away from the restrictions of traditional construction methods. Applications are beginning to reveal the benefits of AM throughout the life cycle of construction projects. It appears that the development of large-scale systems has been focused on concrete printing. Both on-site and off-site systems are now commercially available. On-site printing systems offer great flexibility in design with the elimination of the need for labor-intensive, costly formworks. However, the process and materials used appear more sensitive to environmental conditions. Off-site systems, on the other hand, offer these similar benefits but operate under controlled conditions. This provides a better opportunity for assuring construction quality and speed. A disadvantage, however, is the need to transport the printed components to the construction site at a significant cost.

While concrete printing has potential in several areas, there is a need for continuing research in areas of materials, mechanical properties of final components, and process control. Processes need to be reexamined to allow for reinforcement of concrete. Mechanisms for integrating different reinforcement materials, patterns, and shapes during component printing need to be considered. Other AM techniques like sheet lamination, which is capable of integrating sensors and electronic monitoring devices may prove useful in construction projects. The need for skilled labor is a challenge facing the adoption of AM in all industries including construction. As public interest in the AM technology increases, more education and training programs are being offered. The more designers and engineers are trained on the technology, the better the chances are for efficient and effective implementation.

8. References

Abderrahim M., Garcia E., Diez R., and Balaguer C., (2003). "A mechatronic security system for the construction site." Journal of Automation in Construction, 14 (4), pp 460–466.

Apis Core, Online: https://www.apis-cor.com, (accessed 30.01.20).

ASTM, (2012). “F2792. Standard terminology for additive manufacturing technologies.” West Conshohocken, PA: ASTM International. www.astm.org

Camacho, D., D., (2018). “Applications of additive manufacturing in the construction industry – A forward-looking review.” https://doi.org/10.1016/j.autcon.2017.12.031

Duxson P., Fernández-Jiménez A., Provis L., Lukey C., Palomo A., and Deventer S., (2006). “Geopolymer technology: the current state of the art.” Journal of Material Science. 42, pp 2917-2933.

Gibson, Rosen, and Stucker (2015). "Additive Manufacturing Technologies: 3D Printing, Rapid Prototyping and Direct Digital Manufacturing", 2nd Ed., Springer, New York

Huashang-Tengda, On-line: https://3dprint.com/138664/huashang-tengda-3d-printhouse, (accessed 30.01.20).

Khoshnevis B., (2003). “Toward total automation of on-site construction an integrated approach based on contour crafting.” Los Angeles: University of Southern California.

Khoshnevis B., (2004). “Automated construction by contour crafting - related robotics and information technologies.” Automation in Construction, 13 (1), pp 5–19.

Khoshnevis, B., Hwang, D., Yao, K-T. and Yeh, Z. (2006). “Mega-scale fabrication by contour crafting.” International Journal of Industrial and Systems Engineering, 1 (3), pp 301–320.

Le T., Austin A., Lim S., Buswell A., Gibb F., and Thorpe, T., (2012) “Mix design and fresh properties for high-performance printing concrete.” Materials and Structures, 45, pp 1221- 1232.

Noguchi, M., (2003). “The effect of the quality-oriented production approach on the delivery of prefabricated homes in Japan.” Housing Built Environ. 18, pp 353–364.

Tay Y., Panda B., Paul S., Mohamed N., Tan M., and Leong K., (2017) “3D printing trends in building and construction industry: a review.” Virtual and Physical Prototyping, 12 (3), pp 261-276.

Thermwood, on-line: http://www.thermwood.com, (accessed 25.03.20) WinSun China, online: http://www.winsun3d.com (accessed 29.01.20)

Xia M, Nematollahi B, Sanjayan J., (2019). “Printability, accuracy, and strength of geopolymer made using powder-based 3D printing for construction applications. Automation in Construction 101, pp 197-189.

Zhou Z., Irizarry J., and Li Q., (2013). “Applying advanced technology to improve safety management in the Construction Industry: a literature review. Construction Management & Economics, 31 (6), pp 606–622.

Zhu Z., Dhokia V., Nassehi A., and Newman S., (2013). “A review of hybrid manufacturing processes—state of the art and future perspectives.” International Journal of Computer Integrated Manufacturing 26 (7), pp 596- 615.

ENTREPRENEURIAL PSYCHOLOGY: REVISING THE PSYCH 101 COURSE

Aryanne de Silva, Ph.D.[1]
Dennis Ridley, Ph.D.[2]
Sarah Green, BAEd (Hons).[3]
[1]Abigail Wexner Research Institute at Nationwide Children's Hospital
[2]School of Business & Industry, Florida A&M University
[2]Department. of Scientific Computing, Florida State Universit
[3]Eldene Nursery and Primary School
aryanne.desilva@nationwidechildrens.org; dridley@fsu.edu; *sargre28@aol.com*

Abstract

Extant courses in psychology typically begin with the assumption that a stable system of government exists which supports the foundation for the development of society and entrepreneurship. This approach excludes the discussion of capital or assumes its existence. The purpose of this paper is to acknowledge a new CDR index that is remarkable for explaining real gross domestic product (GDP) adjusted for purchasing power parity (G) and use it to introduce new suggestions for teaching psychology as it applies to entrepreneurship theory. It is assumed that the pertinent psychology course is available as a required or optional course for business administration students. The benefit to psychology students is to expand their knowledge about growing areas within psychology and to enable them to join the entrepreneurship community even if they are not themselves entrepreneurs. The benefit to business students is to join the entrepreneurial community as well as to become successful entrepreneurs, as well as help further develop those critical thinking skills that are typically advanced in psychology courses.

1. Introduction

The field of psychology is largely defined as the study of human thoughts and behavior. It naturally follows then that psychology largely involves the investigation of how humans make decisions and how their thoughts and behaviors are shaped by their individual and cultural circumstances. The study of psychology is important for society through the promotion of certain critical thinking skills as well as better understanding of how others operate, which may be related to increased empathy (Deal, 2003), inherently providing a social benefit. Such a social benefit is of utmost importance globally, particularly in capitalist societies where corporations and business entities may have greater ability to exploit opportunities to serve the needs of the populace. Therefore, there may be a societal advantage towards strengthening associations between psychology coursework and entrepreneurial education. Besides a societal benefit, there may also be a theoretical overlap between the study of psychology and entrepreneurship, which interestingly has not been thoroughly capitalized on in the past. Understanding how humans make decisions and the motives behind human actions would substantially overlap with entrepreneurial education, which is focused on understanding demand for goods and services. Lastly, a new and growing discipline within psychology is that of Industrial/Organizational psychology, which involves understanding factors that affect employee behavior and productivity. Such a subspecialty within psychology is an even more obvious opportunity for associations with entrepreneurial learning.

Entrepreneurship is the process of starting a business, typically a startup company offering an innovative product, process or service (see also Schumpeter, 1928). It is borne out of human ideas of imagination and creativity. Products of the mind. And, a mind once stretched by an original idea never regains its original shape. Entrepreneurship education is largely focused on ushering students toward innovative business opportunities, through the creation or modification of existing businesses. Historically, entrepreneurship courses have been siloed in schools focused on business and management, but there has been a more recent push towards introducing entrepreneurial education in other types of curricula, including engineering, economics, and mathematics (Llaugel and Ridley, 2018; Luryi et al., 2007, Ngnepieba et al., 2018, Ridley, 2018). The benefits of incorporating entrepreneurial education into various types of curricula, including the psychology curricula may therefore further support such societal benefits. The advantage of including entrepreneurship learning into other types of curricula has been very appealing on a societal level, but the process has been slow due to resource and knowledge limitations (Llaugel and Ridley, 2018). Due to the benefits, however, some governments, such as in the Dominican Republic, have supported the introduction of entrepreneurial learning in engineering and through such governmental support have overcome many of the limitations involved.

Recent research demonstrated that the only source of wealth is entrepreneurship systems of Capitalism in the presence of Democracy and Rule of Law (CDR). In practice, capital begins as human capital ideas of imagination and creativity. Successful entrepreneurship requires the uninhibited development of children if they are to become the future entrepreneurs. The result is a capital stock of products and services that raise standards of living. Corrupt dictators tend to suppress the development of entrepreneurs leading to widespread poverty. Oppressed children, if that is their only exposure, will themselves become oppressors and corrupt dictators, and the cycle continues. The future of entrepreneurship relies on breaking this cycle through the development of an understanding of CDR and communication of same through instructional curricula. Psychology curricula often focuses on developing certain critical thinking skills and a sense of ethics in students.

The remainder of the paper is organized as follows. Section 2 reviews traditional psychology pedagogy. Section 3 proposes a modern psychology pedagogy. Section 4 contains some conclusions and recommendations for future research. A nomenclature is given to help beginning students understand various terminologies used in psychology and entrepreneurial economics, particularly the elements of the CDR index.

2. Traditional psychology pedagogy

Traditional psychology 101 courses are focused on introducing students to the basic theories of the four main disciplines within psychology: cognitive, developmental, social and clinical psychology. Psychology 101 courses almost universally include a cognitive segment on the brain and nervous system. A developmental segment discussing theories of human development from prenatal development through infancy, childhood, adolescence and into adulthood. A social segment looking at group behaviors. And lastly, a clinical segment that is focused on different types of psychological disorders. The four main disciplines within psychology are the most common ways in which research in psychology has also been traditionally focused, as the four disciplines help us understand the historical tenets of human behavior that have been of interest to researchers and academics.

Such a traditional format for psychology 101 courses has been a fantastic opportunity for students to learn about the basic topics within psychology and provide them a better understanding of the breadth of psychological study. As the world has become more industrialized, however, there has been a growing push for greater study into human behavior as related to employee relations and productivity, and as such the industrial/organizational discipline within psychology has been gaining momentum with several graduate schools offering advanced degrees in the field. As of yet though, few psychology 101 courses have industrial/organizational topics in the syllabus.

3. A Modern Psychology Pedagogy

Given the foregoing traditional psychology pedagogy that has been focused on the four main disciplines within psychology, a more modern approach would be to further develop the curricula to include industrial/organizational topics. Industrial/organizational psychology is a growing field with significant interest, as can be seen from the growing number of graduate programs offering advanced degrees in industrial/organizational psychology. Incorporating more industrial/organizational psychology topics into the psychology 101 syllabus would therefore be a welcome addition for many students who might otherwise be unaware of the field.

Additionally, incorporating industrial/organizational psychology topics into the psychology 101 syllabus would allow for a larger discussion related to business functioning, and therefore allow a natural transition to include entrepreneurial learning.

3.1 The Source of Wealth: Intangible Versus Tangible

In order to explore the source of wealth we examine the factors that determine annual gross domestic product adjusted for purchasing power parity for all the countries of the world for which there are data (see Ridley, 2017a, 2017b; Ridley and Khan, 2019, Ridley and Llaugel, 2020). The G=f(C,D,R) model for year 2014 data and 79 countries that represent practically all the people in the world is reproduced in the appendix from Ridley (2020). The CDR epistemology comprises a regression model and corresponding vexillographical chart. A capitalist is a person who deploys his personal capital so as to maximize his benefit. Therefore, every rational person is a capitalist. C is the degree of capitalism practiced in a country. It includes exogenous human capital of imagination and creativity and that which was converted to endogenous capital stock (machines, devices, recordings, knowledge, etc) from prior years, less depreciation and obsolescence. Exogenous capital represents 85% of total capital. See Ridley and Khan (2019) for the first ever calculation of the value of ideas. Total capital is measured by total market capitalization. It is the amount of capital available for investment. An example of the organization of capital is Wall Street, United States of America (USA). It is not entirely clear that we can measure imagination and creativity in individuals. Still, we think we know expressions of human imagination and creativity when we see them. And, we expect that on the whole, psychology has a bearing on imagination and creativity (Naylor and Simonds, 2015) and therefore C. R is the degree of rule of law. Rule of law is the opposite of corruption and reflects the degree to which contracts are enforced and property rights are protected. Low property rights, especially corruption repels capital. R creates the stability that attracts capital. This is entirely psychological in that it appeals to trust and confidence in the enforcement of justice. D is the degree to which democracy is practiced and additional pathways are created so as to deploy C optimally. This is entirely psychological in that it appeals to trust and confidence in the partial ownership of ideas. The fitted CDR function is

$$\text{CDRindex} = 1.53C + 0.14D + 0.23R - 1.21C{\cdot}D{\cdot}R + 0.38N,$$

$$\text{where } G = \text{CDRindex}(\text{highest } G\text{-lowest } G) + \text{lowest } G,$$

$$\text{highest } G=\$83{,}066 \text{ and lowest } G=\$1{,}112.$$

This function serves as an index for computing *G* in any year for any country where *C*, *D* and *R* are known, and the highest and lowest *G* in the world are known. We refer to this model as the CDR model because *C*, *D* and *R* are policy variables and natural resources (*N*) is not. As shown in Figure 1 in the appendix, all the countries of the world fall approximately on a straight line. When natural resources and latitude (the absolute distance from the equator reflects the effect of geography) are included in the model, 90% of the variation in *G* is explained. Natural resources contribute 6% and latitude contributes 4%. These are negligible, unlike what is commonly believed to be true. See also, Ebrahim-zadeh, 2003, Auty, 1993, Sachs and Warner, 2001, Ross, 2001, Sala-i-Martin and Subramanian, 2003, Humphreys, 2005, Wadho, 2014 and Ridley, 2017b for discussions on how the natural resource curse can destroy an economy. The CDR contribution to explaining variation in *G* is (0.83-0.06)/0.06≅13 times the contribution from *N*. The residuals (not shown) are random and normally distributed, implying that there is no omitted variables bias.

The implication of the CDR model is that after adjusting for country factors of production, the conversion of *C* to *G* is global time invariant. That is, the same for all countries in all years. The conversion process is governed by the laws of natural science. *G* is not determined by high or low country productivity, but by the amount of capital a country attracts. What is often described as a country's high productivity is actually its ability to attract capital. That brings us to the observation that the only source of wealth is entrepreneurial human capital of imagination and creativity.

Ten percent of the people in the world are rich and ninety percent live on 2 to 3 US dollars per day. Almost invariable, the rich countries are high CDR rule of law democracies and the poor countries are low CDR corrupt dictatorships. See also Acemoglu, Johnson and Robinson (2005); Acemoglu and Robinson (2012); Acemoglu, Naidu, Restrepro and Robinson (2014). High CDR countries are where ideas go to fly. Low CDR countries are where ideas go to die. Even within the high CDR USA, there are students who lack an entrepreneurial family background. They think that the source of wealth is already established in existing factories that are places for the exploitation of workers. They see no relationship to their life. They are discouraged from entrepreneurship (see also Celuch, Bourdeau, Winkel (2017),Tognazzo, Gubitta and Martina (2016)). Adam Smith (1776) recognized that the real tragedy of the poor is the poverty of their aspirations. The CDR model establishes the source of wealth as human ideas of imagination and creativity (see also https://cdrindex.blogspot.com/). When exposed to entrepreneurship and related psychology, through the modified course, even the poorest students might see themselves as potential entrepreneurs (see also Ridley, 2017c). Many brilliant ideas have humble beginnings. Sometimes it is the people who no one imagines anything of, who do the things that no one can imagine!

3.2 Entrepreneurship: Information Theory of Psychology

It is common psychology to believe that the best ideas are easily communicated and understood. That they fit perfectly into the current knowledge base of facts and experiences. It turns out that the best ideas are entrepreneurial, surprising and unexpected. Unsurprising ideas that are expected cannot be innovative and therefore cannot add to economic growth. Psychologically, this is counterintuitive. In the CDR economic growth model *C* is converted to g in the presence of catalysts (Berzelius, 1835) *D* and *R*. Signals of ideas containing *C* must first be heard before they can be evaluated. This is very difficult when the channel of communication is noisy (Gilder, 2013; Romer, 1990). A noisy channel exists when *D* and *R* are low and the status quo of corruption and infighting prevails. Ideas go unnoticed. As *D* and *R* rise, noise drops, information signals of ideas containing *C* are communicated and received, *R* attracts *C* and *D* deploys *C* optimally.

Recent developments in thought pattern recognition by functional magnetic resonance imaging (fMRI) were reported on by Stahl (2019). The patterns are detected, analyzed and color diagramed on a computer display screen. One might think of this as a type of mind reading. The results are independent of native language or cultural background of the subject. There is a possibility that this type of communication when applied to international team members might be used to enhance entrepreneurship. It might also play a role in political elections. Politicians running for political office can be asked to include an fMRI in their medical checkup. Made available to the electorate, voters can make better informed choices, including the avoidance of corrupt dictators as defined by Ridley and de Silva (2019).

One of the obstacles to communications along potential pathways is legitimate logical and binary thinking, and group thinking. Music can connect that which people have in common. Music and sports can facilitate political and other problem solving by cutting through polarization and binary logic. People who might otherwise not think to associate might come together through music and sports. Consider also, the discovery by Hibbing, Smith and Alford (2014) that there are genes that make people predisposed to liberalism and other genes that make people predisposed to conservatism. In the advanced democracy of the USA, it is quite astonishing to find that these characteristics appear nearly equally bifurcated in the population. Due to the remarkably peaceful political party and presidential transitions, it is reasonable to assume that democracy is responsible for the meaningful deployment of liberal and conservative traits. And that both traits are required for the economic growth that we observe. In contrast, the absence of democracy could permit one of these two traits to dominate the economy. The outcome may be a reduction in diversity and the lack of economic growth that persists in poor countries.

3.3 Overview of How Psychology is Taught in The US

Although psychology originally developed in Europe, since the late 1800s it has become an international field of study. In 1881, a manuscript titled Project for an International Congress of Psychology was published, calling for an international meeting of psychologists (Sabourin and Cooper, 2014). Soon after, the First International Congress of Physiological Psychology was held in Paris in 1889 with participants from over 20 countries. This congress was one of the first international psychology conferences (Sabourin and Cooper, 2014). William James, the "Father of American Psychology" was one of the most influential American psychologists of his time attended this conference. William James was a Harvard professor and the first to offer a psychology course in the USA. Much of what was taught in this first course reflected the observational nature of psychology at the time, which was largely influenced by the many European thinkers that were present at the conference. Soon after this conference, in 1892, the American Psychological Association was founded. Still, psychology was very much a European field, heavily influenced by the many French, German and English psychologists who were popular at the time. After World War I, though, many prominent psychologists migrated to the USA, and the USA became a major hub for the field, growing even further post World War II. Following Freud's influential writings based on his psychological observations, empiricism began to grow with greater interest placed on the importance of research, as a means of testing the influential theories of the time. Such empiricism has continued to grow in the USA, which is reflected in the heavy reliance on empirical research in introductory psychology courses.

Psychology in the USA is often considered a social science, as the investigations into human thoughts and behavior are conducted using the scientific method. Graduate studies in psychology often have a strong research component, and research-based strategies and therapies are often seen as the gold-standard for treating psychopathology. In the clinical psychology realm, Cognitive-Behavioral Therapy, for example, is often seen as the most effective therapy for a variety of mental health disorders in the USA, due to the strong empirical research support for the therapy in diverse populations. In many other countries, such as France, however, observation-based psychology, akin to Freud's use of psychoanalysis, continues to be the most utilized form of clinical psychology.

The introductory psychology courses in the USA reflect the heavy reliance on empiricism that is found in the US in the field of psychology. Many introductory psychology courses contain a unit that is based on research methods and statistics. Psychology majors around the USA typically are required to take a full course on research methods and statistics before graduation. Similarly, various topics typically covered in the psychology 101 course, such as the brain, infant development, and language development, all are taught with numerous examples from research findings. Textbooks for introductory psychology courses also detail many examples from research conducted by other American psychologists.

3.4 Entrepreneurship Education for Children

Lifetime outcomes can begin with exposures at an early age. From an educational viewpoint, children from lower socioeconomic backgrounds often do not have the role models at home to strive for a more entrepreneurial way of learning. Parents who have not been exposed to entrepreneurial education themselves may not place value in it and that has a knock-on effect with their children, an epigenetic transgenerational sequela. Parents from low income backgrounds who may have had a poorer experience with education do not necessarily see why it is important to teach their children skills such as resilience and risk taking.

Consider for example, schools in the United Kingdom (UK). Children will often find it difficult to cope with making mistakes and thinking of ways to solve problems themselves. Statistically, children from lower income households also experience word poverty, a diminished lexicon. They are exposed to fewer words (Harley,2018; Hart and Risley, 2003). These children will be unable to converse successfully with others and will therefore be at a disadvantage from the start. Schools have to actively teach entrepreneurial skills while dealing with a widespread oracy problem in order to try to increase the equity between children from disadvantaged backgrounds and children from typical middle-class backgrounds.

However, if this support is not continued at home and parents do not hold these skills in high regard then their children will not see the value in it and the downward spiral will continue. They will likely be unable to access the skills needed to enter into higher levels of work or even be able to converse adequately at jobs. They will therefore only be able to enter into low paying, low skill job sectors. This can then impact future generations and the cycle continues. Broadening these disadvantaged students' entrepreneurial opportunities through greater access to education such as the university psychology 101 course may be beneficial. A university can help overcome these deficiencies in childhood exposure through high school community outreach activities such as a summer camp for entrepreneurship.

3.5 Revising the university psychology 101 course

In order to incorporate the new discovery of the CDR model (Appendix), it is necessary to revise the extant psychology curricula in a small number of ways. The heavy reliance on empiricism should remain, but there should be some additions to incorporate more business-related topics. Although only few, the implication of the revisions is profound. In order to fit the limited number of pages in this paper, only the overlapping topics from two syllabi for introductory psychology courses will be considered. The selected course is based on Psychology 101 courses taught at a small university in the Midwest and one in the Southern USA. Other courses can be revised similarly. Some items to include involve an introduction to industrial/organizational psychology, a greater focus on ethics, and lastly, a focus on society with discussion of the CDR model and the role of democracy. The original topics in the syllabi are listed in the left column in Table 1. The center and right columns list descriptions and reasons for the revisions, respectively.

Table 1. Introductory Psychology Course

Traditional Topics	***Proposed Topics in italics***	**Rationale for change/addition/removal**
1. Introduction to psychology -Definition of psychology -History of psychology -Overview of different areas in psychology	**1a. Introduction to psychology** -Definition of psychology -History of psychology -Overview of different areas in psychology ***1b. Modern trends in Psychology*** *-Industrial/Organizational psychology entrepreneurship*	Industrial/Organizational psychology is a growing field with many applications in business fields, that may be of relevance to a wide range of students. A discussion of industrial/organizational psychology is also an opportunity to discuss entrepreneurship, which is the process of starting a business, and the link between entrepreneurship and psychology.
2. Research Methods in Psychology -Validity and Reliability -Research Ethics	**2a. Research Methods in Psychology** -Validity and Reliability -Research Ethics ***2b. Applied ethics*** *-Ethics for justice in society*	How our understanding of ethics from the history of psychological research has changed over the decades, can inform how we interact with others in other realms beyond research. Implications can be seen in how businesses operate (industrial/organizational psychology) and in how we may promote ethical business practices in accordance with rule of law for the attraction of capital.
3. Statistics in Psychology	**No change**	
4. Nature and Nurture -Nature: heredity and evolution -Nurture: culture and society	**4. Nature and Nurture** -Nature: heredity and evolution -Nurture: culture and society *(Discussion of how environment and capitalism shape our thoughts and behaviors)*	The capitalist society in which we operate has a direct impact on how we function and develop. How we are shaped by society is also related to how entrepreneurship occurs. The environment that is not only conducive but required for entrepreneurship comprises capitalism (C), democracy (D) and rule of law (R).
5. Cognition -The brain and the central nervous system -Memory and Sleep	**No change**	
6. Human Development -Prenatal, infant, and early child development -Adolescence and adulthood	**No change**	
7. Emotional and Language development	**No change**	
8. Social Psychology -Social thinking, conformity, obedience,	**8a. Social Psychology** -Social thinking, conformity, obedience, group behavior, prejudice and discrimination	Understanding how others are influenced by social groups and the media can help students understand how businesses use psychology to become profitable. Such topics also invite

group behavior, prejudice and discrimination	***8b. Influencing others*** *-the role of democracy*	discussion for how entrepreneurs can identify and capitalize on social trends. Democracy for positive influence on entrepreneurial teams for the optimal deployment of capital should also be discussed.
9. Personality	**No change**	
10. Psychopathology - Mood and anxiety disorders, schizophrenia, eating disorders	**No change**	

4. Concluding Remarks

Excellence is never an accident. It is the result of high intention, sincere effort and intelligent execution. It represents choosing among a wide range of options. Choice, not chance determines one's destiny (Aristotle, 384-322BC). This leads us as university teachers to consider carefully the sequential consequences of our thoughts, words, actions, habits, character, and destiny (Lao Tze, 1601). Teachers touch the future into eternity. Any institution or activity that oppresses the creativity of children and young adults diminishes them and must destroy human capital, the only source of wealth. This is typical of the corrupt dictatorship (Ridley and de Silva, 2019) where new ideas are the natural born enemies of the way things are. This is potentially harmful to the entrepreneurial self-efficacy of future descendants for generations to come. This epigenetic transgenerational sequela is especially harmful if left untreated (Weber-Stadlbauer, 2017). And, people who are hurt, hurt people. A relatively low wealth community could persist. Modern psychology curricula that speak to this concern through incorporating greater focus on ethics as well as more business-related topics can do more good than harm. Incorporating more business-related topics with a greater focus on Industrial/Organizational psychology and entrepreneurial education may then have an increased societal benefit over the traditional psychology 101 curriculum.

Psychology and risk taking are related to wealth. A country that knows where it is going will only get that far. A country willing to create an entrepreneurial environment of risk taking and investment in the unknown may experience unlimited growth.

5. Nomenclature

Capitalist	A person who deploys his personal capital so as to maximize his benefit.
Capitalism	Mechanism for the collection and assembly of capital.
Catalysis	The creation of alternative pathways to enable a process.
CDR index	The vector inner product (dot product) of the global constant [1.53 0.14 0.23 -1.21] and the country [C D R C·D·R].
Company	The instrument of capitalism for the profitable investment of capital.
Democracy	Private work force idea participation and periodic election of public representatives (catalyst for the process of generating G from capital).
Endogenous	Generated from within a system.
Entrepreneurship	The process of starting a business, typically a startup company offering an innovative product, process or service.
Epigenetic	Relating to or arising from nongenetic influences on gene expression.
Epistemology	The investigation of what distinguishes justified belief from opinion.
Exogenous	Generated from outside a system.
Gross domestic product	The monetary value of all the finished goods and services produced within a country's borders in a specific time period.
Intrapreneurship	The employee practice of entrepreneurial activity inside a large business without incurring the associated risk.
Micro intrapreneurship	The low skill employee practice of micro entrepreneurship in variance reduction, quality improvement or customer relations at a business by virtue of proximity to a task.
Natural resource rents	Surplus value of natural resources after all costs and normal returns are accounted for.
Property rights	Property is a legal expression of an economically meaningful consensus by people about assets, how they should be held, used and exchanged.
Rule of Law	Reverse of corruption (protection of shareholder and other property rights) (catalyst for the attraction of capital).
Sequela	A condition which is the consequence of a previous disease or injury.
Transgenerational	Acting across multiple generations.
Virtue	Self-governing human property that promotes fairness and justice without the need for central government.

6. Appendix: The Source and Mechanism of Wealth

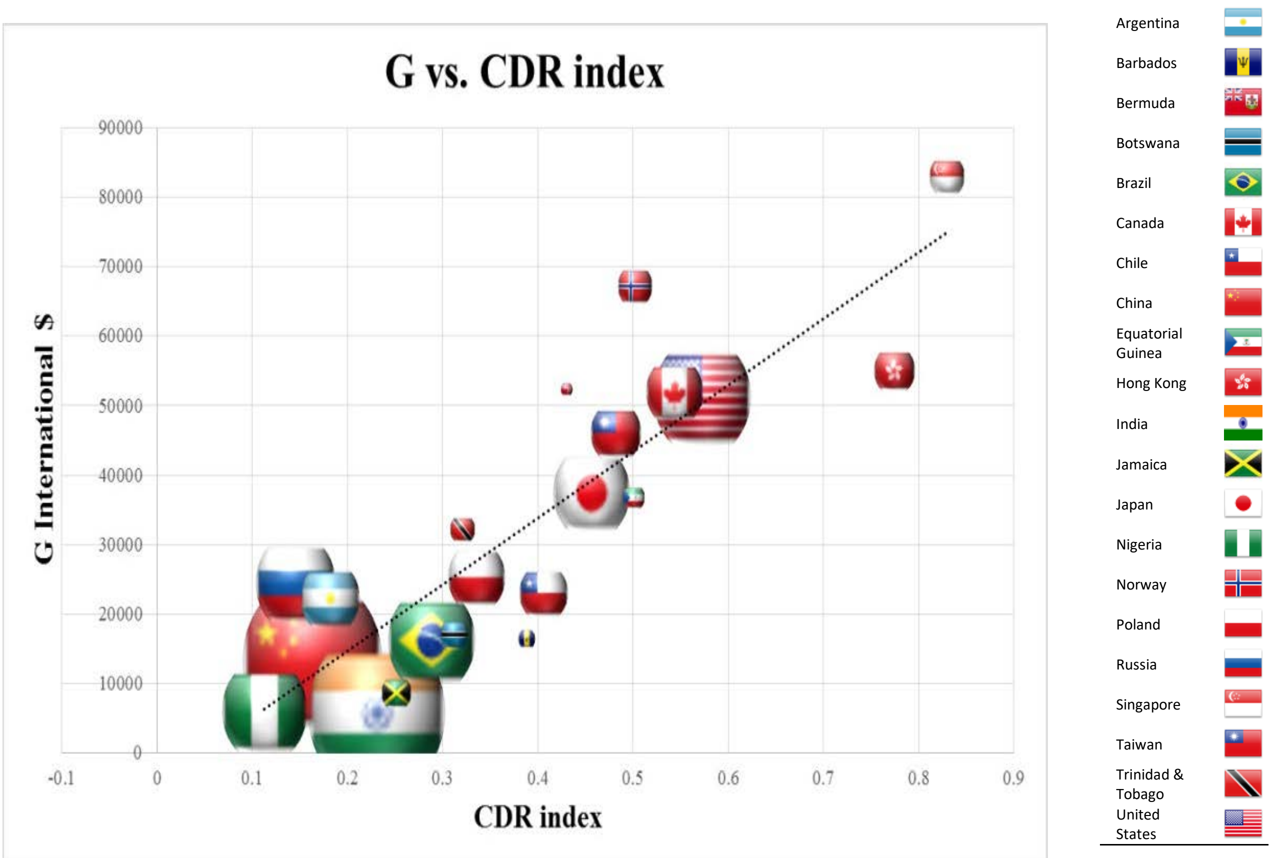


Figure 2. Year 2014 G vs CDR Index for 79 countries (line).

Bubble size (21 countries) is the square root of population. This model was re-estimated for years 1995-2016 with similar results. For additional comments on the countries see Ridley (2017a, 2017b).

Standardized g model

The ordinary least squares g model is specified as follows:

$$g = \beta_0 + \beta_C C + \beta_D D + \beta_R R + \beta_{CDR} C \cdot D \cdot R + \beta_N N + \varepsilon$$

where, the intercept β_0 and the coefficients β_C,β_D, β_R, β_{CDR},β_N are all dimensionless,
ε is a random, normally distributed error with a mean of zero and constant standard deviation,
and where all model variables are standardized as follows:

g $= \frac{G - \text{lowest } G}{\text{highest } G - \text{lowest } G}$

G = per capita real gross domestic product per capita (PPP)

C (Capitalism) $= \frac{\text{per capita capitalization} - \text{lowest per capita capitalization}}{\text{highest per capita capitalization} - \text{lowest per capita capitalization}}$

D (Democracy) $= \frac{\text{lowest democracy rank} - \text{democracy rank}}{\text{lowest democracy rank} - \text{highest democracy rank}}$

R (Rule of law) $= \frac{\text{lowest corruption rank} - \text{corruption rank}}{\text{lowest corruption rank} - \text{highest corruption rank}}$

N (Natural resources) $= \frac{\text{per capita total natural resource rents} - \text{lowest per capita total natural resource rents}}{\text{highest per capita total natural resource rents} - \text{lowest per capita total natural resource rents}}$

These transformations standardize the variables and ensures upper and lower bounds on $0 \le g, C, D, R, CDR, N \le 1$.

Democracy and corruption are rank ordered, where the highest = 1 and the lowest = the number of countries. *G* is measured in $/capita/year.

$\hat{g}$= 1.53*C* + 0.14*D* + 0.23*R* - 1.21*C·D·R* + 0.38*N*

t= (6.60) (1.69) (2.60) (4.40) (5.59) F ratio = 81.

Partial correlations (contributions to R^2_{adj}*):*

59% 5% 10% 3% 6% R^2_{adj} *= 83%.*

where ^ denotes estimated or fitted value and *G* can be estimated from

$\hat{G}$= $\hat{g}$ (highest *G*-lowest *G*) + lowest *G*.

*H*ighest *G*=83,066. Lowest *G*=1,112.

The CDRindex = 1.53*C* + 0.14*D* + 0.23*R* - 1.21*C·D·R* comprises positive *C*, *D* and *R* effects and a negative component due to friction from democracy that reduces *G* from what it might otherwise be if there were perfect agreement amongst decision contributors. The contribution from *N* is negligible and can be dropped from the model.

7. Disclaimer

Any opinions expressed are those of the authors and not of Nationwide Children's Hospital.

8. Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

9. Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

10. References

Acemoglu, D., Johnson, S., and Robinson, J. A. (2005). "Institutions as a fundamental cause of long-run growth." Handbook of economic growth, 1A, 385-472.

Acemoglu, D., and Robinson, J. A. (2012). Why Nations Fail. Crown Publishers, Random House, NY.

Acemoglu, D., Naidu, S., Restrepro, P., and Robinson, J. (2014). Democracy Does Cause Growth. NBER working paper #20004, National Bureau of Economic Research, 1-64.

Auty, R. (1993). Sustaining Development in Mineral Economies: The Resource Curse Thesis. London: Routledge.

Berzelius. J. J. (1779–1848). https://en.wikipedia.org/wiki/J%C3%B6ns_Jacob_Berzelius.

Deal, K. H. (2003). "The Relationship Between Critical Thinking and Interpersonal Skills: Guidelines for Clinical Supervision." The Clinical Supervisor, 22(2), 3–19.

Ebrahim- zadeh, C. (2003). "Back to Basics: Dutch Disease. Too Much Wealth Managed

Unwisely." Finance and Development. 40(1), 50–51.

Naylor, E., R. and Simonds, J. (2015), "Adult Imagination: Development and Psychometric Examination of the Imaginative Involvement Questionnaire." Imagination, Cognition and Personality: Consciousness in

Theory, Research, and Clinical Practice, 35(2), 112-136. DOI: 10.1177/0276236615608751.

Gilder, G. (2013). Knowledge and Power: The Information Theory of Capitalism and How it is Revolutionizing our World. Regnery Publishing, Inc., Washington, DC.

Harley, J. (2018). "Why Closing the Word Gap Matters: Oxford Language Report," Oxford University Press. https://www.oup.com.cn/test/word-gap.pdf.

Hart, B. and Risley, T. R. (2003) "The Early Catastrophe: The 30 Million Word Gap by Age 3," American Educator. https://www.aft.org/sites/default/files/periodicals/TheEarlyCatastrophe.pdf.

Hibbing, J. R., Smith, K. B. and Alford, J. R. (2014) Predisposed: Liberals, Conservatives, and the Biology of Political Differences, Routledge, N. Y.

Llaugel, F., and Ridley, A. D. (2018). "Entrepreneurial Engineering: revising the engineering 101 course." International Journal of Economics, Business and Management Research, 2(5), 258-271.

Luryi, S., W. Tang, N. Lifshitz, G. Wolf, S. Doboli, A. Betz, Maritato, P. and Shamash, Y. (2007). "Entrepreneurship in Engineering Education". 37th Annual Frontiers in Education Conference-Proceeding. 10-13 Oct.

Ngnepieba, P., Ridley, A. D., Stephens, D., Johnson, L., and Edington, M. (2018). "Entrepreneurial Mathematics: revising the math101 course." International Journal of Economics, Business and Management Research, 2(5), 272-286. http://ijebmr.com/uploads2018/IJEBMR_02_287.pdf

Ridley, A. D. (2017a). "Wealth is all in the mind." International Journal of Social Science and Economic Research, 2(5), 3369-3379. http://ijsser.org/uploads/ijsser_02__208.pdf

Ridley, A. D. (2017b). "Division of Human Capital creates Surplus Wealth." Bulletin of Political Economy, 11(1), 1-20.

Ridley, A. D. (2017c). "Micro Intrapreneurship." International Journal of Social Science and Economic Research, 2(9), 4729-4743. http://ijsser.org/uploads/ijsser_02__299.pdf.

Ridley, A. D. (2018). "Entrepreneurial Economics: revising the econ 101 course." International Journal of Economics, Business and Management Research, 2(5), 183-207. http://ijebmr.com/uploads2018/IJEBMR_02_280.pdf

Ridley, A. D. and de Silva, A. (2019). "Game Theoretic Choices between Corrupt Dictatorship Exit Emoluments and CDR benefits: is there a Nash Equilibrium?" The American Economist, 65(1), 1-27. https://doi.org/10.1177/0569434519878858.

Ridley, A. D. and Khan, A. (2019). "Decoupling Entrepreneurship Capital from Capital Stock," Bulletin of Political Economy, 13(1), 77-86.

Ridley, A. D. (2020). "Capitalism/Democracy/Rule of law Interactions and Implications for Entrepreneurship and per capita real Gross Domestic Product adjusted for purchasing power parity," Journal of the Knowledge Economy, DOI: 10.1007/s13132-020-00632-6.

Ridley, A. D. and Llaugel, F. (2020). "Advances in the CDR economic theory of entrepreneurship and GDP," forthcoming.

Romer, P. M. (1990). "Endogenous Technological Change," Journal of Political Economy, 98(5), S71–S102.

Ross, M. (2001). "Does Oil Hinder Democracy?" World Politics. 53(3), 326–361.

Sachs, J., and Warner, A. M. (2001). “Natural Resources and Economic Development: The Curse of Natural Resources.” European Economic Review. 45(4–6), 827–838.

Sala- i-Martin, X., and Subramanian, A. (2003). “Addressing the Natural Resource

Curse: An Illustration from Nigeria.” NBER Working Paper No. 9804.

Schumpeter, J. A. (1928). The Entrepreneur. in M. Becker, T. Knudsen, and R. Swedberg

(eds.) 2011. The Entrepreneur: Classic Texts by J. Schumpeter. Stanford, California:

Stanford University Press, 227-260.

Smith, A. (1776, 2010). An Inquiry into Nature and Causes of the Wealth of Nations. Tantor.

Stahl, L. (2019). https://www.cmu.edu/news/stories/archives/2019/november/just-to-appear-on-60-minutes.html.

Sabourin, M., & Cooper, S. (2014). "The first International Congress of Physiological Psychology (Paris August 1889): The birth of the International Union of Psychological Science." International Journal of Psychology, 49(3), 222-232.

Wadho, W. A. (2014). “Education, Rent seeking and the Curse of Natural Resources.” Economics and Politics, 26(1),128–156. DOI: 10.1111/ecpo.12029.

Weber-Stadlbauer, U. (2017). “Epigenetic and transgenerational mechanisms in infection-mediated neurodevelopmental disorders,” Translational Psychiatry, 7(5), e1113, doi: 10.1038/tp.2017.78

A SYSTEMS ENGINEERING FRAMEWORK FOR INTEGRATION OF COGNITIVE ERGONOMICS IN SMART FACTORY CONFIGURATIONS

Lauren Drankoff[1]
Sharon Bommer, Ph.D.[1]
[1]University of Dayton
Drankoffl1@udayton.edu

Abstract

The current manufacturing industry is increasingly compelling manufacturing firms to turn to technology to improve operations. A relatively unexplored area is the role that cognitive ergonomics with the effective implementation of technology has on smart factory configurations. The manufacturing industry continues to increase the information processing content operators engage in (i.e., multitasking) to perform their work tasks. Cognitive ergonomics are known predictors of preventing injuries and accidents in industry and drive towards improving work conditions, human performance, safety, and avoiding human errors. Cognitive ergonomics evaluates the interaction of the human operator with the task they are performing and their mental capabilities. Although these principles are currently being explored in manufacturing, it is still uncertain how cognitive ergonomics can be applied to smart factory configurations. A smart factory combines technology to create a network that has the capability to communicate and act independently to advance manufacturing systems further. Smart factories can assist people and machines by interacting with the environment surrounding to make tasks more comfortable and more efficient. By investigating current literature on cognitive ergonomics, the goal of this work is to pave the way for further research into human-systems integration of technology implementation in smart factory configurations. The technology methods discussed will seek to aid the human operator in the decision-making process throughout their work task. Also, a framework is developed for the integration of cognitive ergonomics in smart factory configurations using the D-E-J-I systems engineering model. In summary, this paper aims to investigate how current cognitive ergonomic principles can be utilized to support the effective implementation of technology to aid the human operator in smart factory configurations.

1. Introduction

In this paper, the use of current cognitive ergonomic principles that can be used to support the effective implementation of technology to aid the human operator in smart factory configurations will be investigated. As new technology arises in the manufacturing industry, cognitive capabilities are becoming an essential work element creating the need to gain further research on human cognitive processes (Bouargane, 2015). Cognitive ergonomic principles can be applied to smart factories in manufacturing. A smart factory combines technology to create a network that has the capability to communicate and act independently to further advance manufacturing systems. They can assist people and machines by interacting with the environment surrounding to make tasks easier and more efficient. The elements smart factories focus on are execution, maintenance, and education functions of manufacturing. Smart factories provide flexible and reconfigurable systems, which are low cost operations, adaptive, and lean. They are still a work in progress and remain in the research investigation stage. Cognitive ergonomics are an essential element to support effective technology implementation to aid the human operator in smart factories to improve performance and reduce mental workload in effort to improve performance (Radziwon et al, 2014). By investigating current literature on cognitive ergonomics, the goal of this literature survey is to pave the way for further research into human-systems integration of technology implementation in smart factory configurations.

2. The need for cognitive ergonomics in manufacturing

2.1. What is cognitive ergonomics?

Cognitive ergonomics are known to be related to the human operator and the task they are performing. It is the human operator's mental processes as they affect work interactions. The mental processes associated with cognitive ergonomics are perception, memory, reasoning, and motor response (Kim, 2016). The goal of cognitive ergonomics is to design tools and systems to support the human operators' mental processes to improve situation awareness, problem-solving, and decision making.

A human operator's mental model allows them to construct a system based on how it functions, providing the operator an understanding of how the process is accomplished (Bouargane, 2015). The basic principle of cognitive ergonomics is that cognitive load should not exceed the human cognitive ability. To manage cognitive workload and increase human reliability, it can provide a user-centered design, introduce a safer design of tools, and implement information technology systems that support cognitive processing. An example application of cognitive ergonomics would be designing a software interface to be easy to use or with controls that prevent catastrophic human errors. They can prevent industry accidents and can prevent injuries from cognitive failures (Nobre F et al., 2008). When used in industry, cognitive ergonomics improves work conditions, human performance, safety, health, and work to avoid human error. Manufacturing plants with mechanization introduce a more significant risk due to increased decision making. This leads to the need for increasing safety to minimize any human errors (Kim, 2016). Cognitive ergonomics can be applied in various methods that focus on user-centric knowledge. These include advanced knowledge acquisition and design, cognitive interviewing, cognitive task analysis, and group task analysis. These methods lead towards solutions that combine cognitive, system design, and practical engineering concepts to aid the human operator to ensure good usability (Berry et al, 2015).

2.2. Cognitive ergonomics in manufacturing

The future needs of manufacturing systems are not being met using the current science principles available. A key element to manufacturing systems is the human operator who may experience cognitive failures during a work task. A cognitive failure is an error that occurs throughout a work task that a worker could perform without error (Kim, 2016). A human operator is continuously facing a variety of tasks involving design, inspection, planning, scheduling, and decision- making principles that require additional human intelligence. A few examples of operator human error include inappropriate use of equipment, pushing the wrong button on a machine, leaving a valve open on a machine, or experiencing a machine maintenance failure. All of these human errors can lead to injuries, scrapped parts and material, or extended machine downtime (Burgess, 2007). Cognitive ergonomics would allow higher degrees of info processing to aid an operator (2008, Nobre et al). Solutions for operators with cognitive disorders such as failures in visual and hearing functions could be implemented using cognitive ergonomics (Kim, 2016). Cognitive ergonomics can be used to identify cognitive failures and prevent human error in the workplace. Current manufacturing systems lack the connection between their environmental complexity and their cognitive capacity. Technology and increased cognitive capability would improve this connection. Cognitive ergonomics can locate sources of human errors to support technology systems to assist the human operator in smart factory configurations. It is necessary to locate the human error and manage the cognitive load of the operator to increase productivity and maintain a safe work environment in manufacturing. The knowledge of cognitive ergonomics in manufacturing is limited due to challenges of developing methods for reducing accidents and injuries in industry using cognitive ergonomic principles (Kim, 2016). This paper seeks to outline how human-systems integration of technology implementation can benefit smart factory configurations.

Although the industry has limited research, cognitive ergonomics has currently been used to test mental workload, human performance, and cognitive load in manufacturing as outlined in Table 1 in six different studies. Using a physiological method such as eye tracking in study 1 can provide information on how a user interacts with a visual display to record if a user looks directly at a feature, ignores it, or has delayed visual timing. Recently, eye tracking has been used to estimate cognitive workload through the changes in pupil dilation of the human operator. By using an index of cognitive activity to measure abrupt changes in pupil dilation, cognitive effort can be detected (Marshall, 2002). Study 2 uses cognitive ergonomics to assess mental workload by taking EEG recordings, monitoring reaction times, and monitoring unrelated movements from a task for manual assembly work. This type of work can be repetitive throughout the day and contain a low mental workload which can be dangerous for the operator. By recording the three physiological methods, mental workload changes can be tracked and measure the cognitive influence it has on the worker (Mijovic et al, 2017). Study 3 indicates a subjective approach for measuring the effect that quality of work-life and cognitive demands has on human performance. The study simulates human operators assembling mechanical and electrical equipment and performs a task analysis using four index variables and conducts a quality of work-life questionnaire. The results suggest human performance is correlated with cognitive demand when performing a task as well as quality of work-life attributes. Cognitive ergonomics can aid in monitoring and increasing human performance when cognitive load is examined (Layer et al, 2009). Study 4 also uses cognitive ergonomics to assess the cognitive load of the operator. Human-machine interface in manufacturing presents the challenge of information overload for the operator. The study tracked eye movement and recorded workers feelings about workload to reveal difficulties when finding a visual button. By performing subjective and physiological approaches using the National Aeronautics and Space Administration Task Load Index (NASA-TLX), the Questionnaire for User Satisfaction (QUIS) and eye tracking, the influence of information overload on the user's cognitive load can be tracked. Cognitive load tracking can reveal how to limit cognitive overload and increase user efficiency (Wu et al, 2016). Study 5 uses a subjective approach using the NASA-TLX for monitoring cognitive load in manual production environments. It focuses on the cognitive limitations that impact production outcome in a manufacturing environment by seeking to develop a method for assessing cognitive burden throughout a task. By developing an assessment tool, cognitive load can be measured and used to prevent cognitive overload. Cognitive ergonomics can help prevent cognitive overload from occurring and limit human operator error (Lindblom et al, 2014). Study 6 combines subjective, physiological, and performance approaches using the NASA-TLX, workload profile, fixation duration, and human error probability to predict human performance and evaluate MWL in manufacturing. The experiment combines the approaches to measure task complexity and monitor resource demands for the simulated tasks. By using more than one approach, more data is created and a framework for measuring mental workload is produced (Bommer, 2016). Mental workload, human performance, cognitive overload, and human error can all be assessed using cognitive ergonomics to measure the human operators' cognitive processes when performing work tasks in manufacturing.

3. Define smart factory

The smart factory industry is a recent component of manufacturing that continues to develop. A smart factory is a collection of technologies that develop strategic innovation for existing product development through combining humans, processes, technology, and information in manufacturing

Table 1. Cognitive ergonomic methods

Study	Author	Year	Lab or Field	Task	Cognitive Ergonomic Method
1	Marshall, S.	2002	Lab	Simulated cognitive tasks involving problem solving and visual search	Physiological method examining pupil dilation to estimate the level of cogntive effort on the user
2	Mijović, P. Milovanovic, M. Kovic, V. Gligorijevic, I. Mijovic, B. Macuzic, I.	2017	Lab	Simulated manual assembly	Physiological method examining mental work load through reaction times, EEG, and the number of task unrelated movements
3	Layer, J. Karwowski, W. Furr, A.	2009	Lab	Simulated assembling electrical and mechanical equipment	Subjective method measuring the effect of quality of work life and cognitive demands on human performance through a quality of work life questionaire and a cognitive task analysis of 4 index variables
4	Wu, L. Zhu, Z. Cao, H. Li, B.	2016	Lab	Simulated performing an interface task for an automated sorting system involving three visual search tasks	Subjective and physiological method examining human machine interface influence of information overload on the user's cognitive workload through eye tracking and a NASA-TLX and QUIS questionaire
5	Lindblom, J. Thorvald, P.	2014	Lab	Simulated workers performing various tasks at a work station	Subjective method using a NASA-TLX to report on cognitive load in manual production environments
6	Bommer, S. Fendley, M.	2016	Lab	Simulated workers performing cleaning inspection and final inspection tasks at a work station	Combines subjective, physiological, and performance methods using NASA-TLX, workload profile, fixation duration and human error probability to predict human performance and evaluate MWL in manufacturing

(Wiktorsson et al, 2018). The goals of a smart factory seek to respond in real time to meet the demands in the factory, supply network, and customer needs using an integrated and collaborative manufacturing system (Wiktorsson et al, 2018). They also can provide flexible processes through the use of automation to create and optimized manufacturing system that reduces labor and resources (Radziwon, 2014).

A smart factory example in Sweden shows how smart factory technology can be applied in the manufacturing industry (Wiktorsson et al, 2018). In this study, the bus and engine manufacturer focus on digitalization for the smart truck and bus production lab. To aid in production and collect data in real time, collaborative robots, automotive guide vehicles, hand tools and various sensors are used to control equipment. The Swedish company developed a smart factory pyramid that describes the goals of how an ideal smart factory would perform. The levels include a standardized process, connected technology, data gathering, data analysis, predictable future, and prescribe action as the highest level. Based on the Swedish smart factory, four implementation areas of smart factory integration with corporate production systems, key application areas within core value-adding processes, transformational inabilities of legacy systems, and smart factory for a sustainable production should be considered. For a company to achieve smart factory capability, the main goal is technology integration and digitalization into their production systems to achieve increased efficiency and cost reduction (Wiktorsson et al, 2018).

A smart factory can also be used in a hyper-connected model that focuses on integration of the production system, integration of the product life cycle, and integration of the inter-company value chain. This is shown in a South Korean manufacturing plant that uses these integrations to aid customized production, optimized production through information exchanges, and on-time production. The information obtained for the manufacturing facility is done in real-time through the Internet allowing continuous changes to be made in the manufacturing method and implements an optimized production system. Hyper-connected technology in smart factories also supports rising technologies such as robots and artificial intelligence support tools. Sensors are used for production control and quality management which implemented in robots. Artificial intelligence software has developed for the use of unmanned vehicles to find optimized paths in warehouses. (Park, 2016).

Smart factories use advanced technology to create an optimized manufacturing system. In the two smart factory studies discussed in this section, they used digitalization, robots, sensors, and artificial intelligence to achieve increased efficiency and improve production processes. As these technologies continue to influence how human operators perform their jobs, it is essential to understand how technology can aid the human operator in smart factory configurations. Cognitive ergonomics should be considered when integrating technology into smart factories to evaluate the human operators' mental capabilities and ensure they can optimally perform their task. Human performance, mental workload, and cognitive load can be assessed to aid the human operator with technology integration.

4. Cognitive machines

As explained in Hu Jin's data processing article, cognitive machines form the technological foundations for artificial intelligence to operate in manufacturing settings. Cognitive computing simulates human thought to function as the human brain would operate using algorithms and pattern recognition. In order to achieve this high-level human functioning in a computer format, cognitive computing operates with continuous learning behavior and interactive ability through data processing. It is important to recognize that cognitive computing is not meant to replace all human functions but is used for broad-scale application and problem-solving. Cognitive machines operate with assistance, understanding, decision-making, and cognition and discovery processes. Machine learning operates with statistical learning and optimization theory born into advanced computers. An algorithm that is commonly used for cognitive computing is the Bayesian estimation for decision making. In a data processing example, cognitive computing is used to eliminate the challenges of decision-making through the mass amounts of information available. A cognitive decision algorithm is developed based on a deep confidence network and linear perceptron. The algorithm operates by using a decision model to establish what information is useful (Jin, 2019). An example application of cognitive computing in manufacturing is for enabling high flexibility and high variability in production. To enable this production model, a software can be used to provide a semantic level of information for the manufacturing system to plan for machining and inspection feedback using cognitive capability to adjust to any unforeseen changes during the production process (Zhao, 2010). As cognitive computing continues to develop, humans and machines are merging into one integrated component. Human, machine, and software agents form together to create hybrid agents that enhance control of a manufacturing system (Jones, 2018). Although cognitive computing research has been developing since the 1980's, there is limited recent literature available, and its' use in manufacturing is a relatively unexplored topic. The cognitive machine and cognitive computing interface could be used in manufacturing to ease cognitive load off the operator and eliminate human error that develops from cognitive overload.

5. Artificial intelligence in manufacturing

Smart factories heavily rely on sensors, computing, and artificial intelligence to optimize cost, labor use, and product quality in manufacturing. Artificial intelligence allows the computing infrastructure to be guided by data collected throughout the manufacturing process to make decisions. An example of the development of an artificial intelligence tool is the intelligent tool condition monitoring system (Lee et al, 2019). The tool seeks to decrease the amount of waste produced throughout the machining process by monitoring the condition of the tool used in manufacturing. The construction of the artificial intelligence tool requires sensing technologies such as vibration to collect data on the condition of the machine tool throughout the manufacturing process. By using artificial intelligence principles such as artificial neural networks, support vector machines, and fuzzy systems, predictions on the tool condition can be made. This tool allows optimal machining conditions to be identified (Lee et al, 2019).

Although the intelligence tool leads to more sustainable manufacturing and decreases waste production through tool condition monitoring, the article fails to discuss how cognitive ergonomic principles could complement the tool in a smart factory setting. The tool itself attracts cognitive ergonomics because it would allow the human operator to decrease human error by making more accurate cuts through monitoring the tool condition and ensuring every cut produces a quality part. However, an intelligence tool like this could cause a mental overload for the operator if the interface is too complex. To integrate an artificial intelligence model such as a tool monitoring system, the interface should be tested using subjective, physiological, and performance cognitive ergonomic methods to ensure the human operator would be successful when using the intelligence device.

Unsupervised learning algorithms can also be applied in manufacturing for the use of setup for machining. The algorithm combines intersecting and nonintersecting features for a machine setup and classifies them in its memory. The best tool sequence is formulated from the algorithm to create the desired setup (Chen, 1994). A semi-supervised example detects faults in additive manufacturing products using a machine learning algorithm. It uses data during training from additive product components that were quality certified and from components that had unknown quality. This decreases quality inspection time and is more cost efficient (Okaro, 2019). A reinforcement learning example uses part-driven distributed learning arrival time control and machine-driven distributed reinforcement learning based control. This enables part controllers to adjust their parts' arrival time while searching for optimal dispatch to achieve just-in-time production. In order to implement learning algorithms in smart factories, the most important element is proper data collection and analysis in a manufacturing system (Hong, 2004).

An additional intelligence example in smart factory configurations is deep learning algorithms (Wang et al, 2017). Deep learning automatically processes data by understanding speech, image recondition, and natural language processing. These automatic processing features allow advanced analytics tools to be available for smart manufacturing. Deep learning is beneficial because it is flexible and adaptable to data which creates optimized smart facilities by reducing production costs, reducing downtime, and creating a flexible environment. An application of a deep learning algorithm would be using the convolutional neural network which uses descriptive analytics for product quality inspection. Prior to deep learning, insufficient inspection defects would occur because the introduction of new products created different surface textures and variations. The convolutional neural network allows an automated defect detection process for inspection that identifies changing defect needs complementing the introduction of new parts or part variation (Wang et al, 2017). Deep learning in smart manufacturing is still a developing research topic but exemplifies how to use and collect data to benefit production.

6. DEJI model for cognitive ergonomic integration with technology for smart factories

The Design-Evaluate-Justify-Integrate (DEJI) systems engineering model is composed of four phases that examines the design, evaluation, justification, and integration process at each stage of system development (Badiru, 2018). The DEJI model assists with facilitating an understanding of the system, diagnosing problems in the systems and creating a systems profile. Figure 1 shows the framework for the DEJI model (Badiru et al, 2017). The model in Table 2 seeks to represent a process for using cognitive ergonomic principles with technology for smart factories in manufacturing using the DEJI model. Table 2 was created based off the cycle in Figure 1 which shows the framework for the DEJI model (Badiru et al, 2017). Most cognitive ergonomic principles in technology are relatively unexplored and the DEJI model in Table 2 is developed to guide implementation.

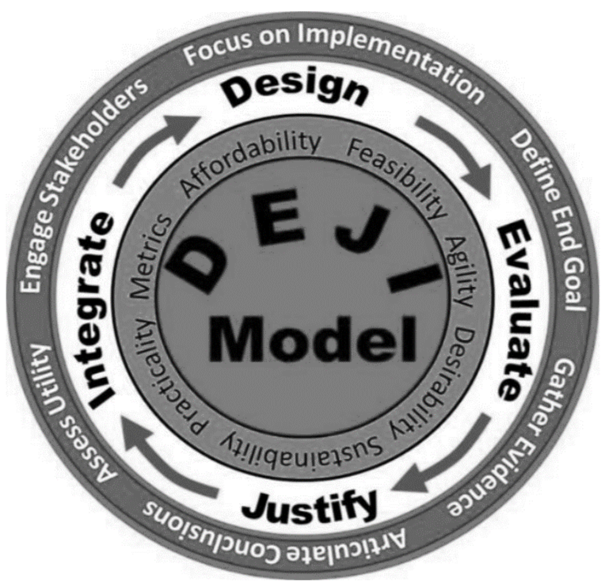


Figure 1. DEJI model: design, evaluate, justify, integrate

The design phase involves selecting the goals of the smart factory, establishing a team, deciding which cognitive ergonomic principles will be targeted, and analyzing a task a human operator performs. The model then goes into the evaluation phase where data is collected, reviewed, and design improvements are identified. The justification phase involves validating the program efforts according to the data collection in the evaluation phase and a technology review. The final integration phase develops a plan, implements technology changes, and conducts training for the human operators using the technology.

Table 2. DEJI model for cognitive ergonomics in technology

DEJI Model	Step	Task
Design	1	Assess smart factory goals
	2	Establish the team that will facilitate the program
	3	Review cognitive ergonomic principles that will be targeted
	4	Analyze a task the human operator performs in manufacturing
Evaluate	5	Collect data on task time, mental load, human error, technological error, etc.
	6	Run cognitive ergonomic tests while operators perform tasks (i.e. NASA-TLX)
	7	Review results to see where there is high mental workload
	8	Review user centric design with human operator and engineers
	9	Identify where the design could be improved
Justify	10	Validate program efforts according to data
	11	Conduct technology review for how it could be implemented to the design
Integrate	12	Develop a continued integration plan
	13	Implement technology changes
	14	Conduct training/ kick-off for human operators using the technology

7. Conclusion and discussion

Cognitive ergonomics and technology implementation in smart factories has limited research available and lacks any implementation studies. The use of technology in smart manufacturing is essential to achieve maximum optimization. The use of cognitive ergonomics would help to identify the human operator's mental workload and lead to a design creation where technology can be implemented in their work task to assist in managing mental overload. This would help decrease injuries in the workplace, increase productivity, and increase profit. Smart factories can assist people and machines by interacting with the environment surrounding to make tasks easier and more efficient. As smart manufacturing continues to evolve, there is an increasing integration of artificial intelligence applications such as sensors that allow extensive data to be collected. As more companies implement software to examine the data they collect throughout their manufacturing facility, it is essential that it can be used to increase optimization. All optimization principles often circle back to the human operator and depend on if they can successfully complete their work task and how to aid them to decrease error and task time. Examining cognitive ergonomic principles allows them to be used with technology to aid the human operator. By investigating current literature on cognitive ergonomics, the goal of this work is to pave the way for further research into human-systems integration of technology implementation in smart factory configurations.

The six articles listed in Table 1 show cognitive ergonomic assessment methods used in manufacturing that can assess mental workload, human performance, cognitive overload, and human error to measure the human operators' cognitive processes when performing work tasks. The technology methods discussed will seek to aid the human operator in the decision-making process throughout their work task. The DEJI model in Table 2 provides an overview to implementing cognitive ergonomic principles with technology for smart factories in manufacturing. By following Table 2, a process could be implemented into a manufacturing task to improve it by monitoring a human operator's mental workload using cognitive ergonomic principles and determining where technology integration could improve the task. This paper aims to investigate how current cognitive ergonomic principles can be used to support the effective integration of technology to aid the human operator in smart factories. In summary, the literature survey conducted was studied to propose a systems engineering approach for the integration of cognitive ergonomics in smart factory configurations.

8. References

Bouargane, L., Cherkaoui, A. (2015). Towards an explicative model of human cognitive process in a hidden hazardous situation and a cognitive ergonomics intervention in railway environment. IEEE: International Conference on Industrial Engineering and Systems Management, 968-976.

Badiru, A., Bommer, S., Grimes, A. (2018). Applications of DEJI Systems Engineering Model in the Development of a New Faculty Member Program in Higher Education. American Society for Engineering Education Annual Conference and Exposition. 1-11.

Badiru, A., Bommer, S. (2017). Work Design: a systematic approach. Boca Raton, FL: Taylor & Francis Group, 7.

Burgess, P., Ruckart, P. (2007). Human error and time of occurrence in hazardous material events in mining and manufacturing. Journal of Hazardous Materials, 142, 3, 747-753.

Berry, N., Prugh, W., Lunkins, C., Vega, J., Landry, R., Garciga, L. (2015). Selecting Video Analytics using Cognitive Ergonomics: A Case Study for Operational Experimentation. Procedia Manufacturing, 3, 5245-5252.

Bommer, S., Fendley, M. (2016). A theoretical framework for evaluating mental workload resources in human systems design for manufacturing operations. International Journal of Industrial Ergonomics, 63, 7-17.

Bouargane, L., Cherkaoui, A. (2015). Towards an explicative model of human cognitive process in a hidden hazardous situation and a cognitive ergonomics intervention in railway environment. IEEE: International Conference on Industrial Engineering and Systems Management, 968-976.

Chen, P., LeClair, S. (1994). Integration of design and manufacturing: solving setup generation and feature sequencing using an unsupervised-learning approach. Computer-aided Design. 26, 1, 59-75.

Hong, J., Prabhu, V. (2004). Distributed Reinforcement Learning Control for Batch Sequencing and Sizing in Just-In-Time Manufacturing Systems. Applied Intelligence, 20, 71-87.

Jin, H. (2019). Data processing model and performance analysis of cognitive computing based on machine learning in Internet environment. Soft Computing, 23, 9141-9151.

Jones, A., Romero, D., Wuest, T. (2018). Modeling agents as joint cognitive systems in smart manufacturing systems. Manufacturing Letters, 17, 6-8.

Kim, IJ. (2016). Cognitive Ergonomics and Its Role for Industry Safety Enhancements. Journal of Ergonomics, 6, 1-3.

Layer, J., Karwowski, W., Furr, A. (2009). The effect of cognitive demands and perceived quality of work life on human performance in manufacturing environments. International Journal of Industrial Ergonomics, 39(2), 413-421.

Lee, W., Mendis, G., Sutherland, J. (2019). Development of an Intelligent Tool Condition Monitoring System to Identify Manufacturing Tradeoffs and Optimal Machining Conditions. Procedia Manufacturing, 33, 256-263.

Lindblom, J., Thorvald, P. (2014). Initial development of a cognitive load assessment tool. 5th Applied human factors and ergonomics conference, 6256-6266.

Marshall, S. (2002). The Index of Cognitive Activity: measuring cognitive workload. IEEE: 7th Conference on Human Factors and Power Plants, 1-7.

Mijović P., Milovanović M., Ković V., Gligorijević I., Mijović B., Mačužić I. (2017) Neuroergonomics Method for

Measuring the Influence of Mental Workload Modulation on Cognitive State of Manual Assembly Worker. International Symposium on Human Mental Workload: Models and Applications, 213-224.

Nobre, F., Tobias, A., Walker, D. (2008). The pursuit of cognition in manufacturing. Journal of Manufacturing Systems, 27(4), 145-157.

Okaro, I., Jayasinghe, S, Sutcliffe, C, Black, K. Paoletti, P. Green, P. (2019). Automatic fault detection for laser powder-bed fusion using semi-supervised machine learning. Additive Manufacturing, 27, 42-53.

Park, H., Kang, H., Lee, J., Choi, S., Kim., H, Son, J., Kim, B., Noh, S. (2016). Smart Manufacturing: Past Research, Present Findings, and Future Directions. International Journal of Precision Engineering and Manufacturing-Green Technology, 3(1), 111-128.

Radziwon, A., Bilberg, A., Bogers, M., Madsen, E. (2014). The Smart Factory: Exploring Adaptive and Flexible Manufacturing Solutions. Procedia Engineering, 69, 1184-1190.

Wang, J., Ma, Y., Zhang, L., Gao, R., Wu, D. (2017). Deep learning for smart manufacturing: methods and applications. Journal of Manufacturing Systems, 48, 144-156.

Wiktorsson, M., Do Noh, S., Bellgran, M., Hanson, L. (2018). Smart Factories: South Korean and Swedish examples on manufacturing settings. Procedia Manufacturing, 25, 471-478.

Wu, L., Zhu, Z., Cao, H., Li, B. (2016). Influence of information overload on operator's user experience of human-machine interface in LED manufacturing systems. Cognition Technology & Work, 161-173.

Zhao, Y., Xu, X. (2010). Enabling cognitive manufacturing through automated on-machine measurement planning and feedback. Advanced Engineering Informatics, 24, 3, 269-284.

ASSESSING THE INTERNET SERVICE PROVIDERS IN EGYPT

Khaled S Abdallah[1]
Yasmin Adel[2]
[1]Arab Academy for Science, Technology and Maritime Transport (AASTMT), College of International Transport and Logistics- Cairo
[2]Production, Energy, and Automatic control Department, The French University in Egypt
Khaled.seif@aast.edu; *Yasmeen.adel@ufe.edu.eg*

Abstract

Internet is no longer a luxury or an optional service. Most of the people nowadays require some sort of internet connection to be able to carry out daily activities and work. Egypt has been investing in upgrading the internet infrastructure with major improvements to the speed and stability of the internet connections. Recently, all the major internet service providers in Egypt introduced plans with faster speeds, more traffic cap, and cheaper prices, but the throttling speeds have been sharply decreased. However, there is some sort of dissatisfaction of the service that is common among different users around. The aim of this research is to assess the performance of the four major internet service providers in Egypt. There is little research on the quality of the internet service providers in Egypt from the perspective of customers, whether residential or commercial customers. In this research, a model is proposed to find the service factors that matter the most to the customers in Egypt. A questionnaire is developed based on the proposed model. The questionnaire is first distributed on a pilot sample of diverse customers to evaluate whether it was easily understood, or not, and to ensure that appropriate answers were received. The questionnaire is then distributed among a large number of commercial and residential customers. The performance of the major Internet Service Providers is assessed and the room for improvement is identified. The paper results include analysis of the factors that contribute the most to the customer satisfaction in Egypt and the priorities of the performance from the customer perspective so that the companies should work on.

1. Introduction

Service industry has become a major contributor to the economy. It differs significantly from manufacturing by being produced at the same time of production. Hence there is little room for fixing quality problems.

Internet is no longer a luxury service and is nowadays a necessity. Distance learning, international business and world-wide economic institutes are examples of businesses that rely on internet. The COVID-19 outbreak worldwide, that had a lot of people to work from home, have put a stress on the world-wide internet services to the extent that some streaming services and video games services had to put some restriction into action to save some bandwidth for the unpredicted demand for internet from people working from home.

Measuring the performance of the internet service providers is a difficult task. There are two ways of analyzing the performance of the internet service providers. The first way is to determine what the gaps are in current performance of the internet service providers from the customer preferences. The other way is to identify the dimensions that matters the most to the customers.

In this research, we are trying to find the service quality dimensions that matters to the customers of internet service providers in Egypt. We propose some dimensions that we think are important to the customers. In addition to the proposed dimensions, we use the SERVQUAL quality dimensions and we evaluate which parameters could be correlated to the customers' satisfaction. We also analyze whether the current internet service providers' priorities are aligned with the customers' preference or not.

The rest of the paper is divided as follows. In section 2, the relevant literature is reviewed. In section 3, the proposed conceptual framework is introduced. The methodology is discussed in section 4 followed by the data analysis in section 5. Finally, the research is concluded in section 6.

2. Literature Review

Despite the importance of service quality, measuring and evaluating the quality of service is a difficult task, and even defining it has been a big discussion in the literature without general agreement on it (Wisniewski, 2001). Service quality is commonly defined as how the customers' needs are met relative to the customers' expectations (Lewis and Mitchell, 1990; Dotchin and Oakland, 1994; Asubonteng et. al, 1996; Wisniewski et. al, 1996). In other words, if the delivered service is less than the customer expectations, then the quality is perceived as not being satisfactory by the customer (Parasuraman et al., 1985; Lewis and Mitchel, 1990)

SERVQUAL is a well-known model that measures the customers' perception of service quality (Parasuraman et al., 1985; Ladhari, 2009). SERVQUAL measures five dimensions: tangible part of the service, reliability defined as the ability to perform the promised service dependably and accurately, assurance defined as the knowledge and courtesy of employees and their ability to convey trust and confidence, empathy defined as the provision of caring and individualized attention to customer, and responsiveness defined as the willingness to help customers and to provide prompt service. SERVQUAL is based on 2 sets of questions: the first set measures what any service provider should provide from the customers' perspective for the five dimensions and the second set measures what the company is actually delivering. The analysis then shows whether there is a significant difference between what is delivered and what is expected. Dinçer et al. (2019) was able to further use SERVQUAL criteria to rank competitors under hesitant decision-making process. Souri et al. (2018) extended the use of SERVQUAL to include the measurement of customers' attitude towards green products.

The service quality dimensions as defined by the SERVQAUL model can be applied to the internet service providers. However, other dimensions contribute to this quality. Cayana et al. (2007) noted that the SERVQUAL dimensions should be assessed to see its impact on different service providers. Boohene and Agyapong (2011) notice that telecom service providers' quality is also concerned by the customer relationship management. Omotayo and Joachim (2011) study of the customers of telecom companies that showed that the service quality has a positive impact on satisfaction and loyalty. This is similar to what Belwal and Amireh (2018) found when applying a similar study on two major telecommunication companies in Oman. Caruana (2002) showed that SERVQUAL dimensions have an effect on service loyalty.

3. Proposed conceptual framework

The independent variables of the study are appearance, empathy, reliability, responsiveness and assurance as defined in SERVQUAL (Parasuraman et al., 1985; Parasuraman et al., 1988; Iwaarden et al., 2003). We also suspect that the performance of the data transmission by the Internet Service Providers (ISP) will also be an independent variable. The dependent variable is the customer satisfaction. The proposed conceptual model is shown in figure 1.

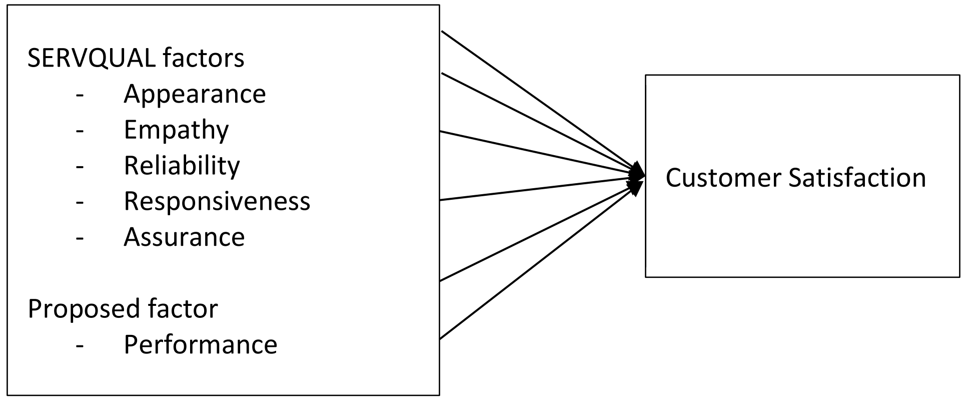


Figure 1. Proposed Model

The model implies a main hypothesis that the SERVQUAL factors along with the performance factors suggested have a significant effect on the customer satisfaction of the ISP companies in Cairo. The main hypothesis is divided into 6 minor hypotheses as follow:

- There is a significant effect of the physical appearance on customer satisfaction of the ISP companies in Cairo.
- There is a significant effect of empathy on customer satisfaction of the ISP companies in Cairo.
- There is a significant effect of reliability on customer satisfaction of the ISP companies in Cairo.
- There is a significant effect of responsiveness on customer satisfaction of the ISP companies in Cairo.
- There is a significant effect of assurance on customer satisfaction of the ISP companies in Cairo.
- There is a significant effect of performance on customer satisfaction of the ISP companies in Cairo.

4. Methodology and Data Analysis

The population of this study is Cairo at large, the main reason behind this choice is the assumption that the capital of Egypt receives most of the funds to improve the internet speed. Hence what applies to Cairo now would be similar to the other cities when they receive similar upgrades.

To ensure the validity of the study, a pilot study was developed by handing the questionnaire to 10 participants with different backgrounds; from being general participant to being scholar in Industrial Engineering. It should be noted that the survey questions are in Arabic to ensure proper understanding of the questions and accurate responses. The feedback received was used to develop a better version of the questionnaire.

The questionnaire collected qualitative data that are divided into 8 sections. Section 1 consists of 6 general questions including ISP, gender, whether the participant pays the ISP invoice by himself or paid by someone else, work status, whether the internet is used at home or at work, and a question asking the participants to order the different performance criteria. Section 2 consists of 4 questions collecting data about the physical appearance. Section 3 consists of 5 questions collecting data about the empathy within the service. Section 4 consists of 5 questions about the reliability of service. Section 5 consists of 4 questions collecting data about the responsiveness of the service provider. Section 6 consists of 4 questions collecting data about the service assurance of the ISP. Section 7 consists of 7 questions collecting data about the performance of the ISP. Section 8 consists of 3 questions collecting data about customer's satisfaction.

The last question in the first section is a question that tries to find the preferences of the Egyptian customers when it comes to the performance of the service. For this question, the participants are asked to order different service performance indicators. The authors propose 8 factors as follows: speed, stability, sustainability, price, value, customer service, technical support, and quota. Speed in this study is defined as the actual average speed the customers receive, not the announced speed. Stability in this study is defined as the deviation from the average speed received. The smaller the variation from the average speed the better. Sustainability of the internet service is defined here as not facing service drops or disconnection. The fewer the disconnection, the better the sustainability. Price is the cost the customer pays for the service. The value is the perception of what the customer receives for the paid cost. Customer service is defined as all the activities provided by the representatives to facilitate the service such as billing, sales and similar activities excluding technical support as defined later. The authors suspect that the technical support, that is defined as activities provided by representative regarding faults or problems facing the customers, should be evaluated as a different factor than the customer service. The last factor is the quota defined as the maximum number of gigabytes allowed for the subscription price. The more gigabytes allowed, the better for the customer (e.g., 120 GB cap for entry service plan).

5. Data Analysis

Cronbach's alpha was used to test the reliability of the data collected. Table 1 provides the Cronbach's alpha for each independent factor, the combined reliability of the independent factors, the reliability of the dependent factor, and the combined reliability of the independent and the dependent questions. The Cronbach's alpha coefficient for each factor is as follows: physical appearance (0.73), empathy (0.77), reliability (0.87), responsiveness (0.76), assurance (0.67) and performance (0.75). The independent variables resulted in alpha of 0.93 and the whole instrument resulted in alpha of 0.94. Alpha values are considered adequate for the study according to Taber (2018).

Table 1. Data validity using CronBach's Alpha

Factor	No of questions	Cronbach's alpha
Physical appearance	4	0.73
Empathy	5	0.77
Reliability	5	0.87
Responsiveness	4	0.76
Assurance	4	0.67
Performance	7	0.75
Factors combined	**29**	**0.93**
Satisfaction	2	0.81
Instrument	**31**	**0.94**

Respondents are distributed among the 4 main ISPs as (68%, 16%, 8% and 8%) which is similar to the percentage of providers' market share. The respondents are divided into 70% males and 30% females, out of the respondents 85% are responsible for paying the bill. For the work status, 75% are employed, 15% are students, 9% unemployed, and 1% retired. Most of the respondents (96%) are using the internet at home and 4% are using it at work.

To analyze the question of what matters the most for the customers, the customers were asked to sort the eight parameters proposed in the performance factor. The null hypothesis of whether there is a significant difference between their mean rank is tested. One-way ANOVA analysis is applied to the results to analyze whether there is a significant difference between the mean rank given to each factor, or not. The descriptive analysis of the mean rank is shown in table 2. The ANOVA results show that there is a significant difference between the different mean ranks at 0.05 significance level. As for the customer preferences, it is observed that speed and sustainability are ranked highest according to the customers' perspective, and there isn't a significant difference between their mean ranks. So, they are considered as the first rank. Stability of service is ranked second. Price is ranked third. Value, Technical support and quota are ranked fourth. And customer service for non-technical services are ranked last without being significantly different from quota.

Table 2. Descriptive statistics of the performance factors ranks

	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean	
				Lower Bound	Upper Bound
Speed	6.8457	1.58447	.11977	6.6093	7.0821
Stability	5.7600	2.01135	.15204	5.4599	6.0601
Sustainability	6.5600	1.03702	.07839	6.4053	6.7147
Price	4.3371	1.70717	.12905	4.0824	4.5918
Value	3.5714	1.64851	.12462	3.3255	3.8174
CS	2.5543	1.24853	.09438	2.3680	2.7406
TS	3.2914	1.73228	.13095	3.0330	3.5499
Quota	3.0800	2.19571	.16598	2.7524	3.4076

To get an idea about the participants' satisfaction with the different dimensions, the degree of acceptance is computed for the different factors as shown in table 3. It is shown that the degree of acceptance for the SERVQUAL dimensions, the performance dimension, the proposed model combining the SERVQUAL and performance dimensions, and the performance (i.e., satisfaction) is 2.69, 3.40, 2.84, and 3.28, respectively. This suggests that the customers may need further improvement in the SERVQUAL dimensions. It also suggests that the proposed performance dimension may turn out to be important for the customers. The high degree of acceptance of the customer satisfaction may be a result of the proposed model factors; and the factors need to be further explored.

Table 3. Degree of acceptance

Factor	Mean	Variance
SERVQUAL	2.69	1.359827
Performance	3.40	1.303739
Combined	2.84	1.378162
Satisfaction	3.28	1.246814

The major hypothesis of the study is "there is a significant effect of the SERVQUAL dimensions along with the performance factors suggested on the customer satisfaction of the ISP companies in Cairo". Pearson Correlation was used to test this hypothesis. The Pearson correlation is 0.803 as shown in table 4 indicating that there is a positive and strong correlation between the proposed factors, i.e., the SERVQUAL factors in addition to the proposed performance factors, and the customer satisfaction.

To test which independent factors affect the satisfaction as a dependent variable, regression analysis is used. The independent factors are the SERVQUAL factors in addition to the proposed performance factors. Table 5 shows the value of R2 as 0.846 which means that the proposed model was able to explain 84.6% of the variation in the customer satisfaction as a dependent factor. It also shows that the Durbin-Watson is 2.054 and the F value is significant at 0.05%.

Table 4. Pearson Correlation

		Sat	SERVQUAL
Sat	Pearson Correlation	1	.803**
	Sig. (1-tailed)		.000
	N	169	144
SERVQUAL	Pearson Correlation	.803**	1
	Sig. (1-tailed)	.000	
	N	144	147

**. Correlation is significant at the 0.01 level (1-tailed).

Based on the above analysis, we fail to reject the null hypothesis and conclude that there is a significant effect of the SERVQUAL and the proposed performance factors on the customer satisfaction of the ISPs in Cairo at 5% significance level.

Table 5. Enter Regression

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.846[a]	.715	.703	.60875	.715	57.336	6	137	.000	2.057

a. Predictors: (Constant), PER, TAN, ASS, RES, EMP, REL

To test the six minor hypothesis, stepwise regression is applied to the data as shown in table 6. R2 values are 0.62, 0.701 and 0.713 for the SERVQUAL reliability, the proposed performance, and the SERVQUAL empathy factors, respectively. Table 7 shows the ANOVA analysis. It could be concluded from the two analysis that we fail to reject the minor hypothesis that the SERVQUAL reliability, the proposed performance, and the SERVQUAL empathy factors have a significant effect on the customer satisfaction of the ISPs service in Cairo at 5% significance level. Also, we reject the minor hypothesis for the remaining factors: physical appearance, responsiveness and assurance.

Table 6. Stepwise Regression

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. F Change	
1	.788[a]	.620	.618	.69043	.620	231.930	1	142	.000	
2	.837[b]	.701	.696	.61525	.080	37.823	1	141	.000	
3	.845[c]	.713	.707	.60430	.013	6.156	1	140	.014	2.054

a. Predictors: (Constant), REL (reliability)

b. Predictors: (Constant), REL (reliability), PER (performance)

c. Predictors: (Constant), REL (reliability), PER (performance), EMP (empathy)

d. Dependent Variable: SAT (customer satisfaction)

To analyze how the ISPs different performance factors are performing from the customers' perspective, the customers are asked to evaluate the different performance factors. Pairwise comparison of the means is applied on the results. The results show that at 5% significance level, stability is evaluated highest by the customers without being significantly different from technical support. Technical support is not significantly different from the remaining factors, so they are all combined as the second rank. It should be noted that the perceived quality of the customer here is totally different from the customer preferences discussed earlier.

Table 7. ANOVA analysis

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	110.560	1	110.560	231.930	.000[b]
	Residual	67.690	142	.477		
	Total	178.250	143			
2	Regression	124.877	2	62.438	164.948	.000[c]
	Residual	53.373	141	.379		
	Total	178.250	143			
3	Regression	127.125	3	42.375	116.039	.000[d]
	Residual	51.125	140	.365		
	Total	178.250	143			

a. Dependent Variable: SAT (satisfaction)

b. Predictors: (Constant), REL (reliability)

c. Predictors: (Constant), REL(reliability), PER (performance)

d. Predictors: (Constant), REL(reliability), PER (performance), EMP (empathy)

6. Conclusion

In this research, the quality of the internet service providers in Cairo as a major city in Egypt is assessed and a proposed model of the parameters that may affect the customer satisfaction is developed. The model is based on the SERVQUAL dimensions in addition to a proposed dimension that assesses the performance of the service. A questionnaire is developed on Microsoft forms and distributed to a wide range of participants. SPSS was used in the analysis of the responses. Cronbach's alpha shows that the responses were indeed valid and fit for the study. The results show that the SERVQUAL reliability, the proposed performance dimensions and the empathy are the main contributors to the customer satisfaction.

The customers were asked to rank the different factors proposed in the performance dimension, then we asked them to evaluate each factor in the performance dimension. The main motive is to find whether the current status of the service is aligned with the customers preferences. The results show that at 5% significance level, stability and technical support are evaluated highest as a company's offer. On the other side, speed and sustainability are ranked highest as customers preferences, followed by stability of service, then price. Hence, we conclude that the speed and sustainability should be improved as opposed to the technical support which is currently the top performance parameter despite having a much lower preference by the customers.

This paper was developed just before the COVID-19 outbreak, and the internet service became essential at this period. This adds to the importance of the internet service all around the world. Further extension to this work includes an extensive study specific to each internet service provider given its own strategic customer target.

7. References

Asubonteng, P., McCleary, K. and Swan, J. (1996). SERVQUAL revisited: a critical review of service quality. Journal of Services Marketing, 10(6), 62-81.

Barbara R. Lewis Vincent W. Mitchell, (1990). Defining and Measuring the Quality of Customer Service. Marketing Intelligence & Planning, 8(6), 11 – 17.

Belwal R. and Amireh M. 2018. Service quality and attitudinal loyalty: Consumers’ perception of two major telecommunication companies in Oman. Arab economic and business journal, 13, 197–208.

Boohene, R., Agyapong, G. K.Q., 2011, Analysis of the Antecedents of Customer Loyalty of Telecommunication Industry in Ghana: The Case of Vodafone (Ghana), International Business Research Vol. 4, No. 1

Caruana, A. (2002). Service loyalty: The effects of service quality and the mediating role of customer satisfaction. European Journal of Marketing, 36(7), 811-828.

Dinçera H., Yüksela S. and Martínezb L. 2019. Analysis of balanced scorecard-based SERVQUAL criteria based on hesitant decision-making approaches. Xomputers & Industrial Engineering, 131(1-12).

Donnelly M., Shiu E., Dalrymple J.F., Wisniewski M. (1996). Adapting the SERVQUAL scale and approach to meet the needs of local authority services. In: Kanji G.K. (eds) Total Quality Management in Action. Springer, Dordrecht

Dotchin, J. and Oakland, J. (1994). Total Quality Management in Services: Part 3: Distinguishing Perceptions of Service Quality. International Journal of Quality & Reliability Management, 11(4), 6-28.

Joachim, A. A.; Omotayo, O. A.; Omotayo, O. J., 2011, Strategic Orientations and Technology Policy: An Empirical Test of Relationship in a Developing Economy. Journal of Contemporary Management Research. 5(1), 1-11.

Ladhari, R. (2009). A review of twenty years of SERVQUAL research. International Journal of Quality and Service Sciences, 1(2), 172-198.

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1985). A Conceptual Model of Service Quality and Its Implications for Future Research. Journal of Marketing, 49(4), 41–50.

Souri M.E., Sajjadian F., Sheikh R. and Sana S. S. 2018. Grey SERVQUAL method to measure consumers' attitudes towardsgreen products - A case study of Iranian consumers of LED bulbs, Journal of Cleaner Production, 177(187-196)

Wisniewski, M. (2001), Using SERVQUAL to assess customer satisfaction with public sector services, Managing Service Quality: An International Journal, 11(6), 380-388.

A CRITICAL ANALYSIS OF THE AIAG-VDA FMEA; DOES THE NEWLY RELEASED AIAG-VDA METHOD OFFER IMPROVEMENTS OVER THE FORMER AIAG METHOD?

Christopher Kluse[1]
[1]Bowling Green State University
ckluse@ bgsu.edu

Abstract

After several years of collaboration, The Automotive Industry Action Group (AIAG) and The German Association of the Automotive Industry (VDA) completed a harmonized version of their respective FMEA methods. Previous to the joint initiative, each organization had their own Failure Modes and Effects Analysis (FMEA) recommended methods & guidelines. The differing methods posed confusion among those automotive suppliers that supplied both markets. As a result, the AIAG & VDA FMEA 1st edition Handbook (2019) was created to mitigate the confusion, improve the overall FMEA development process, and to improve efficiency developing accurate, robust and effective FMEA's. The AIAG-VDA 1st edition (2019) may or may not have improved the method or offered efficiency gains to the FMEA development process. Early critique of the new method offered by Harpster (2018) stated "If the AIAG-VDA DFMEA methodology (which uses structure, function, and failure analyses to determine critical DFMEA content) is implemented, it will result in an automotive DFMEA process that is both considerably less effective and much more inefficient than the current methodology described in the AIAG 4th Edition FMEA manual".

This result of this research is a critical, holistic review of the recently published (June 2019) AIAG-VDA FMEA Handbook along with summarized advantages and disadvantages associated with the new automotive FMEA development process. While the new FMEA AIAG-VDA process offers several advantages, and results in a more robust FMEA, the increased complexity, need for dedicated software, and the learning curve associated with the new process may represent a significant barrier to improving the FMEA development process in the automotive community.

1. Evolution of the AIAG-VDA first edition FMEA

The risk assessment tool, Failure mode and effects analysis (FMEA) has been in existence since 1949 when the military released MIL-P-1629 and then MIL-STD-1629A. Even though this standard was cancelled in 1998, the FMEA process used to assess risk in the design and manufacture of goods and services was adopted by several industries including, but not limited to, automotive, aerospace and healthcare. As a result, several guidance documents that both refined and outlined the original military FMEA process were developed since the release of MIL-STD-16929. Some of the notable, recent guidance documents include SAE J-1739 Potential Failure Mode and Effects Analysis in Design, Potential Failure Mode and Effects Analysis in Manufacturing and Assembly Processes, AS 13004 Process Failure Mode and Effects Analysis (PFMEA) and Control Plans, The Automotive Industry Action Group (AIAG) publication Failure Modes Effect and Analysis Reference Manual 4th edition (2008) and The German Association of the Automotive Industry (VDA), Product and Process FMEA

2nd edition (2012) revised June 2012. The AIAG FMEA handbook was originally published in 1993 and has since been revised on three occasions with the most recent revision, the 4th edition in June 2008. The AIAG FMEA process has been generally accepted as one of the most referenced guides in the United States to developing both design and process FMEA's and many of the notable standards have been derived from the AIAG FMEA handbook. The exception being the VDA methodology, which in brief, used a different approach (often referred to as the tree structure method) to FMEA development. Consequently, in the automotive community, a disparity existed between the two methods causing confusion for organizations that supplied both sectors. As a result, the AIAG and VDA felt it was necessary to join forces, develop a working group to harmonize the FMEA approaches and ease some of the confusion in the supply base. As cited by Kymal and Gruska (2019), "the goal of the workgroup was to develop a single, robust, accurate and improved methodology that satisfied both AIAG and VDA FMEA guidelines". The result is the AIAG-VDA FMEA 1st edition released June 2019. While the intent was to commonize and improve the FMEA process, at this point, and only 9 months since the official release, debate is ongoing as to whether or not this new process is an improvement over the latter AIAG handbook. Richard Harpster (2018), a noted expert and consultant in the FMEA discipline, asserted during the draft review phase of the document that, "There is still considerable resistance to the VDA tree methodology in the US. If it is ever adopted, it will put FMEA's back 20 years". Though some would agree with this statement, others such as Levinson (2020) claim the new approach to FMEA development represents a "major advance" in this new FMEA process. This researcher's primary motive is to determine if this new AIAG-VDA process, offers an overall improvement to the former AIAG FMEA process, which was often executed incorrectly, or resulted in poorly developed FMEA's. In order to assess the effectiveness of the new FMEA process, this research will present the significant changes to the FMEA process as described by this author and other notable authorities in the FMEA field, and the perceived advantages and disadvantages of this new FMEA process. From this, conclusions regarding the overall effectiveness and efficiency of the new FMEA process are offered.

2. Objectives, Methods, and scope of the research

The intent of this research is to present the holistic changes to the FMEA process as outlined in the AIAG-VDA Failure Modes and Effects Analysis Handbook, 1st edition (2019). The papers' focus is on the overall process, policy and procedures that affect the creation of a robust FMEA. A critique of specific technical aspects such as the new severity, occurrence and detection criteria are not in the scope of this research nor are detailed examples of completed FMEA's. The scope of the research is specifically targeted to compare and contrast the AIAG-VDA 1st edition with the AIAG 4th edition (2008) FMEA development process. A comparison to the VDA, June 2012 FMEA process is not within the scope of this paper. Thus the objectives of this research are:

1. Explain and critique the AIAG-VDA method as compared to the former method, the AIAG 4th edition.
2. Differentiate the significant aspects between the AIAG-VDA and AIAG FMEA method.
3. Illustrate the advantages and disadvantages offered by the AIAG-VDA method.
4. Provide the reader with additional sources offering a perspective on the AIAG-FMEA process.

As source for clarification, throughout the paper the terms "former" or "prior" method refers to the AIAG 4th edition (2008) and the "new" method refers to the AIAG-VDA 1st edition (2019) method.

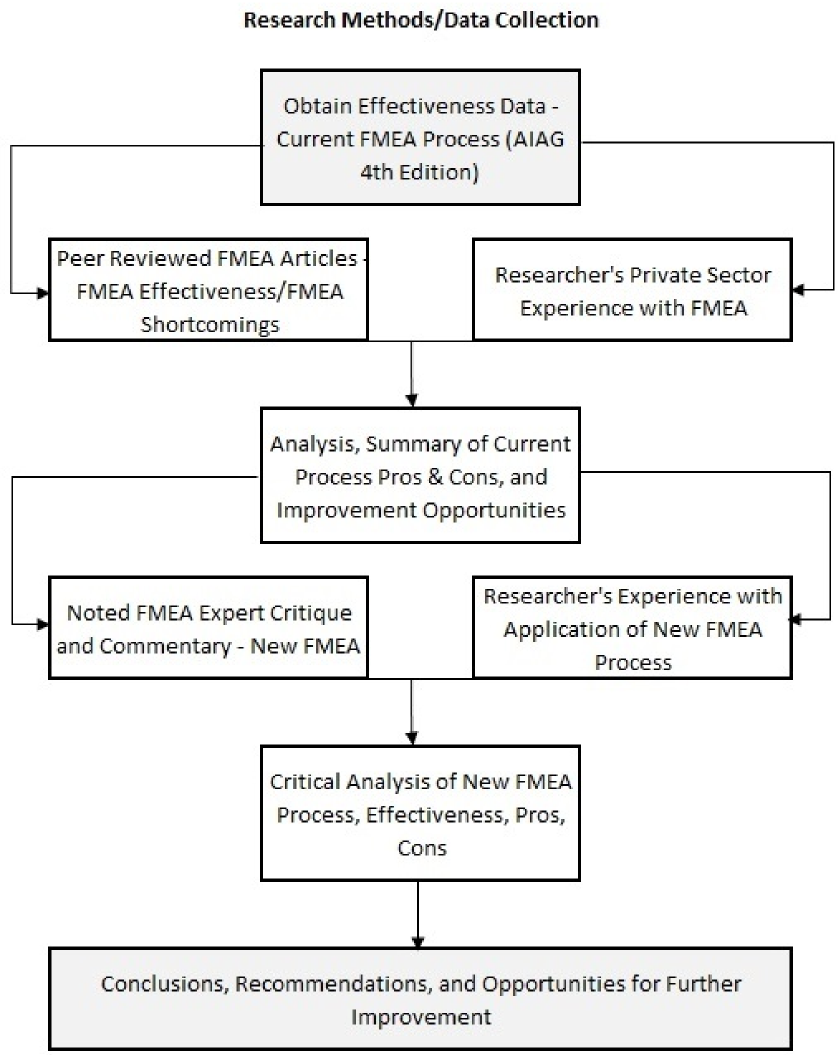


Figure 1. Data Collection & Methods

3. Significant changes to the FMEA process

In general, by harmonizing two methodologies, it is hard to argue that the end result, especially for the automotive supply base, is a unified standard FMEA process that eliminates redundant efforts resulting in improved process efficiency. On that premise, the focus on changes will specifically lie with the methodological and process changes in the AIAG-VDA FMEA as compared to the former process defined by the AIAG methodology (4^{th} edition) that has been in place since 1993.

3.1 Methodological & philosophical changes to FMEA process

The initial chapter in the AIAG-VDA handbook provides substantial guidance on overall FMEA activities segregated into five areas which are; (a) FMEA purpose and description, (b) limitations and objectives of a FMEA, (c) integration of the FMEA into the entire organization, (d) types of FMEA's and specific guidance on when to apply, and (e) FMEA development planning. A noteworthy addition to the 1st edition AIAG-VDA FMEA (2019) purpose and description is the clear distinction that the risk assessed using a FMEA is limited to "analyzing the technical risks to reduce failures and improve safety in products and process" (pg. 15). Other aspects of risk such as financial, time related risk and strategic risks are defined as out of scope for the FMEA process. This indicates that the FMEA alone is not the sole risk assessment tool or strategy used by an organization and this limitation of scope compliments the current ISO 9000 Standard requirement regarding risk assessment. Scope as defined in the former process merely outlined the boundaries of what is appropriately contained in each type of FMEA with no mention of the limitations of risks identified.

The *limitations and objectives* section specifically points out that the FMEA process is a *qualitative* method that relies on the FMEA's team knowledge that may or may not accurately predict performance or predict/assess risk. While this may seem either intuitive (i.e the subjective nature) or insignificant, it is an important statement regarding the FMEA and its inherent subjective nature often the source of criticism in the former FMEA process. Criticism of the former FMEA process focused heavily on the subjective nature of the RPN ranking process and the subjective evaluation of the FMEA when subject to third party and/or customer audits. This assertion that the FMEA is a subjective tool clearly indicates that the FMEA and predicted risk associated with a process and product is only as good as the team knowledge. This indicates the necessity for proper selection of an experienced facilitator and knowledgeable team members to develop an accurate, robust FMEA. However, with the introduction of the action priority (AP) number to replace the former risk priority number (RPN), it is clear that the process is not completely subjective. The emphasis for accuracy is highly dependent on the FMEA planning process.

While the *integration into the organization* section includes important aspects of the FMEA process such as senior management commitment, transition strategy from the former process to the new process, and a discussion outlining foundational and family (generic, baseline FMEA's) FMEA importance and purpose, the important aspect defined in this section is the clarification of realistic failures when identifying potential failure modes. The AIAG-VDA 1st edition (2019) clearly delineates that extreme events such as loss of power to the manufacturing facility are not to be considered in the FMEA and failure resulting from intentional misuse are failures that are not be considered in the FMEA process. Furthermore, explicit language states that a complete FMEA does not conceal any potential failures based on documenting or exposing to much information for those users or customers of the FMEA document. Concerns or issues of proprietary nature are to be worked out between the customer and supply chain.

The final section offering significant methodological change is the project planning section. The planning strategy for FMEA development is described by "five T" method that is short for In**T**ent, **T**iming, **T**eam, **T**ask and **T**ool. While planning was addressed in the AIAG method through use of a multidisciplinary team, it was not as detailed and thorough nor was it considered a specific, documented stage in the FMEA development process. By adhering to the 5T planning method, the team can limit the subjectivity and assure accuracy of the final FMEA.

Most of these changes are related to the planning phase of the FMEA. While scope and planning were addressed in the former AIAG method, the new approach represents an enhanced preplanning phase giving way to several benefits. The pre planning phase now demands a decisive focus and will save time by having a clear scope, clear boundaries as to what is, and what is not, to be included in the FMEA, baseline data in the form of a foundational FMEA, and a strict identification of required roles and responsibilities. However only through a dedicated commitment from management in the form of resources and time will these benefits be realized.

3.2 Specific FMEA Process Changes

In addition to the planning changes outlined above, changes have been incorporated into the FMEA format and all aspects of the FMEA creation process. The figures below represent the former

AIAG version and the revised AIAG-VDA version.

Company	Failure Mode and Effects Analysis									FMEA Number Indentification				Page of			
Part Number (s) or Part Family			Design or Process Responsibility			Prepared by and their Title						Telephone # / Email Address					
Process/Design			Team Members			FMEA Creation Date						Latest FMEA Revision Date					
Process Step/Input or Design Item	Potential Failure Mode	Potential Effect(s) of Failure	SEV	Potential Cause(s) / Mechanism(s) of Failure	OCC	Current Process Controls to Prevent Failure Mode	Current Process Controls to Detect Failure Mode	DET	RPN	Recommended Actions	Person Responsible for Actions	Target Completion Date	Actions Taken	SEV	OCC	DET	RPN

Figure 2. AIAG FMEA – Typical FMEA Header

PLANNING AND PREPARATION (STEP 1)		
Company Name	Subject	
Manufacturing Location	PFMEA Start Date	PFMEA ID Number
Customer Name	PFMEA Revision Date	Process Responsibly
Model Year(s)/Program(s)	Cross Functional Team	Confidentiality Level

Figure 3. AIAG-VDA FMEA Header

	Continuous Improver	STRUCTURE ANALYSIS (STEP 2)			FUNCTION ANALYSIS (STEP 3)			FAILURE ANALYSIS (STEP 4)			
Issue #	History/Change Authorization (As Applicable)	1. Next higher Level	2. Focus Element	3. Next Lower Level or Characteristic Type	1. Function of the Process Item Function of the System, Subsystem, Part Element, or Process	2. Function of the Process Step and Product Characteristic (Quantitative value is optional)	3. Function of the Process Work Element and Process Characteristic	1. Failure Effects (FE)	Severity (s) of FE	2. Failure Mode (FM) of the Process Step	3. Failure Cause (FC) of the Work Element

Figure 4. AIAG-VDA FMEA – Optimization Steps 2 thru 4

RISK ANALYSIS (STEP 5)							OPTIMIZATION (STEP 6)													
Current Prevention Control (PC) of FC	Occurrence (OC) of FC	Current Detection Control (DC) of FC or FM	Detection (D) of FC/FM	PFMEA AP	Special Characteristic	Filter Code (Optional)	Prevention Action	Detection Action	Responsible Person's Name	Target Completion Date	Status	Actions Taken with Pointer to Evidence	Completion Date	Severity (S)	Occurrence (O)	Detection (D)	SpProd Char	PFMEA AP	Remarks

Figure 5. AIAG-VDA FMEA – Optimization Steps 5 thru 6

By brief examination of the forms (figures 3 thru 5), it is evident the process has changed and a level of complexity added. The new AIAG-VDA version implements a 7 step process; these steps as seen in figure 3 are: (1) planning and preparation, (2) structure analysis, (3) function analysis, (4) failure analysis, (5) risk analysis, (6) optimization and (7) results documentation. It is important to note that the former AIAG method did not prescribe a formal tiered (step) approach to FMEA creation. In contrast to the 4th edition FMEA format, the AIAG-VDA format introduces far more columns to populate than the AIAG format leading to pre planning work not typically required in the former AIAG method. As seen in figure 3, the header represents step 1, and structure analysis, function analysis and failure analysis are steps 2 thru 4 respectively. Each of these main groupings are then segregated into three subgroupings that are defined as the focus element, the higher element and the lower element. By AIAG-VDA manual (2020) definition, the focus element is: "The subject of the analysis. In a hierarchically described system, a focus element has a next higher-level element and at least one next lower element. A focus element may be a System Element (item), a function of a system element, or a failure to provide a function as specified." (p. 236). The higher element for a process FMEA is the process item, or the end product being produced. The focus element is the process step or station (element) used to produce the process item, and the lowest element is the process work element. The process work element is the category of causes that could, in turn, affect the focus element. This tiered structure, or development of the structure tree, is completely new to the PFMEA process and adds a level of complexity over the former AIAG method. However, this systematic structure to identify failure will encourage the FMEA team to address both the higher and lower level elements and not bypass or skip to conclusions based off experience or intuition. It forces the thought process. A simple, generic function tree as developed by Thurnes, Zeihsel, Visnepolschi, and Hallfell (2015), reproduced under creative commons license CC BY-NC-ND 4.0 is shown below in figure 5. This type of tree activity is required as a pre-planning activity to both the DFMEA and PFMEA.

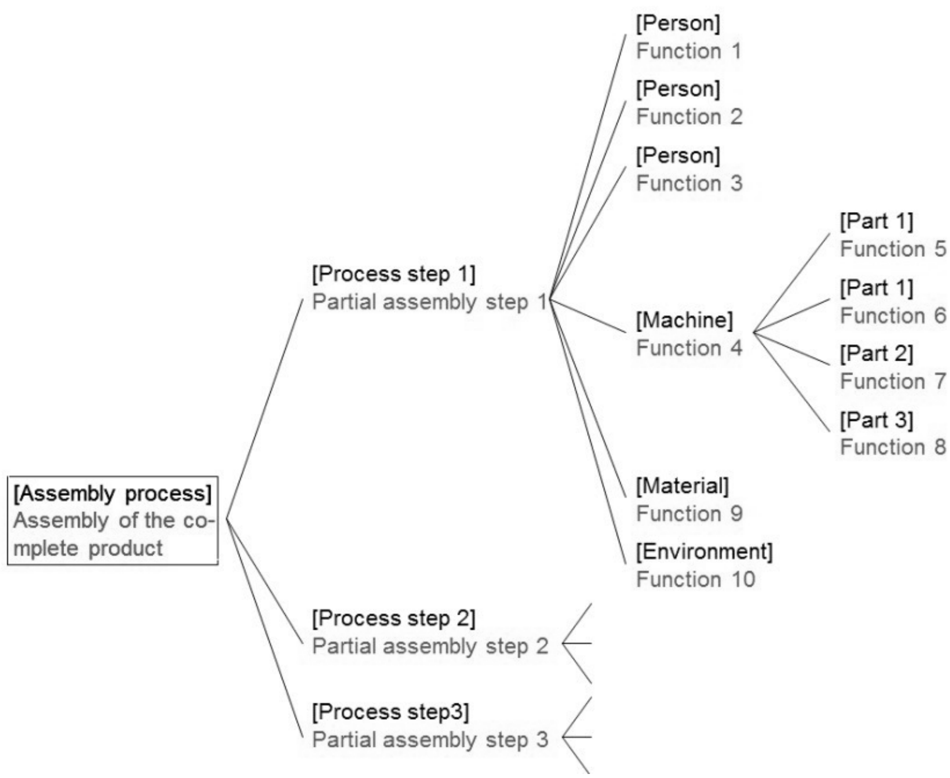


Figure 6. System Structure and Function Tree.

It is not until these elements are established in this fashion can the FMEA team evaluate failure modes for *each* of the elements. In contrast, the former FMEA allowed a rapid progression directly from operation to failure mode with potential to ignore elements now identified by the new AIAG-VDA method. In the former method, the path was essentially linear with potential to jump to failure modes that are inherent or obvious and overlook failure modes related to both the higher and lower elements. At this point in the FMEA creation process, the team is only in step #2 of the 7-step process. Under the former method, the team would be approximately 66% complete with the FMEA. Following the structure tree analysis, the team will proceed to step #3, the function analysis, by listing (a) the function of the process item in the structure tree, (b) the function of the process step(s) in the structure tree, and (c) the function of the process work element(s) in the structure tree. In this third step, the FMEA team is outlining, in detail, process function, operation, and the work elements (man, machine, methods, material, measurement, environment, etc.) by describing what each element is intended to do or contribute to the focus element. Directly from this analysis, the team can evaluate the negative effects of each of these three areas (function, operation, and element) and complete step #4, or failure analysis. While failure analysis was similar in the former method (meaning failure effects, mode and cause are evaluated), and was a significant part of the former FMEA process, it was not derived from a detailed structural and functional analysis, thus the new process may promote the development of a (more) comprehensive, accurate FMEA. Furthermore it is not until the detailed 3-stage functional analysis is complete does the team determine the severity of the failure effect (FE). Previously, the severity value was directly assigned after a mere listing of the potential failure effect, and often lacked consideration for other focus elements now specified in the functional analysis. With the additional consideration of the work elements and associated failure modes, the potential for added robustness to the FMEA exists, however the tedious aspect of the added steps and evaluation criteria may present a burden to those responsible for FMEA development.

At the completion of the failure analysis phase, step 5, the risk analysis phase is conducted. It is in this phase that the assignment of severity, occurrence and detection values occur along with the identification of current controls to prevent and detect failure within the failure chain established in step 4. While this step remains somewhat consistent with the AIAG 4th edition, the timing in which it occurs (after all failures have been identified) and the move from the use of the risk priority number (RPN) to the action priority (AP) numbers represent a significant change from the former process. It should also be noted that the severity, occurrence and detection criteria has changed considerably since the AIAG 4th edition. In the former AIAG process, the RPN value was a product of the severity, occurrence and detection values, and each are weighted equally. This gave the possibility of calculated RPN's being identical for a process or design step that are far more severe than process or design step due to the equal weight factor. The action priority number (AP) has seemingly addressed the issue by weighting severity at the highest priority for initiation of action followed by occurrence and then detection. Actions based from RPN values focused on magnitude of the RPN without proper consideration for severity. The resulting actions driven by the new action priority number (AP) are severity driven. An action priority (AP) table, new to the AIAG-VDA method essentially identifies priority of actions and places them into a high, medium or low priority category. For example, if a failure effect (FE) carries a severity of ranking of a "7" indicating *high* severity to the customer or plant, and the associated occurrence ranking is rated as a "4" indicating a *moderate* chance of occurrence, and a detection of "7" indicating a *low* chance to detect, the resulting action priority category results in an "H" or *high action* priority driven by the severity ranking of 7. In the former AIAG 4th edition method, with such a low occurrence value, the need for corrective action as determined by RPN may have not considered since other RPN's could carry a higher overall value yet the severity may be as high or higher than a 7. While each category of severity, occurrence, detection still carry a subjective 1 – 10 rating scale, and are selected by the FMEA team, the combination of these values as discretely listed in the action priority table clearly provides advantage over the former method. Specific numeric combinations (defined in the table) dictate whether the action is of a high, medium or low level – it is not dictated by RPN magnitude with severity, occurrence, and detection carrying the same weight. At the conclusion of step 5, the severity, occurrence, detection values have been assigned and the action priority category determined. This leaves the FMEA team with a clear direction leading into step 6, optimization.

The optimization step 6, while similar in the intent as the former AIAG "recommended actions" FMEA phase, exhibits key differences that distinguish this phase from the latter. A fundamental difference lies in the focus of the actions; in the AIAG method, actions were not focused on the both the prevention and detection aspect of the failure cause (FC) and/or the failure mode (FM). Using this approach, the team can decide upon those actions that have the ability to reduce occurrence (prevention) of the failure cause (FC) and increase the chance of identifying (detection) a given failure cause (FC) and or failure mode (FM). The optimization phase now has a tiered approach to developing and implementing actions to mitigate risk – this is a strength of this phase. Since severity ultimately dictates the action priority category, the strategy that is built into this new process is to prioritize mitigation of the failure effect (FE) followed by a focus on reducing the occurrence of the failure cause (FC) and then a focus on detecting the failure mode (FM) or the failure cause (FC). The process may appear as intuitive, however the former AIAG method was not prescriptive enough to encourage this strategy thus recommended actions may have resulted in trivial actions or actions lacking consideration for severity of the failure. In addition to an emphasis on prevention and detection, the optimization process now assures that the FMEA team not only identifies the specific actions to be taken, the responsible individual and the target implementation date, (similar to AIAG method), this phase requires a documented status indicating if the suggested actions are (a) not yet defined, (b) action defined but pending approval, (c) pending implementation, (d) completed and (e) not implemented based on practical factors weighed against risk/technical capabilities and/or cost consideration. In the former method, actions are often left as blank (status unclear) or pending which often left the FMEA as static document vs. a living document. The new optimization phase helps improve this most important aspect (and criticism) of the FMEA – specific actions to mitigate risk. Once actions have been implemented, the action priority (AP) is re-evaluated based off those action taken and documented on the FMEA form. What differs and has strengthened the process is and added requirement for "pointers of evidence" which is placed directly on the FMEA form itself. This column serves as validation or proof that the actions have truly been implemented and the place or location where this evidence can be located.

A final phase new to the FMEA process is stage 7, the final report or results documentation that is intended for management. This is not part of the FMEA form, however is an integral part of the new process. Documentation of results is intended to summarize all FMEA related activities and communicate results, actions, concerns or needs to management. While there is no specific format specified, suggested items that should be included are documentation of the 5T's, scope summary, function summary and development process for the functions, an overview of AP values especially those that pose the highest risk requiring action, outline of actions taken (pending or awaiting approval), and the overall plan and milestone timing to implement actions that have resulted from the FMEA development process.

4. Pros and cons of the AIAG-VDA FMEA process

The intent of the team that developed the AIAG-VDA FMEA handbook was to commonize the FMEA process for those suppliers who needed to satisfy both markets that consequently caused confusion and redundant work. Additionally, the team set out "to develop a single, robust, accurate and improved methodology that satisfied both AIAG and VDA FMEA guidelines" (Kymal, Grusaka 2019). A critical look as the shortcomings and advantages of this process as compared to the former, AIAG 4th edition FMEA process can give some insight whether or not the process has advanced robust FMEA development. It is important to note that in the US automotive industry most were more familiar with the AIAG method vs. the VDA method of FMEA development.

4.1 Shortcomings of the new AIAG-VDA FMEA process

The new process is complicated and will require significant time to complete. With the addition of the 7-phase process, the time to complete the FMEA from planning (step 1) to results documentation (step 7) has increased significantly. An estimate of the increased time (using working examples) requirements to complete a FMEA from this researches perspective is approximately 60% additional time. Furthermore, the added complexity has significant potential to reduce FMEA accuracy and completeness. It was noted from previous research "While the PFMEA can be a valuable tool, due to tedious development, vast & intricate guidelines, subjectivity, complicated RPN's the PFMEA method is not properly or utilized developed as suggested by current guidelines" (Kluse 2017, pg. 112). Regarding the new FMEA method, Harpster (2018, paragraph 2, conclusions) makes the important observation that "When the proposed AIAG-VDA Handbook DFMEA method is used to construct DFMEAs which address the Throttle Positioner example found in the 2012 VDA Manual, the following must be created; structural analysis, function analysis, failure analysis, and 4 DFMEA's with a total of 12 lines". A similar case can be made for the PFMEA thus further supporting the added complexity of the AIAG-VDA FMEA process. Establishing that the former FMEA process was viewed as tedious, the new process undoubtedly adds a level of complexity and an increased tedious aspect to the FMEA development process thereby increasing the possibility of an incomplete, inaccurate FMEA that may not lend itself well to risk management. Harper (2018, 2019) further expands on this complexity and shortcoming of the AIAG-VDA FMEA.

Successful completion of a FMEA using the new process is heavily dependent on FMEA specific software. After reviewing the FMEA process and developing the FMEA using a MS Excel created template, it is evident that dedicated FMEA software, while not required to complete the FMEA, is necessary to complete a FMEA in an *efficient* manner. Considering the additional time required to complete the FMEA, and complete the additional worksheets required for functional analysis, element analysis etc., software will become a necessity and Excel templates represent a burden or barrier to completion of an accurate FMEA. Both Harpster (2018) and Gruska (2020) mention the prerequisites needed to complete FMEA's using the new method; while these preliminary development activities can be performed manually and/or with a complicated MS Excel template, software will surely streamline this process. This was not required when developing FMEA with the former method. Harper (2019) further asserts "This column will examine several key issues, including why I believe the AIAG-VDA FMEA Handbook European committee members' allegiance to existing VDA FMEA software and the AIAG decision to become a seller of VDA FMEA-based software has led to the creation of an FMEA handbook that, if implemented, will severely increase implementation time while seriously degrading the effectiveness of FMEAs in the North American automotive industry." (Harper, 2019, paragraph 1). While this may seem insignificant, it is a barrier to robust FMEA development. The Ford, E. C., Smith, K., Terezakis, S., Croog, V., Gollamudi, S., Gage, Keck, J., DeWeese, T., & Sibley, G. (2014) study leads to the conclusion that a FMEA can be effectively completed with only moderate use of resources, it was also demonstrated that a FMEA process can be streamlined to produce an effective, value added FMEA without some of the tedious requirements (FMEA forms, long meetings, etc.) of the traditional method.

Management commitment in the FMEA process is now far more critical. Management commitment has always, and will continue, to be a critical aspect of proper FMEA development. The augmented preplanning and preparation process outlined in the AIAG-VDA FMEA now demands an increased level of resource commitment and involvement in the FMEA process. In addition to defining the FMEA scope during the planning phase, the new process suggests developing family and/or and baseline FMEA's, identifying a FMEA project plan by documenting the 5T's, discretely defining team members and roles, and creating structural and functional tree structures upfront requires a significant investment in resources, mainly in FMEA development time and FMEA dedicated resources. It is known from research that management support, while critical, is a reason for poor FMEA development. "PFMEA development is poor due to lack of resources, minimal emphasis or placed on the PFMEA and PFMEA development process, lack of management support and minimal benefits realized from the PFMEA" (Kluse 2017, pg. 112). As a result, management support must increase in order to assure the AIAG-VDA method concludes in an accurate, robust FMEA.

The AIAG-VDA FMEA may not manage risk as well as the AIAG 4th edition. Working through a PFMEA for automotive component manufacture (fascia systems), this author found that by following the AIAG-VDA process, several failures within a process step lead to failure causes in the FMEA, while they were an important and valuable aspect identified, they did not necessarily represent root cause thereby resulting in the FMEA team not considering prevention or detection actions of the root cause for failure. In cases such as this that seem to exist in many examples, risk mitigation is not as robust as in the former AIAG 4th edition method. Supported by a detailed analysis presented by Harpster (2018) who states "Without knowing the root causes of why the process step failed, a prevention control cannot be defined nor can the PFMEA line be used for risk management", therefore the case for a less effective method for risk management can be valid.

4.2 Advantages/improvements offered by the new AIAG-VDA FMEA process

The new action priority (AP) value leads to true action priority vs. the former RPN metric. It was well documented that the RPN method used in the AIAIG method was flawed in several ways. The most notable flaw was that each constituent that went into the RPN calculation, severity, occurrence and detection, are all selected based off a subjective scale ranging from 1 to 10. The product of these 3 numbers became the RPN value. The problem was that several process within the FMEA may end up with an identical RPN value yet the severity of failure to the next operation and/or end user differed in the impact felt – in other words there was no weight value on severity, occurrence and detection. As noted in the discussion of FMEA changes, (section 2.2 of this paper) the new action priority number places a greater weight on the severity factor thereby ranking actions with high severity whereas with the former RPN system to rank risk, a process with high severity may not have been the focus of actions due to another process with low detection and moderate occurrence. According to Narayanagounder & Gurusami (2009) since RPN is a product of the severity, occurrence and detection, in any given FMEA, one could have identical RPN values yet the true risk potential may differ. Thus, this new action priority (AP) ranking suggests improvement over the former RPN method since it lacked a true, discrete risk rank system which is one of the main reasons that FMEA's are developed.

Changes to the severity, occurrence and detection ranking criteria. Closely related to the action priority number are the changes made to the severity, occurrence and detection ranking criteria. While each still carries the somewhat subjective 1 – 10 scale, the severity table has added columns for severity impact of failure on the production plant/operation, the ship to plant or immediate external customer, and the end user. For a tier 1 automotive component supplier, this would mean that severity impact would be considered for the producing plant and subsequent operations, the OEM that would use the component in vehicle assembly, and the end user who ultimately purchases the vehicle from a dealer. Severity criteria no longer are determined by warning whereas before, severity increased if the failure occurred and the user/affected party realized the failure without warning. High severity rank is also tied directly to safe operation and lack of compliance with regulations. Occurrence ranking criteria that were formerly quantitative based and were conducive, in particular, to high volume operations are now primarily based on the type or classification of control these being "behavioral" control or "technical" control. The latter are those controls that are human dependent driven by policy and/or procedural control. Technical control relates to more sophisticated methods such as error proof techniques or automated techniques that reduce chance of failure occurrence. The detection criterion not only considers the likelihood of detecting a failure but also now enhances the detection rank by considering the maturity of the detection system. Thus, regardless of perceived effectiveness of a detection method, consideration must be given to evidence based, historical performance, of the detection method. For example, if a barcode scan operation is used to determine that the correct component is installed during an assembly operation, in the previous ranking criteria this was often considered a moderate to high assurance of detection. However, if the FMEA team knows from history that the barcode scan process is easily bypassed or has failed to detect installation of wrong components, the resulting detection value, based on proven maturity, would be far lower and therefore require actions to improve and mitigate weak detection presumed to be strong. This greatly improves FMEA accuracy since assumptions based on established systems (bar code scan, color coding, manual inspection, RFD tag, vision system, etc.) often carry the same detection value regardless of performance or capability. These detection criteria would also lend well to the implementation of new technology (AI or VR) that offers perceived benefits over traditional detection methods.

Consideration for structural, functional and element analysis should lead to a robust overall FMEA development process. While it has been noted that the preliminary work developing tree structures and the requirement for conducting structural, functional and element analysis adds complexity, the benefit, when conducted properly, may increase robustness of the FMEA. By considering failure possibilities, or failure cause of the element (i.e. 4M's, man, machine, method, material, measurement, environment) in both the structural and functional analysis, this alone may add an element of robustness to the FMEA that previously may have not been considered or inadvertently overlooked.

Optimization and formal results documentation. The final phases added to the FMEA process adds an element of thoroughness, a check and balance for management and a record of FMEA development activities. A clear difference and apparent advantage is the shift from merely identifying "recommended actions" as in the AIAG method, but to now define both preventive and detection actions to mitigate effect of the failure, prevent or reduce how often the failure occurs, and improve the chance of failure detection. The former recommended actions in the AIAG method often did not include all three of these aspects and even though and RPN was reduced, the overall risk to creating a flawed design or producing a defective product may have gone unchanged. Optimization in the AIAG-VDA FMEA, similar to the former AIAG method requires documentation of responsible party for the actions along with planned date of completion. What has been enhanced are the requirements for providing evidence that the actions took place (if AP number will be lowered) and the status of actions which includes (a) open/pending, (b) decision regarding actions pending, (c) action implementation pending, (d) actions completed, and (e) not implemented/completed due to risk vs. reward scenario. Status a, d, and e are considered mandatory while b and c are optional. Regardless of the optional categories, introduction of the action status requirement is a significant gain in the FMEA process and lends itself to prompting the "living" intent of the FMEA. Results documentation, an added final phase in the FMEA process, has the intent to inform stakeholders of the final conclusions at FMEA completion, establish the FMEA as a living document, provide a clear record and status of actions resulting from the FMEA process, documentation and communication of risk mitigation activities, and provide a record of the quantifiable change in risk to the level deemed acceptable by management and/or the customer. A completed FMEA was always the intent of the former process, however the results documentation phase adds documentation requirements to assure all objectives of the FMEA have been captured and may be used as an input into the lessons learned process.

Supplemental FMEA for monitoring and systems response (MSR-FMEA). The FMEA-MSR is a new category of FMEA introduced with the AIAG-VDA process. As noted in the AIAG-VDA 1st edition handbook (2019), the intent of this FMEA is for "monitoring and system response, potential failure causes which might occur under customer operating conditions are analyzed with respect to their technical effects on the system, vehicle people and regulatory compliance. The method considers whether or not failure causes or failure modes are detected by the system, or failure effects are detects by the driver. Customer operation is to be understood as end-user operation or in-service and maintenance operations" (pg. 125). At time of this writing, this type of FMEA has not been mandated to be completed by suppliers of automotive OEM's (DFME's and PFMEA's are mandatory), however this represents a significant aspect for consideration of consumer safety when operating a motor vehicle. In retrospect, a MSR FMEA being completed may lead to a reduction in vehicle recalls if consumer operation and interaction are fully explored. The failure mode leading to the GM ignition switch recall, responsible for 124 deaths and 266 injury claims, (Automotive News, 2015) may have been evident if such a FMEA strategy was utilized.

Table 1. Summary of Major Advantages/Disadvantages

Advantages/Improvements	Disadvantages/Shortcomings
AP Value (replaces RPN) - Promotes/ranks priority using (S) significance as base	Complicated, tedious to complete accurately & correctly
Severity, Occurrence, Detection (S O D) criteria - improved objectivity and targeted focus	Necessity to utilize dedicated FMEA specific software
Required structural, functional, & element analysis improve FMEA robustness	Management commitment - requires increased Mngt support in planning and action to AP list
Documentation of optimization & formal FMEA results	Risk Management - failure causes (FC) not necessarily root cause
MSR-FMEA - Addresses operating condition risk(s)	Substantial training required to acclimate to new process

5. Final thoughts and conclusions regarding AIAG-VDA FMEA process

As mentioned in the scope and objectives of the paper, the intent for this research was to present a critical look at the holistic aspect of the new AIAG-VDA FMEA process to evaluate if the process has advanced the knowledge and robustness of the FMEA process. In speaking with colleagues in the automotive community, the jury is still out on whether or not this process will improve effective FMEA development. Reviewing a recent blog site, Elsmar Cove (2020), that has a dedicated board to comment on the new AIAG-FMEA process, comments are predominantly negative and the theme surrounds the complexity of the new process – it is tedious and time consuming. This researcher has placed these comments for review in the appendix of this paper. Previous research into the FMEA process by this researcher (Kluse 2017, Kluse 2018) mainly concluded (along with other factors) that the tedious nature of the FMEA process is a barrier to accurate and robust FMEA development. While this researcher fully supports the new AIAG-VDA process as it offers far more technical and procedural aspects that surely improve the FMEA process, it is hard to ignore the increased complexity and tedious aspect that has been placed upon FMEA development teams and may again represent the most significant barrier to a robust, value added FMEA. The reader is encouraged to visit Harpster (2020) for a case study and thorough technical review of the AIAG-VDA process as well as Levinson (2020) and Gruska (2020) for alternate perspectives regarding the AIAG-VDA FMEA process.

6. Appendix –Elsmar Cove Forum for Quality Professionals

The comments below were retrieved on March 13, 2020 from https://elsmar.com/elsmarqualityforum/threads/new-aiag-fmea-process-how-to-complete-the-new-format.76552/. These comments below are selected since they represent a critique of the new FMEA process, others have not been included that are outside the scope of the paper. The reader, however, may visit the site to review all posted comments.

Comment #1

Ok so the new "process" has been published and I have (tried and tried again) to read it and understand it. So far 5+ reads later and I still do not have a clue how to complete the new format. It seems to be way over the top compared to the previous, if slightly flawed editions which at least after a quick read most people could get cracking with the generation of a PFMEA.

It now appears likely that you will need your own "expert" just to understand how to complete / co-ordinate completion of the form which to my mind completely defeats the object of getting process owners to "do it for themselves".

I think this is for the benefit of "quality consultants" who will be rubbing their hands with glee and deciding how they are going to spend all the extra money they will now be getting on the back of all the new training that will be required. (Probably an "acceptable" cost for larger organizations, but for those smaller organizations with relatively simple processes it seems a completely unnecessary expense).

All in all a major backward step that will achieve nothing but confusion for many people involved in simpler processes, and there should have been the option of a more streamlined FMEA process available for processes that involve less complex parts / processes.

Comment #2

I always thought that the main reason for FMEA's was to identify process variables, how they can go wrong, and what you can do to mitigate them. There are far far better, simpler and more user friendly ways to achieve this objective than what is written in the latest specification. I fear the new process will detract from the main reasons for carrying out an FMEA process as people will be focusing more on how to fill out the form!

Comment #3

I started doing FMEA's 30 + years ago, (pre computer and hand written), the reasons for doing them are sound, but even now there seems to be much confusion as to what constitutes a "failure mode", or a cause, (of a cause etc.), resulting in more time spent discussing how to complete the formant and what goes in what box and completely missing the point of doing an FMEA in the first place.

Comment #4

The new approach say it is too complicated to do in Excel and that you will need special software, and those same people work for the companies that plan to provide that software.

Comment #5

I'm not sure I can even get through the 1st page. The presentation you posted shows the confusion and complication of it. And I have already one Customer asking to submit PFMEAs in the *new format.*

Comment #6

I am composing a presentation on process FMEA (Design is a little more challenging and I have not worked in that field), and the new manual actually makes sense. The concepts are pretty much the same but it stresses that you start with the failure mode as the focus element, and go from there to the failure effect (consequence) and the failure cause (why the failure mode happens, formerly known as the mechanism). One improvement is that the Occurrence rating is based on the process step's prevention controls as opposed to an estimate of the probability of failure, which is often hard to obtain unless you have an established process with a known process performance index. And Ppk doesn't account for assignable causes that result in non-conformances.

They also use an action priority matrix which is better than the risk priority number (the product of three ordinal numbers). And they also look at three classes of failures: (1) poor quality, the traditional one, (2) failures that affect continuity of operations and (3) failures with occupational health and safety issues. I am actually starting to like it.

One drawback (my opinion) is that the work elements considered are normally only four of the six traditional categories of a cause and effect diagram, with Method and Measurement omitted. I found an example in which the Method (drill speed and feed rate) might in fact be instrumental in preventing the failure in question, namely broken drill bits (which result in undrilled parts which can be detected before they leave the workstation). I am working off a Shigeo Shingo case study in error-proofing.

Another is that it is possible to get a Low action priority for a 9-10 severity failure and occurrence and detection ratings greater than 1 (1=the failure cause cannot cause the failure mode, period, for the Occurrence rating), which should not be possible, but that is just my opinion.

In any event, I really like the manual and it is definitely an improvement on how things were done 20 years ago.

Comment #7

Quick feedback on this new revision coming from applying in an automotive electronics project.

Seems the structure breakdown in essence is a great approach however I found rather difficult to apply to electronics.

My sneaking suspicion at the moment, unless I am utterly wrong, that it is more likely for mechanical structures where the connection of components kind of straight forward due to the obvious sequential nature by physics, for electronics however representing the physical structure blows up a horrendous 3D spider net to represent in those 3 columns.

Finally I have dropped the structural approach and follow the journey of electricity on the schematics to nail this structure.

Also, word of caution is to use fmea in conjunction with FTA or other approach that provides more functional insights as a complement to dfmea. Ratings for severity/occurrence/etc seem academic as usual for designs, we can use that, adds little to none to design analysis

7. References

Automotive Industry Action Group, (2008). Potential Failure Mode and Effects Analysis, Reference Manual, 4th edition. Chrysler LLC, Ford Motor Company, General Motors Corporation.

Automotive Industry Action Group, (2019). Failure Mode and Effects Analysis, Handbook, 1st edition. AIAIG, Southfield, MI 2019.

Ford, E. C., Smith, K., Terezakis, S., Croog, V., Gollamudi, S., Gage, Keck, J., DeWeese, T., & Sibley, G. (2014). A streamlined failure mode and effects analysis. Medical Physics, 41(6), 061709-n/a. doi:10.1118/1.4875687.

Gruska, G. (2020). The AIAG-VDA DFMEA, Understanding the Changes. Omnex Navigator, Winter 2020, p. 1-6. Retrieved on March 11th 2020 from https://www.omnex.com/members/newsletters/omnex-2020-navigator.aspx.

Harpster, R. (2018). The Case Against the AIAG-VDA DFMEA. Quality Digest, January 31st, 2018, 14:42. Retrieved on January 14th 2020 from https://www.qualitydigest.com/inside/risk-management-column/case-against-aiag-vda-dfmea-013118.html.

Harpster, R. (2019). Opinion: Should the AIAG Be in the FMEA Software Business? Quality Digest, January 10th, 2019, 13:03. Retrieved on March 11th 2020 from https://www.qualitydigest.com/inside/operations-column/opinion-should-aiag-be-fmea-software-business-011019.html

Harpster, R. (2020). http://www.harpcosystems.com, 2020.

Kluse, Christopher (2017). Failure Modes and Effects Analysis (FMEA): Factors Affecting Execution and Implementation of the FMEA and an Alternate Method for Process Risk Assessment. The Journal of Management and Engineering Integration, 10(1) Summer 2017, 106 – 116, Association for Industry, Engineering, and Management Systems.

Kluse, Christopher J. (2018) The 25th Anniversary of the AIAG FMEA Reference Manual: A Systematic Literature Review of Alternative FMEA Methods. The Journal of Management and Engineering Integration 11(2), Winter 2018: 37 – 45.

Kymal, C., Gruska, G. (2019). Introducing the AIAG-VDA DFMEA. Quality Digest, June 24th, 2019, 12:03. Retrieved on January 14th 2020 from https://www.qualitydigest.com/inside/management-article/inntroducing-aiag-vda-dfmea-062419.html.

Levinson, W.A., (2020). AIAG/VDA's FMEA Manual Is a Major Advance. Quality Digest, January 15th, 2020, 13:03. Retrieved on January 14th 2020 from https://www.qualitydigest.com/inside/lean-article/aiagvda-s-fmea-manual-major-advance-011520.html.

Military Specification MIL-STD-1629 Rev A, Notice 3 (1998). Procedures for Performing a Failure Mode Effects and Criticality Analysis, Department of Defense.

Naughton, N. (2015). GM Ignition Switch Death Toll Reaches 124. Automotive News, July 13th edition, Crain Communications.

Narayanagounder, S., Gurusami, K. (2009). A New Approach for Prioritization of Failure Modes in Design FMEA using ANOVA. World Academy of Science, Engineering and Technology, International Science Index 25, International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering, 3(1), 73 - 80.

Thurnes, Christian & Zeihsel, Frank & Visnepolschi, Svetlana & Hallfell, Frank. (2015). Using TRIZ to Invent Failures – Concept and Application to go Beyond Traditional FMEA. Procedia Engineering. 131. 10.1016/j.proeng.2015.12.439. Creative Commons license CC BY-NC-ND 4.0. http://creativecommons.org/licenses/by-nc-nd/4.0/

A PROBABILISTIC DYNAMIC PROGRAMING MODEL FOR DETERMINING OPTIMUM INVENTORIES

John Wang[1]
Steve Bin Zhou[2]
Zhi Pei[3]
[1]Montclair State University
[2]University of Houston-Downtown
[3]The University of Texas at Tyler
wangj@montclair.edu; drbinzhou@yahoo.com; zpei@uttyler.edu

Abstract

The occurrence and size of snowfall, and therefore, the level of demand for salt, are uncertain for each week. Therefore, a probabilistic inventory model is required. This will be a periodic review model - the inventory of salt at the beginning of each week will determine the optimal quantity to be purchased to satisfy expected demand. Dynamic programming with a backward recursion is utilized to determine weekly optimal costs/inventory level.

1. Introduction

The United States is significantly dependent on its road system to support the rapid, reliable, reasonable movement of people, goods, and services. In many states, this requires substantial planning, training, manpower, equipment, and material resources to clear roads and streets through the winter (EPA, 1999). The most common deicing method is the use of sodium chloride (salt) in the form of crushed rock salt, which is inexpensive, abundant, and easy to mine, store, distribute, and apply. Salt brines are increasingly used in some areas, but the vast majority is still rock salt.

In areas inclined to winter precipitation, transportation network must be able to rapidly respond to snow and ice on roadways. Ice removal is a vital service in these communities. Deicing chemicals melt ice by lowering the temperature at which it melts. They can also prevent new ice from forming and improve traction. Salt is a popular deicing chemical because it is cheap and abundant. As (Lilek, 2017) indicated that the United States produces 22.9 and needs 24.5 million metric tons of rock salt for roadway deicing. Top consuming states are New York, Ohio, and Illinois, the total cost of rock salt used for roadway deicing was $1.18 billion.

Winter road maintenance is a critical function of public works and transportation agencies in states with harsh winter climate or in cold regions. The use of salts represents a cost-effective snowfighting technique, having beneficial impact on public safety, essential mobility and travel costs (Autelitano, Rinaldi, and Giuliani, 2019).

In line with Tiwari and Rachlin (2018), in areas of North America that receive snow, road-salt runoff leads to many problem, such as groundwater salinization, loss or reduction in lake turnover, and changes in soil structure. It also poses threats to birds, mammals, and roadside vegetation. Furthermore, road-salt runoff can affect biotic communities by causing changes in the composition of fish or aquatic invertebrate assemblages.

Having faced an increase in extreme winter weather, many places in the U.S. have experienced increased demand for roadway snow and ice control (RSIC) operations. According to Sullivan, Dowds, Novak, Scott, and Ragsdale (2019), as the number and severity of extreme weather events increases, the costs associated with winter roadway maintenance materials, plow operator time, equipment maintenance and replacement, and fuel use will also increase.

The process involved in salting the roads of a county in Northern New Jersey included the involvement of thirteen inspectors. During the winter periods these inspectors are all assigned to certain parts of the county, and inspect the conditions of the roads during a snowstorm. Once they see the conditions of the roads and determine if they need to be salted, they send trucks out. These trucks are contracted out for their services. The county uses around 25 to 30 contractors paying them ninety dollars an hour for their services. These trucks are filled with 10 tons of mineral rock salt and 24 gallons of calcium chloride which is sprayed onto the salt and acts as a catalyst for melting the snow. The county also uses a mixture of pea gravel with salt when the conditions of the roads are not that bad or when they are low on salt. The percentage of pea gravel to salt is 3:2. In salting the roads, the hills are the first priority followed by the main streets and level areas.

When the snow conditions are less than 3 inches the county salts the roads. If there is more than three inches, they usually do one run before the snow starts to accumulate and then plow the roads, and if necessary the roads are resalted again when the snow stops depending on the intensity of the storm. This process helps set up a base on the roadway preventing a dangerous situation called rib ice and provides better traction for the motorists. After they salt the roads, the salt will last for about an hour and fifteen minutes before they have to resalt. Again, this situation is based on the intensity of the storm and it may or may not have to be done. Some of the costs involved in this process includes the purchasing of the salt, the construction of the dome and all costs associated with it, and other related costs involved with the salting process.

2. The Model

When it comes to winter weather specifics, the only sure thing is that nobody is sure. Any scientist will tell you it is risky to predict temperature more than thirty days in advance and impossible to predict precipitation. Snow, especially the heavy one, has caused too much trouble and frustration.

Many scholars focus on the specific methods and/or materials for deicing. For instance, Hossain, Fu, & Lake (2015) conducted approximately 300 tests in a real-world environment, which cover four alternative materials, and 21 snow events. They presented the results of an extensive field study on the comparative performance of alternative materials for snow and ice control of transportation facilities. Each of the alternatives tested were compared to regular rock salt in terms of snow melting performance — bare-pavement regain time.

In order to design a road de-icing device by heating, Bernardin and Munch (2019) consider in the one dimensional setting the optimal control of a parabolic equation with a nonlinear boundary condition of the Stefan– Boltzmann type. This control problem models the heating of a road during a winter period to keep the road surface temperature above a given threshold. Their model allows to quantify the minimal energy to be provided to keep the road surface without frost or snow.

However, there is a shortage of research articles dealing with the specific dilemma that is faced by the county level officials. The problem is to determine the optimal inventory of salt (required to melt snow) that the Essex County should carry on a weekly basis during the winter months (November to April). The optimal inventory will be determined based on the inventory on hand at the beginning of the week, and the optimal amount to be purchased during the week can be determined based on the expected demand for salt for the week. The expected demand will be determined using historical probability distribution of snowfalls in the county for each week.

The occurrence and size of snowfall, and therefore, the level of demand for salt, are uncertain for each week. Therefore, a probabilistic inventory model is required. This will be a periodic review model - the inventory of salt at the beginning of each week will determine the optimal quantity to be purchased to satisfy expected demand. Dynamic programming with a backward recursion is utilized to determine weekly optimal costs/inventory level.

Dynamic programming is both a mathematical optimization method and a computer programming method (Beuchat, Georghiou, & Lygeros, 2020; Sauré, Begen, & Patrick, 2020). Operational management of reservoirs at hourly/daily timescales is challenging due to the uncertainty associated with the inflow forecasts and the volumes in the reservoir. Ramaswamy and Saleh (2020) apply a multi-objective dynamic programming model and obtain optimized release strategies accounting for the inflow uncertainties. Their study provides perspectives on the benefits of the proposed forecasting and optimization framework in reducing the decision-making burden on the operator by providing the uncertainties associated with the inflows, releases and the water levels in the reservoir.

2.1. Basis and assumptions

1. Salt can be ordered at the beginning of each week and delivered immediately.
2. Inventory management is over a 24-week period, from the second week in November to the third week in April.
3. Bulk purchase of salt is allowed at the beginning of week one or any other week and there is limited storage capacity.
4. There is a holding cost of $h/ton/week and a shortage cost of $s/ton/week. Holding costs include cost of funds tied up in financing inventory, cost of storage, insurance and maintenance costs. Shortage costs include additional costs incurred when salt (or substitutes such as sand, gravel, etc.) has to be obtained in an emergency, the political costs and damage to the reputation of the county management, accident costs, etc., when there is not enough salt to melt snow in the event of a snowstorm. (Some of these costs are difficult to quantify.)
5. Salt may be purchased at a unit cost of $u per ton; there are no fixed costs.
6. Inventory may be left over at the end of the 24-week period and can be stored and used to begin the next season.
7. There are no financial constraints regarding the quantity of salt that can be purchased in any week.

2.2. The mathematical model

There are twelve (discrete) possible levels of snowfall. The probability that the level of snowfall will be j during any particular week is P_j (j = 1, … n). The corresponding expected salt demand for that level of snow is D_j so that for any week t, the expected demand for salt will be

$$E(Dt) = \sum_{j=1}^{n} PjDj$$

i.e., for any week:

Level of Snowfall, j:	1	2	3	…	n
Probability, P_j:	P_1	P_2	P_3	…	P_n
Salt Demand, D_j:	D_1	D_2	D_3	…	D_n

Let

$u = unit cost per ton of salt

$h = holding cost per ton of salt per week

$s = shortage cost per ton of salt per week
c(x) = u * x = purchase cost of x tons of salt
i = inventory (in tons) of salt on hand at the beginning of week t

Then, the minimum cost of meeting demands for weeks t, t + 1, …, 24, given by $f_t(i)$, is

(i) For week twenty-four

$$f_{24}(i) = \min[c(x) + \sum_{i+x \geq Dj} Pj(i + x - Dj) * \square + \sum_{i+x<Dj} Pj(Dj - i - x) * s\,]$$

(ii) For week t ≤ 23

$$f_t(i) = \min[c(x) + \sum_{i+x \geq Dj} Pj(i + x - Dj) * \square + \sum_{i+x<Dj} Pj(Dj - i - x) * s + \sum_{i+x \geq Dj} Pj * f(t+1)(i + x - Dj) * \square + \sum_{i+x<Dj} Pj * f(t+1)(0) * s\,]$$

$$(i, x, D_j \geq 0)$$

The future cost component for $i + x < D_j$ is f_{t+1} (0) because, in the event of a shortage, beginning inventory in the following period will be zero.

2.3. Data

Snowfall distribution

Weekly snowfall figures were compiled from daily records of snowfall obtained from the National Weather Service. Data was obtained from Winter 1993/94 through Winter 2014/2015 (21 years), each period beginning first week in October, i.e.. Week 1, is October 1 through 7 of each year, Week 2 is October 8 through 14, etc. The end of the 24th week is within the third week in April.The snowfall levels (in inches) were simply summed up for the seven days in each week and a probability distribution was derived from the frequency distribution of snowfall levels in each week over the 21-year period. Three snow size/level ranges (points of the discrete probability density function) were obtained which suited the data reasonably. These ranged from zero (or trace of snow) to 30 inches or more.

Salt demand

Exact figures of salt required to melt each inch of snow could not be obtained from the county. According to a county official, this was difficult to estimate because salt use may continue one or two days after a heavy snowstorm (for maintenance) and salt is also used for fire-fighting. However, from charts of annual snow costs provided by the county, we were able to estimate that an average of 1920 tons of salt were used per inch of snow in Winter 2014/15. The county official confirmed that this figure (approximately 2000 tons) may be a reasonable estimate. Salt demand of 2000 tons/inch was therefore considered reasonable for use in the calculations.

Salt costs

A purchase cost of $31.25 per ton of salt was given by the county official.

Holding costs

Storage costs are relatively low because salt is easily stored in a simple warehouse. The county has two major warehouses, each with a storage capacity of 10, 000 tons. The approximate cost of each building was $150, 000. If we assume that the buildings can be used for 25 years, and ignoring time value of money considerations, the cost of each building (amortized over 25 years) is

$150, 000/25= $6, 000/year

Assuming insurance and maintenance costs of $5000 a year, we obtain

Cost of storage per ton per week = ($6, 000 + $5, 000)/(10,000 x 52) = $0.021

If the opportunity cost of funds used to finance inventory is assumed to be 13% p.a., then

Financing cost per ton per week = (31.25 x .13)/52 = $0.078

Addition of the above two costs results in a holding cost of approximately 10 cents per ton per week, which is the cost used in the computations.

Shortage costs

The county has no figures on these costs that could be made available, because of the difficulty in quantifying certain components of these costs. For this reason, shortage cost is initially assumed to be about twice the unit cost of salt ($60/ton). This translates into a total shortage cost of $120, 000 per inch of snow not cleared by the county. It can be argued that political costs alone will be higher than this amount, especially with heavy snowfall. (The shortage cost - snow size relationship is, in reality, nonlinear, but a linear relationship is assumed here.) Therefore, a range of higher (and also lower) shortage costs are used to sensitize the analysis.

3. Computations

A computer program was written to undertake the recursive computations. The following costs have been used in the computations:

Unit cost of salt - $31.25/ton
Holding cost - 10 cents/ton/week
Storage cost - $60/ton/week

The first set of data is the printout of the data file containing the probability distributions for each snow level for the 24-week period.

The computations are shown for total inventory each week (beginning inventory (BEG. INV.) plus purchases (QUANT) for the week not exceeding 25, 000 tons.

An initial run of the program had shown that even for a shortage cost as high as $l, 000/ton, the maximum expected weekly salt requirement would be 22, 000 tons during periods with the highest probabilities of heavy snowfall. The restriction also saves computer time and eliminates redundant data for inventory levels over 25, 000 tons.

It is assumed that minimum purchases of salt will be 1, 000 tons at a time and larger purchases will be in multiples of 1, 000 tons, so increments in purchases/Inventory levels are by 1, 000 tons.

4. Results and discussion

Results show that for Weeks 18, 17, and 16, it is optimal not to make any purchases for any level of beginning inventory, including zero. The Expected total Cost (TOT. COST) which is equal to Purchase Cost (PCOST) plus Expected Holding/Shortage Cost (H/S COST) plus Expected Future Cost (FUT. COST) is minimum for purchase quantity of zero, and is shown in the last column as Optimal Cost (OPT. COST). The optimum cost for each inventory level is highlighted. For subsequent weeks, the optimal level of inventory increases from 2000 tons/week in week 15 to 20,000 tons/week during week 8 and then down to 2000 tons/week in week 1.

The results here show that it is optimal to begin the season by purchasing 2000 tons of salt in the first week if no inventory was left over from last season (BEG. INV. = 0) or to purchase 1000 tons if the inventory left over was 1000 tons. If 2000 or more tons were left over, then no purchases need to be made. For the following weeks, salt can be purchased as required to maintain the optimal inventory level required for each week.

It is observed that, for this shortage cost of $60/ton, it is optimal not to keep any inventory of salt at all (in the case where beginning inventory = 0 and purchases = 0) even though there is a small probability (0.06) that there might be an average snowfall of up to 1 inch during weeks 16, 17, and 18. Because of the inability to quantify shortage costs especially political costs, it is reasonable to assume that the county cannot afford any salt shortage whatsoever and must therefore keep a minimum inventory at all times sufficient to melt at least one inch of snow, i.e., 2000 tons. Two alternative approaches were used to deal with this:

1. A modified program was run to give optimal total weekly inventory values for different shortage costs. The lowest shortage cost which gives optimal total weekly inventory of at least 2000 tons for weeks 16, 17, and 18 was found to be $300/ton. This translates into total shortage cost of $600, 000 per inch of snow not melted. The original program was then run with shortage cost = $300 and holding cost remaining at 10c.
2. A restriction was introduced into the program to ensure that the minimum total inventory for each week is 2000 tons. (For ease of computation, increments in inventory of 2000 tons at a time were used in the recursions here.)

The results show only the optimal values for each amount of beginning inventory, for shortage costs of $100, $200, and $300 per ton. Table 1 shows a comparison of the minimum optimal level of total inventory required each week for various shortage costs ranging from $30 to $1, 000. Shortage cost of $1, 000 is included to illustrate the effect of imposition of a much larger shortage cost, to imply very little allowance for shortage. The inventory levels in Table 1 are the results for the case without restriction of minimum inventory to 2000 tons.

The results show that for all shortage costs, an optimum inventory level of 2000 tons should be maintained in week 1 as in the case with S = $60. From then on, as would be expected, higher optimal levels are required as the probabilities of occurrence and size of snowfall increase in the following weeks. The optimal inventory levels are not much different during peak snowfall periods for the different shortage costs ranging from $100 to $1, 000. They are somewhat lower for a shortage cost of $60. For periods of sparse snowfall (the beginning and ending weeks) the differences across shortage costs are more obvious especially between S = $60 and S =$1, 000.

The B column for S = $60 shows the inventory levels with the restriction of minimum inventory to $2, 000 tons. The difference is that for those weeks in the unrestricted case where inventory equals zero, the inventory levels are now 2000 tons.

The maximum level of inventory recommended by the model is 22,000 tons during the peak periods of snowfall. Optimal inventory levels determined by the model generally correspond to the probability density functions. This is easily observable in week nine, even though it is a period of relatively high snowfall, the optimal inventory is 8000 tons, compared with 22,000 is because the probability distribution cuts off at snow level 5 (3-4 inches of snow) where the salt demand would be 8000 tons. On the other hand, optimal inventory for weeks 8, 10, 11, and 12 amount to 22,000 tons because their probability distributions are spread out up to snow level 12 where the salt demand would be 22,000 tons. The model determines this maximum because of the relatively high shortage costs. It is seen from Table 1 that for lower shortage costs ($60 or less), the optimal inventories are less than 22,000 tons (down to between 6000 and 20,000 tons at S = $40). Howler, the optimal inventories for weeks 6 and 7 remain at 20,000 tons across all shortage costs from $40 to $1000 because they have the most disperse probability distributions as well as being two of the weeks with highest probabilities of snowfall.

Table 1. Optimal inventory levels (in thousand tons) for different shortage cost ($S/ton)

Week	$30	$40	$50	$60 (A)	$60(B)	$100	$200	$300	**$1000**
18	0	0	0	0	2	0	0	2	4
17	0	0	0	0	2	2	2	2	2
16	0	0	0	0	2	2	2	4	10
15	0	0	0	2	2	8	10	10	14
14	0	2	2	4	4	14	16	16	18
13	0	4	4	6	6	14	14	14	14
12	0	6	8	8	10	18	20	22	22
11	0	12	12	12	12	22	22	22	22
10	0	14	14	16	16	22	22	22	22
9	0	8	8	8	8	8	8	8	8
8	0	20	22	22	22	22	22	22	22
7	0	20	20	20	20	20	20	20	20
6	0	20	20	20	20	20	20	20	20
5	0	8	8	8	8	S	8	8	8
4	0	6	6	6	6	6	6	6	6
3	0	10	10	10	10	10	10	10	10
2	0	14	14	14	14	14	14	14	14
1	0	2	2	2	2	2	2	2	2

The results obtained for the other weeks can be similarly explained, e.g., inventory is 14, 000 tons for week 2 and 10, 000 tons for week 3 because there are (small) outlying probabilities of snowfall at snow levels 8 and 6, respectively. The "outliers" seem to bias the model's results, but this is due to the dominant effect of shortage costs which represents the avoidance of shortage in practice, however small the probability and/or size of snowfall.

For a shortage cost of a $30/ton ($60, 000 per inch of snow), the model recommends that no salt be kept. This is because this amount is lower than the purchase plus holding costs of salt per ton, and therefore justifies the assumed shortage costs of $60 or more as realistic.

In view of the likelihood that holding costs may be higher or lower than the 10c per ton assumed, the program was modified to undertake computations for holding costs varying from 5c to 20c, for shortage costs of $60 as well as $300. Results are summarized in Table 2.

Not surprisingly, for a given holding cost, the inventory levels determined by the model are the same for both S = $60 and S = $300, since the shortage costs component, being relatively much higher, dominates holding cost. There would be a significant difference if holding and shortage costs were much closer, which is not the case with the realistic costs in this study.

For the same reason, for a given shortage cost, the variation in holding costs from 5c to 20c does not give significantly different results. For s=$60, there is a difference at week 14 where optimum inventory is 4000 tons for h = 5c, 10c, and 15c, and 2000 tons for h = 20c. For both s = $60 and s = $300, the optimum inventory at week 6 is 22, 000 tons when h = 5c/ton, 20, 000 tons when h = 10c and 18, 000 tons at h = 15c and h = 20c. This shows the effect of the relative dominance of shortage cost (in higher inventory levels when holding costs are lower).

From the above, it can be deduced that any error in estimating holding costs will not significantly affect the results; the major source of problems is the estimation of shortage costs.

Table 2. Optimal inventory levels (in 10^3 tons) for different shortage & holding costs

Shortage	$60	$60	$60	$60	$300	$300	$300	$300
Holding	5c	10c	15c	20c	5c	10c	15c	20c
WK 18	0	0	0	0	2	2	2	2
17	0	0	0	0	2	2	2	2
16	0	0	0	0	4	4	4	4
15	2	2	2	2	10	10	10	10
14	4	4	4	2	16	16	16	16
13	6	6	6	6	14	14	14	14
12	8	8	8	8	22	22	22	22
11	12	12	12	12	22	22	22	22
10	16	16	16	16	22	22	22	22
9	8	8	8	8	8	8	8	8
8	22	22	22	22	22	22	22	22
7	20	20	20	20	20	20	20	20
6	22	20	18	18	22	20	18	18
5	8	8	8	8	8	8	8	8
4	6	6	6	6	6	6	6	6
3	10	10	10	10	10	10	10	10
2	14	14	14	14	14	14	14	14
1	2	2	2	2	2	2	2	2

5. Conclusions

The model's results are reasonable, to the extent that holding and shortage costs can be reasonably quantified. A major limitation is that because of the dominance of the relatively high shortage costs, the model determines a high inventory level even when there is only a very small probability of heavy snowfall, e.g., in weeks 10, 11, and 12, the probability is only .06 that snowfall of over 10 inches will occur (one winter out of 21 years in the sample). Since it is unrealistic to just drop out such isolated probabilities, the model can only be improved by using a larger sample, say with data for 50 to 100 years, which may provide lower probabilities for such isolated cases, e.g., if this occurs only once in 50 or 100 years. The model's results may then be more realistic, but would still depend on the relative dominance of shortage costs.

In conclusion, the model is applicable in practice; even if certain costs are difficult to quantify, it is possible to undertake sensitivity analysis under various cost estimates, as has been done in this study. Overall, an optimal solution was reached, but further study needs to be done to juxtapose this optimal solution with real values obtained over some time period in the future.

6. References

Autelitano, F., Rinaldi, M., Giuliani, F. (2019). Winter highway maintenance strategies: Are all the sodium

chloride salts the same? *Construction & Building Materials.* Nov. 2019, 226, 945-952. DOI: 10.1016/j.conbuildmat.2019.07.292

Bernardin, F. and Munch, A. (2019). Modeling and optimizing a road de-icing device by a nonlinear heating.

ESAIM: Mathematical Modelling & Numerical Analysis, 53(3), 775-803. DOI: 10.1051/m2an/2018056

Beuchat, P. N., Georghiou, A., Lygeros, J. (2020). Performance Guarantees for Model-Based Approximate Dynamic Programming in Continuous Spaces. Transactions on Automatic Control, 65(1), 143-158.

Chi, A. and Lohse, D. (1996, Jan. 10). *Hoist It, Heap It and Haul It: Managing Snow Removal.* Wall Street Journal, p B1.

Hossain, S.M.K., Fu, L.P., Lake, R. (2015). Field evaluation of the performance of alternative deicers for winter maintenance of transportation facilities. Canadian Journal of Civil Engineering. Jul2015, 42(7), 437-448.

Lilek, J. (2017). Roadway deicing in the United States: How a few industrial minerals supply a vital transportation service. AGI Critical Issues Program: www.americangeosciences.org/critical-issues

Ou, J. and Wein, L.M. (1995). Dynamic scheduling of a production/inventory system with products and random yield.

Management Science, 41,1000-17.

Peet, J. (1995, No. 10). Weather or not: predicting winter's impact is cloudy science. *The Star Ledger.*

Puterman, M.L. (1995). Markov decision processes: Discrete stochastic dynamic programming. *Interfaces,* 25, 93-4.

Ramaswamy, V. & Saleh, F. (2020). Ensemble Based Forecasting and Optimization Framework to Optimize

Releases from Water Supply Reservoirs for Flood Control. Water Resources Management, 34(3), 989-1004. DOI:10.1007/s11269-019-02481-8

Underwood, E. (1991). Every time it snows, it's pennies from Heaven. AdWeek's Marketing Week, 32, 9.

White, D.J. (1995). Finite Horizon Markov Decision Processed with Uncertain Terminal Payoffs. *Operations Research,* 43, 862-9.

Sauré, A.; Begen, M., Patrick, J. (2020). Dynamic multi-priority, multi-class patient scheduling with stochastic service times. European Journal of Operational Research, 280(1), p254-265.

Sullivan, J. L., Dowds, J., Novak, D. C., Scott, D. M., Ragsdale, C. (2019). Development and application of an iterative heuristic for roadway snow and ice control. *Transportation Research Part A: Policy & Practice,* 127, 18-31. DOI: 10.1016/j.tra.2019.06.021

Tiwari, A. and Rachlin, J.W. (2018). A review of road salt ecological impacts. Northeastern Naturalist, 25(1), 123-142. DOI: 10.1656/045.025.0110

APPLICATION OF MARKOV MODELS FOR DECISION MAKING UNDER UNCERTAINTY IN THE ELECTRIC UTILITY INDUSTRY

Tettey Anyama[1]
Hensley Kim, PE[1]
Gholston Sampson, PhD[2]
[1]Huntsville Utilities
[2]University of Alabama
aht0005@uah.edu; *kim.Hensley@hsvutil.org*; *gholts@uah.edu*

Abstract

Planning in the Power Systems distribution involves a formal decision-making process of identifying and prioritizing network improvement projects such as the construction of substations, interconnecting feeder links and general upgrade works, with the aim is of providing an efficient service to the customer without damage to the utility's equipment and customer's property or personnel. This task is plagued with uncertainties associated with load growth and demand due to changing weather conditions and other unforeseen developments, which affects the ability to maximize efficiency and utilization. This paper proposes the use of Markov Models as a more effective technique by utility planners and managers for their decision-making efforts under such uncertainties. The authors develop a load flow modeling approach that takes into consideration the stochastic nature of customer demand and uses the distribution network profiles as fitness values to be optimized. A ranking of the criteria of interest based on the decision makers preferences is the result of the optimization algorithm. This provides a formal process for decision making by the management of utility companies.

1. Introduction

Planning in power systems involves a decision-making process of identifying and prioritizing network improvement projects with the aim of providing an efficient and reliable service to the customer without damage to the utility's equipment and customer's property or personnel.

The decision is usually established after extensive analysis of the network conditions using network models, historical transformer and feeder peak loadings as well as expert knowledge considerations. The work in this paper reviews existing approaches used by power system planners to achieve this process and proposes the use of Markov Models to make the process more efficient.

A schematic of the subsystems for a Power System in the United States (US) is shown in Figure 1. The system can be broadly divided into generation, transmission network, distribution network for sub-transmission, and the Low Voltage distribution substations that connect directly to the customer. The complexity associated with transmitting quality power reliably through to each of the subsystems increases with a downward descent in Figure 1. At the generation level the work involves generating and delivering bulk power to meet both the base and peak load demands of bulk and utility customers. Usually, a good energy forecast helps with planning purposes at this level.

At the transmission and distribution levels however, the expanse distribution network, load diversity due to differences in individual demand patterns, industrialization and urbanization, time of day and day of the week behavioral patterns introduce a new dimension of variability. Weather related randomness is another factor that has nothing to do with human behavior. These variabilities make it challenging to transmit power efficiently and reliably through the several installed distribution feeders to customers.

It is worth noting that there is always the need for huge capital investments in infrastructure for a utility company and the distribution planner or decision maker is always faced with the dilemma of how to get the best value for his or her money.

In this research work, a review of existing literature on efficiency and utilization in the electrical distribution network is done. Data from a Utility Company is then used to show evidence of stochasticity in the load behavioral patterns of distribution feeders. Markov Models and optimization techniques are used to illustrate an effective and formal Decision-Making process for distribution network planners in locations where the weather significantly introduces randomness in load behavior.

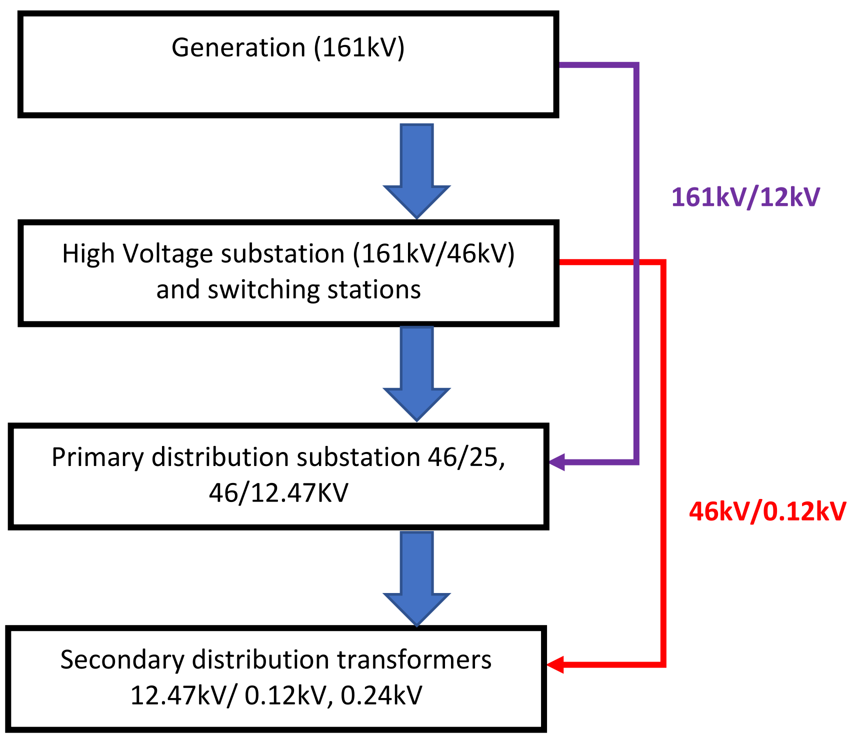


Figure 1. Schematic layout of an electrical power system in the US

2. Literature Review

Management of organizations are always faced with the challenge of putting in place processes and the necessary infrastructure to improve on the quality of the services they provide to customers. Decisions about customer requirements however, almost always come with trade off considerations (Da Silva et al., 1990; Park & Kim, 1998). Most of the time there are negative correlations among requirements of interest to customers.

For example, literature has shown that reliability is one of the most important attributes of service quality as perceived by customers (Babakus & Inhofe, 2015; Gupta, 2018; Singh et al., 2017).

Increasing reliability however will almost always result in cost increase (Benezech & Coulombel, 2013; Burns & Gross, 1990), and this contrasts what customers want: reduced prices. Such observed trends and the importance of satisfying the customer have led to several research work and techniques such as multi-objective optimization to reduce waste and improve efficiency in all systems. In the power system industry, the distribution subsystem particularly stands to benefit from such techniques aimed at improving efficiency and utilization.

(Kramer, 2016) stated that " The US Federal Energy Regulatory Commission (FERC) estimates that tens of billions of dollars per year could be sliced from consumers' utility bills if the transmission system could deliver electricity more efficiently." The question is, 'how can electricity delivery be made more efficient'? In their work in 2009, an executive committee of industry stakeholders identified reduction in system transmission and equipment losses, as well as an increase in line and system utilization as the approaches to help improve efficiency in the industry (Forsten, 2010).

From the basic equations of Ohm and Kirchhoff's laws, we can safely attribute the bulk of system technical losses in a Power System to the distribution network. Apart from the need for an extensive distribution network extending everywhere human settlements are, voltage levels of the distribution network, as well as the variabilities in load demand associated with individual behavioral patterns and weather related uncertainties (Fan et al., 2007) make it a great challenge to put in place a distribution system with high utilization and reduced losses.

There are extensive software packages and techniques used for network modeling and the establishment of network profiles like voltages, system losses, percentage loading of feeders etc. (Coutto Filho et al., 1991; Da Silva et al., 1990; Jasmon & Lee, 1991). For those in the power system industry, the process is popularly known as load flow studies. The load flow results coupled with knowledge of feeder specific load growth estimations is then used to determine the necessary improvements and expansion works needed to maintain the desired network profiles by the utility.

Traditionally, load flow studies use the peak loads of feeders as a representation of the worst-case situation that can ever occur (Da Silva et al., 1990). As will be shown in the next section, seasonal variations experienced at locations like Huntsville, Al and other parts of the US play a significant role in the load behavior (Burns & Gross, 1990), it is therefore important to assign an expectation to the occurrence of feeder specific peak loading for each distribution feeder.

If the peak loading of feeder 'A' reaches 90% of the installed capacity every four years, and feeder 'B' reaches 90% every year, preference should be given to feeder 'B' when it comes to prioritizing improvement projects. However, this is not the usual practice, the annual peak is applied for load flow studies irrespective of the frequency of occurrence of the peak. Silva et al applied probabilistic techniques to the generation and transmission subsystems of the power system (Da Silva et al., 1990) but no such work has been done for the distribution subsystem where the bulk of the losses occur.

Knowing the expectation of a loading condition occurring and planning with that in mind to optimize resources will address both aims of reducing losses and increasing utilization in the distribution network.

3. Evidence of stochasticity in load behavioral patterns

The seasonal variations here in most parts of the United States renders very interesting electricity demand patterns in the utility's network. Depending on the types of load and operations in an area, one can observe feeders winter peaking, summer peaking or both. In figure 2, we show the 15min interval loadings for PERP 214, a summer peaking feeder in the distribution network of Huntsville Utilities. The line 'y=730' shows that loadings in the range of 730A and above should not be expected during the winter temperatures for this feeder. The situation will be totally different for a winter peaking feeder, and this requires an extensive assessment of load behavior on a feeder by feeder bases.

It is worth noting that such stochastic trends in loading due to weather may not be significant or may not even exist in locations where the seasons do not experience such extreme temperature variations. Locations such as Ghana in West Africa do not experience such at all.

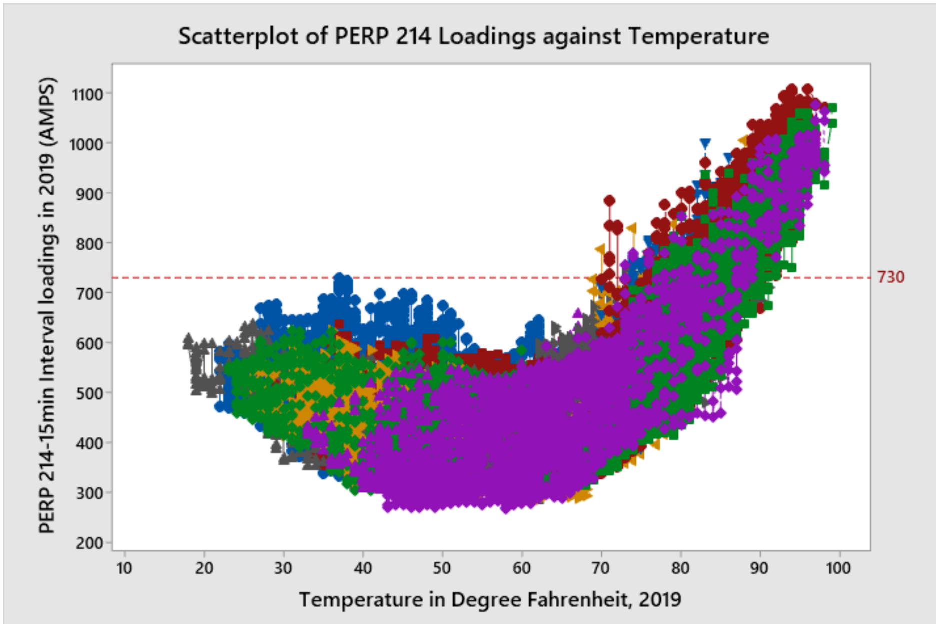


Figure 2. Plot of feeder loading against temperature in 2019

3.1 Regression and correlation analysis

A p-value <0.0001, an R-square and R-square adjusted of 0.8346 and 0.8345 respectively show a good regression model fit for PERP 214 feeder loads against temperature in the summer. The Minitab results reported in Table 1. below shows a beta coefficient of 24.06 for the model. This means the distribution planner should plan for about 24Amps increase with every unit rise in temperature.

Table 1. Regression model output for PERP 214

Model summary			Coefficients					
R-sq	R-sq(adj)	R-sq(pred)	Term	Coef	SE Coef	T-Value	P-Value	VIF
			Constant	1261	16	-80	0	
83%	83%	83%	TEMP AUG	24	0	122	0	1

A detailed assessment of the correlation of PERP feeders and temperature is shown in Table 2. As will be expected from the regression analysis, there is an interesting observation of a very significant correlation of PERP 214 with hotter temperatures but no significance with the winter temperatures. PERP 224 on the other hand shows fairly consistent and significant correlations with temperature throughout the year. Thus, PERP 214 is summer peaking whilst 224 is peaking both in the summer and winter, showing negative correlation for lower temperatures and positive correlation for higher temperatures.

Table 2. Correlations of PERP feeders with temperature calculated for every month in 2019

Month	Min temp	Max temp	CORR (TEMP, PERP 214 AMPS)	CORR (TEMP, PERP 224 AMPS)	CORR (TEMP, PERP 234 AMPS)
January	22	67	-0.127	-0.723	-0.615
February	28	75	0.144	-0.483	-0.269
March	23	76	0.085	-0.599	-0.351
April	33	84	0.694	0.471	0.574
May	49	93	0.875	0.876	0.884
June	56	94	0.882	0.852	0.878
July	62	96	0.905	0.861	0.891
August	60	98	0.914	0.86	0.903
September	62	99	0.935	0.907	0.936
October	35	98	0.814	0.766	0.786
November	18	69	-0.153	-0.675	-0.447
December	26	71	-0.309	-0.657	-0.522

From Table 2, it is not surprising to see a flatter loading curve for PERP 234 compared to PERP 214 as shown in Figure 3 below. Again, this has to do with their correlations with temperature and it brings us to an interesting part of the discussion.

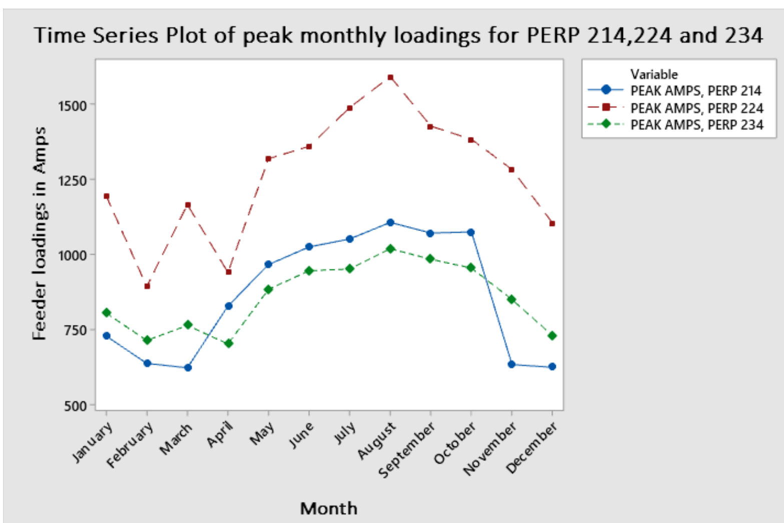


Figure 3. Plot of monthly peak loadings for PERP feeders

Even though PERP 214 has a higher peak load than 234, 234 will have a much higher load factor, which is an indication of higher utilization. This shows that working with the peak values alone (as usually done during load flow studies) does not really guarantee improved utilization.

Finally, summer temperatures in Huntsville have been very consistent over the past 6 years, but the situation has been very different for the winter temperatures. This is depicted in Table 3. 2018 was the last time we experienced a single digit temperature value in Huntsville and so for feeders with high negative correlations with temperature, the distribution planner needs to be prepared to experience a significant increase in load if such an extreme winter temperature occurs again.

Table 3. Extreme winter and summer temperatures in Huntsville since 2015

Extreme summer temperatures in °F						Extreme Winter temperatures in °F						
Year	2015	2016	2017	2018	2019	Year	2015	2016	2017	2018	2019	2020
Max	99	98	96	95	98	Min	8	17	12	7	22	22

The discussion above shows that establishing an improvement plan for the distribution network should not just be about the yearly peak recorded. The effect of the weather and frequency of occurrence of extreme conditions become critical factors to consider when expansion and improvement works are being considered.

3.2 Markov Models: A effective approach to address stochasticity

Markov models are named after a Russian mathematician, Andrey A. Markov who extensively studied Markov Processes at the beginning of the twentieth century. His works led to the branding of some stochastic processes as having the Markovian property. Processes that evolve over time in a probabilistic manner are called stochastic processes. For a stochastic process X_t to have the Markovian property,

$P\{X_{t+1} = j|X_0 = k_0, X_1 = K_1, \ldots, X_{t-1} = k_{t-1}, X_t = i\} = P\{X_{t+1} = j|X_t = i\}$, for $t = 0, 1, \ldots$ and every sequence $i, j, k_0, k_1, \ldots, k_{t-1}$.

In other words, the Markovian property says that the conditional probability of any future event, given any past events and the present state $X_t = i$, is independent of the past events and depends only upon the present state. This is usually referred to as the lack of memory property.

We have established that extreme weather temperatures drive the peak loading patterns for the feeders in Huntsville Utilities' network. We can therefore establish the probability of a loading condition occurring based on the probability of the occurrence of a certain temperature, this idea is used to establish transition probabilities for the Markov models in this research.

If we statistically establish August and January as the months with extreme temperatures, the probability of the occurrence of a certain temperature for a day in August, for example P(T>90) is only dependent on the current state being 'a day in August', and has nothing to do with what happened two months ago.

As an example, we can discretize the daily temperature recorded at a 15 minute time interval in August into 2 or 3 bins; 0='=<90', 1='>90', or 0='<=89', 1='90-94', and 2='>95' respectively for a 2x2 and 3x3 transition probability matrices respectively. The transition probability matrix will show the probability for an expected state of the system starting at a discrete point in time. The choice of the time boundary depends on the researcher and the load characteristics of interest. 12:00 o'clock noon has been used for this study since it serves as a boundary between the first and last halves of the day.

In August, we observed the behavior of the state of the system at the 12:00 noon time daily and used 8 years of temperature data to come up with the transition probabilities. For the transition probabilities shown in Table 4, the first row shows the probability of having a temperature below 90°F before 12 noon and staying below 90°F for the rest of the day in column 1 or transitioning above 90°F in column 2. In row 2, we start off the day with temperatures above 90°F before noon and transition to below 90°F in column 1, or above 90°F in column 2.

The choice of 90°F boundary has to do with temperatures often seen in Huntsville during the summer. This becomes a guide to the distribution planner in that if low voltages and high losses are seen below 90 °F, that should be a major issue to be addressed because that will lead to a lot of inefficiency and customer complaints. The preliminary results for a for a 2X2 transition matrix is shown next.

4. Preliminary Result and discussion

Three summer peaking feeders in Huntsville Utility's network have been used as the case study. The first step was to come up with the 2X2 transition probabilities matrix. The probability of Loadings =< 90°F and > 90°F was used. The loadings at the specified states were used as inputs in the load flow studies to come up with the fitness measures for those states. Records of the line losses in kW for each state of the feeder (at temperatures below and above 90 °F) are shown in Table 4 below.

Table 4. Transition probability matrix and load flow outlook for 3 selected feeders

Transition Probability matrix				Technical losses based on loadings of different states					
				PERP 224		MAD 214		GAR 224	
	Temp =< 90°F	**Temp > 90°F**		**Load @ temp =< 90°F**	**Load @ temp >90°F**	**Load @ temp =< 90°F**	**Load @ temp >90°F**	**Load @ temp =< 90°F**	**Loss @ temp >90°F**
Temp =< 90°F	**0.44**	**0.56**	**Load @ temp =< 90°F**	**1097**	**1340**	**1407**	**1612**	**980**	**1062**
Temp > 90°F	**0.14**	**0.86**	**Load @ temp >90°F**	**1296**	**1360**	**1310**	**1636**	**1052**	**1070**
Loss figures from load flow results in kW				**343**	**1110**	**75**	**233**	**167**	**457**

The method of paired comparison was used to rank the feeders in order of maximum loss as PERP 224>90 °F, GARTH 224>90 °F, PERP 224≤90 °F, a tie of Mad 214>90 °F and Gar 224≤90 °F, and Mad 214≤90 °F. Multi-objective optimization will be used when more than 2 objectives are being considered.

5. Conclusions and recommendations

The different correlation patterns between demand and temperature, even for feeders that are not geographically far apart, call for the need to address the issue of stochasticity in load flow studies.

The peak feeder loading ever recorded is usually used as the worst-case scenario in load flow studies but we have shown that to improve efficiency and utilization, an expectation of the occurrence of extreme loading condition coupled with the expected load growth should be modeled. Markov models are great tools that can give the distribution network planner the network profiles to expect depending on the season. These expectations are used to establish a ranking of the Decision Makers preferences.

Given the fitness measures from the load flow analysis, various optimization techniques can be used to rank alternatives based on the preference of the decision maker. The authors used the method of paired comparisons to rank losses (a single objective) for 3 summer peaking feeders. Ranking was done for 6 alternatives because a 2x2 transition matrix was established for the Markov

Models. This gave an added advantage of comparing feeder profiles at temperatures both below and above 90ºF. The choice of 90ºF temperature boundary has to do with temperatures most often observed in all summer months.

The approach has the added advantage of coming up with appropriate and cost-effective improvement plans. Knowing that a peak occurs very infrequently, may require a basic improvement measure such as the installation of a capacitor bank that switches on load. In fact, this measure has been adopted for the PERP 224 feeder, whilst plans for major improvement works in the area are being considered.

Future studies will use multi-objective optimization to address multiple preferences of the decision- maker. Such objectives may include Goodness of Fit measures for a feeders' relationship with temperature throughout a summer or winter season.

6. References

Babakus, E., & Inhofe, M. (2015). Measuring Perceived Service Quality as a Multi-Attribute Attitude. In M. Levy & D. Grewal (Eds.), Proceedings of the 1993 Academy of Marketing Science (AMS) Annual Conference (pp. 376–380). Springer International Publishing

Benezech, V., & Coulombel, N. (2013). The value of service reliability. Transportation Research Part B, 58(C), 1–15. https://doi.org/10.1016/j.trb.2013.09.009

Burns, S., & Gross, G. (1990). Value of service reliability. IEEE Transactions on Power Systems, 5(3), 825–834. https://doi.org/10.1109/59.65911

Coutto Filho, M. D., Da Silva, A. L., Arienti, V., & Ribeiro, S. (1991). Probabilistic load modelling for power system expansion planning. 203–207

Da Silva, A. L., Ribeiro, S. P., Arienti, V., Allan, R., & Do Coutto Filho, M. (1990). Probabilistic load flow techniques applied to power system expansion planning. IEEE Transactions on Power Systems, 5(4), 1047–1053.

Fan, S., Methaprayoon, K., & Lee, W. J. (2007). Short-term Multi-Region Load Forecasting Based on Weather and Load Diversity Analysis. 562–567. https://doi.org/10.1109/NAPS.2007.4402366

Forsten, K. (2010). Tomorrow's T&D: the most economical energy savings might be found in grid efficiency. Public Utilities Fortnightly (1994), 148(2), 22

Gupta, H. (2018). Evaluating service quality of airline industry using hybrid best worst method and VIKOR. Journal of Air Transport Management, 68, 35–47. https://doi.org/10.1016/j.jairtraman.2017.06.001

Jasmon, G., & Lee, L. (1991). Distribution network reduction for voltage stability analysis and loadflow calculations. International Journal of Electrical Power & Energy Systems, 13(1), 9–13

Kramer, D. (2016). Models poised to boost grid efficiency. Physics Today, 69(9), 25–27. https://doi.org/10.1063/PT.3.3293

Park, T., & Kim, K.-J. (1998). Determination of an optimal set of design requirements using house of quality. Journal of Operations Management, 16(5), 569–581. https://doi.org/10.1016/S0272-6963(97)00029-6

Singh, B., Grover, S., & Singh, V. (2017). Evaluation of benchmarking attribute for service quality using multi attitude decision making approach. International Journal of System Assurance Engineering and Management, 8(2), 617–630. https://doi.org/10.1007/s13198-016-0485-4

OMIR EVENT TYPE AND FREQUENCY AT A COLLEGIATE FLIGHT PROGRAM

Brooke Wheeler[1]
Kevin Redden[1]
Davor Bonilla[1]
Gregory Fox[1]
Isaac Silver[1]
[1]College of Aeronautics, Florida Institute of Technology
bwheeler@fit.edu; kredden2015@my.fit.edu; dbonilla2015@my.fit.edu; gfox@fit.edu; isilver@fit.edu

Abstract

Safety is the priority in aviation; policies adopted by 14 CFR Part 141 Flight Schools aim to increase safety awareness while conducting flight operations. The additional evaluations through knowledge tests and check rides at each stage of flight training can may contribute to student stress. Collegiate Part 141 flight schools have students with increased workloads and stress from the combination of courses and flight training. This ex post facto research evaluated the types of events and the frequency of Operational Mishaps and Incident Reports (OMIRs) at a collegiate, Part 141 flight school in the southeastern US. The study reviewed all reports during a three-year period and classified them into four categories: Federal Aviation Regulation (FAR) violation, Flight Operations Manual (FOM), safety concern, or miscellaneous. Reports were evaluated across four-week intervals of the fall and spring semesters. There was no statistical difference between the number of OMIRs submitted by 4-week interval during the semester. This indicated that OMIRs are submitted in a relatively constant quantity throughout the duration of each semester. There was a significant difference in number of OMIRs by event type. Most OMIRs were Flight Operations Manual violations or safety concerns. This is a positive outcome from a safety perspective because the most frequent reports were not violations of FARs. These results also indicated that pilots are actively participating in the safety management system, submitting reports, even for relatively minor concerns.

1. Introduction

Safety management systems (SMSs) ensure that safety is a key goal for 14 CFR Part 141 Flight Schools. Part 141 flight schools that offer collegiate degree programs in tandem with flight training have students balancing the workload and stressors of courses and flight training over the semester. Consequently, it is important to understand the relationship between safety issues and time of the semester. The purpose of this study is to evaluate the type of OMIRs reported from spring 2015 through fall 2017 at a Part 141 collegiate flight school in the southeastern U.S. and when these OMIRs occur during the semester. Grouping will be determined by reading the OMIRs and descriptions and determining whether each OMIR infringes 14 Code of Federal Regulations (CFR) Aeronautics and Space or the Flight Operations Manual (FOM) (FITA, 2018). The intervals of the semester will be divided as follows, weeks one to four, weeks five to eight, weeks nine to twelve, and weeks thirteen to sixteen.

This analysis will provide information on the types of OMIRs collected and the timing of OMIRs throughout the spring and fall semesters, from the spring semester of 2015 through the fall semester of 2017. The findings obtained from this research will also provide a greater understanding of how the grouping of OMIRs changed across these semesters. This information increases our understanding of which parts of the semester have increased risks for mishaps and when student pilots are most likely to experience an OMIR-worthy event at this flight school. Flight schools could use this information to increase awareness for voluntary and self-reporting during the semester and implement another layer of risk management by having instructors at key points during the semester. Each flight school has a unique safety culture and SMS; therefore, the results are applicable to this flight school. However, the results of this study are relevant to all 14 CFR Part 141 flight schools that offer degree programs and provide flight training in connection with college courses because Part 141 schools have curricula that are highly standardized by the FAA and similar student populations.

2. Background

Safety culture can be studied through the effectiveness of the safety reporting system at a flight school and how it may influence student pilots to report incidents. The influences in submitting an OMIR can be simplified into two categories: cognitive aspects and safety culture. The cognitive aspects of submitting an OMIR are examined as ethical and psychological influences.

Ethical behavior is a key part of being a professional pilot. Pilots are expected to have a high sense of morality and should report all unsafe incidents. They are trained, educated, and paid highly to pilot aircraft with passengers onboard, yet it is expected that some pilots will not voluntarily report any safety violations for fear of punishment and destruction of their livelihood (Stanford & Homan, 1999). One way some collegiate flight schools have attempted to instill a higher sense of responsibility in their flight programs is to require students to take an ethics course on campus. This may seem a fitting solution, but according to Gonzales, Walter, and Morote (2011), a simple ethics course did not change a student’s behavior towards more ethical conduct. The study was conducted using five accredited higher education institutions, two with a required ethics course for pilots and three with no required ethics course. They utilized a survey including 32 questions on academic dishonesty and ethical standards and received 150 responses. One outcome that differed was that students who did take the ethics course rationalized and took responsibility for their actions, while those without the ethics course shifted the blame (Gonzales et al., 2011). This improvement in ethical conduct merits the idea of requiring Part 141 flight schools to incorporate ethics into their curricula to increase student responsibility, and therefore, voluntary reporting of safety issues. Although submitting OMIRs is the ethical response and could help to prevent future incidents or accidents, student pilots may be wary of the repercussions reporting may have on their future.

Part 141 flight school students have many pressures while training. Students must balance classes and flight training to succeed. Acute stress is a key psychological factor in collegiate aviation driven by classwork and flight training. Seiler (2013) found that the most frequently failed lessons for student pilots were those before an important milestone such as first solo flights and end of course check-rides. In a study of 67 students, 45% failed the lesson prior to the end-of-course check-ride. Pilot performance can be affected by three types of stress: environmental stress, acute stress, and life stress (Stokes & Kite, 1994). These external stressors could be from other areas of life, as is the case when academics takes priority over flight training, putting flight training as a secondary priority to remain in good academic standing and continue Part 141 flight training. Thus, managing stressors is critical to safety and success of student pilots.

If instructors took the time to assess students’ physical and mental state before making the go/no-go decision, learning and satisfaction between student and instructor could be improved (Lindseth, 1994). Robertson and Ruiz (2010) found that students in flight training perceived check-rides and written exams to be the most stressful situations, followed by finances and workload. As soon as students near milestones, they may understandably become stressed. Research has shown that a negative mood state significantly impacts a pilot's ability to recall complex information (Wendell, Recascino, DeLuca, & Kirton 2003), thus impacting performance. The stress of reporting OMIRs can be daunting for student pilots and relates perfectly to the ethicality of reporting, especially when other factors of school and workload can affect available time and willingness to report.

The safety culture of a flight school is a critical factor in OMIR submissions. A well-established safety culture can ensure students, instructors, and management feel a sense of responsibility and a desire to maintain a safe atmosphere at the flight school. On a local level, a safety culture at an organization is built from the top down. Management needs to be involved to facilitate an organization-wide movement. Flight instructors play a key part in instilling the safety culture and must have an established set of values and safety guidelines that they pass on to student pilots (Freiwald, Anderson, & Baker, 2013). As flight instructors operate as the interface between the flight school and the student, they must impart the safety culture of the organization to maintain a safe outlook and desire to continually improve safety at the school.

Incidents are often the result of errors and/ or questionable judgments of the flight crew (Stanford & Homan, 1999). Aviation in the United States and its safety culture have improved over the past three decades, due in part to the Aviation Safety Reporting System (ASRS). ASRS was developed for the FAA to obtain voluntary reporting data that could be used to prevent accidents and incidents from occurring by sharing information. To revisit the ethicality of reporting systems, the FAA has had to entice those who would not voluntarily submit safety reports by offering immunity from any penalties that may be imposed after an investigation (Stanford & Homan, 1999). The fact that the FAA has had to go to such lengths to solicit safety reports gives insight into pilots' concerns of punitive measures. For safety to grow, the aviation community must work together to increase safety.

The Aviation Safety Action Program (ASAP) is currently part of the SMS utilized by the flight school. The ASAP is a proactive safety program aimed at identifying any unsafe operations and implementing rules to correct them (Cusick, Cortes, & Rodrigues, 2017). ASAP reporting is voluntary and offers the same non-punitive reprimand as an ASRS SMS system. The SMS is built from the ground up by the airline or flight school, incorporating ASAP and relying on the safety culture. ASAPs are reviewed periodically by the Event Review Committee (ERC) to determine patterns and discuss possible solutions to prevent future reports of similar nature (Cusick, Cortes, & Rodrigues, 2017). The SMS is built with the flight school in mind and focuses on issues that the school faces, working to improve safety and resolve any possible hazards before any accident or incident can occur.

The FAA understood that pilots would be unwilling to self-report anything that may lead to enforcement action. The introduction of the ASRS and the ASAP has proved beneficial when it comes to safety; it emphasized the current safety culture in the industry. This can also be applied to Part 141 flight schools' OMIRs. As the OMIR has a chance for "enforcement action", whether it be by the school or by the government, the immunity and non-punitive measures of self-reporting encourage pilots to report incidents.

Voluntary reporting is dependent on a well-maintained safety culture. The fear of the FAA and other enforcement agencies can easily detract from safety culture. Pilots are concerned that a report will be used as an admittance of guilt and lead to legal trouble (McMurtrie & Molesworth, 2018). The FAA has worked to provide means of immunity for those who decide to report their inadvertent errors. The reports submitted voluntarily provide a detail of situations and incidents that may not have otherwise been reported. A study conducted with 234 airline pilots found one third of pilots had no confidence in their company's SMS system and decided not to report any safety incidents.

Approximately 59% of the pilots felt that information they submitted would be used against them in enforcement action proceedings (McMurtrie & Molesworth, 2018). This fear of reprisal is a major deterrent to pilots who would submit reports, and underscores the benefits offered by reporting systems within a strong safety culture and/or with immunity protections.

SMS voluntary reporting such as ASAP and ASRS are tools that the FAA has implemented to foster responsibility and continual improvement of safety in aviation. Stress is a major factor in performance. As Part 141 flight students continue to take collegiate course concurrently with flight training, research is needed on both the types of errors and times of the semester that these mishaps are likely to occur.

3. Methodology

This ex post facto study collected deidentified, archival data on OMIRs from a Part 141 collegiate flight school in the southeastern U.S. to address two questions: 1) is there a difference in number of each OMIR event type submitted, and 2) is there a difference in the frequency of OMIRs submitted based on 4-week intervals of the semester?

OMIR data was collected continuously and maintained in an archive for the flight school's SMS. The information requested was received in an Excel file and had been deidentified by management at the flight school before sharing. The Institutional Review Board approved an exemption (18-168). Records and student pilot data were protected by de-identification prior to release, keeping the data files secure during the research process, and only reporting data in aggregate.

Grouping was determined by reading each OMIR and classifying it by type of infringement: 14 CFR (FAR), the FOM, a safety concern, or a miscellaneous issue. The four intervals were selected to cover the most stressful periods of the semester: beginning (weeks 1-4), pre-midterm (weeks 5-8), post-midterm (weeks 9-12), and end of the semester and final exams (weeks 13-16).

The data included all OMIRs submitted by flight students and instructors at the flight school during the calendar years 2015 through 2017. We divided the OMIRs based on which interval of the semester they occurred in using Microsoft Excel and the predetermined semester intervals. Inferential statistics were calculated in R Studio version 3.5.3. We used a one-way ANOVA to test for a difference in OMIRs by event type. A one-way ANOVA was also used to test the effect of 4-week semester intervals.

4. Results

A total of 689 OMIRs were received during a 3-year period: 259 in 2015, 231 in 2016, and 199 in 2017. One record from 2016 was excluded because it had mostly missing values, leaving 688 records for analysis. The descriptive statistics for OMIRs by event type are shown in Table 1 and Figure 1. FOM violations and safety concerns. The ANOVA conducted in the first research question was significant: $F(3, 140) = 19.4$, $p < 0.0001$. A post hoc, Tukey's pairwise comparison showed that 4 out of 6 event type pairs had significant differences in frequency: FOM – FAR, Safety Concern – FAR, Miscellaneous–FOM, and Miscellaneous–Safety Concern. These event type pairings had the largest statistical difference between them, which corresponds with Figure 1. The eta-squared (η^2) was 0.29, which represents a small to medium effect size.

Table 2 and Figure 2 show the descriptive statistics for OMIRs by four-week intervals during the fall and spring semesters. As can be seen in Figure 2, there is a slight difference in the number of OMIRS submitted across the semester. The ANOVA on OMIRs by 4-week interval was not significant: $F(3, 20) = 0.202$, $p = 0.9$. The eta-squared (η^2) was 0.03, which is a very small effect size.

Table 1: Descriptive statistics for OMIRs by event type

	FAR Violation	FOM Violation	Safety Concern	Miscellaneous
Mean	2.7	6.8	7.4	2.2
Mode	1.0	4.0	8.0	0.0
Median	2.5	5.5	7.5	1.0
Maximum	7.0	28.0	16.0	7.0
Minimum	0.0	1.0	1.0	0.0
Standard Deviation	2.0	5.7	3.7	2.2

Note: OMIR categories are Federal Aviation Regulation (FAR), Flight Operations Manual (FOM), Safety concern (any general safety concern), and Miscellaneous (does not fit any other category).

Table 2: Descriptive statistics for number of OMIRs by four-week interval during fall and spring.

	Weeks 1-4	Weeks 5-8	Weeks 9-12	Weeks 13-16
Mean	23.5	22.7	21.5	20.2
Median	23	18	21	19.50
Mode	0	18	0	0
Range	14 - 34	15 - 43	11 - 31	15 - 26
Standard Deviation	8.5	10.5	6.8	**5. 4.5**

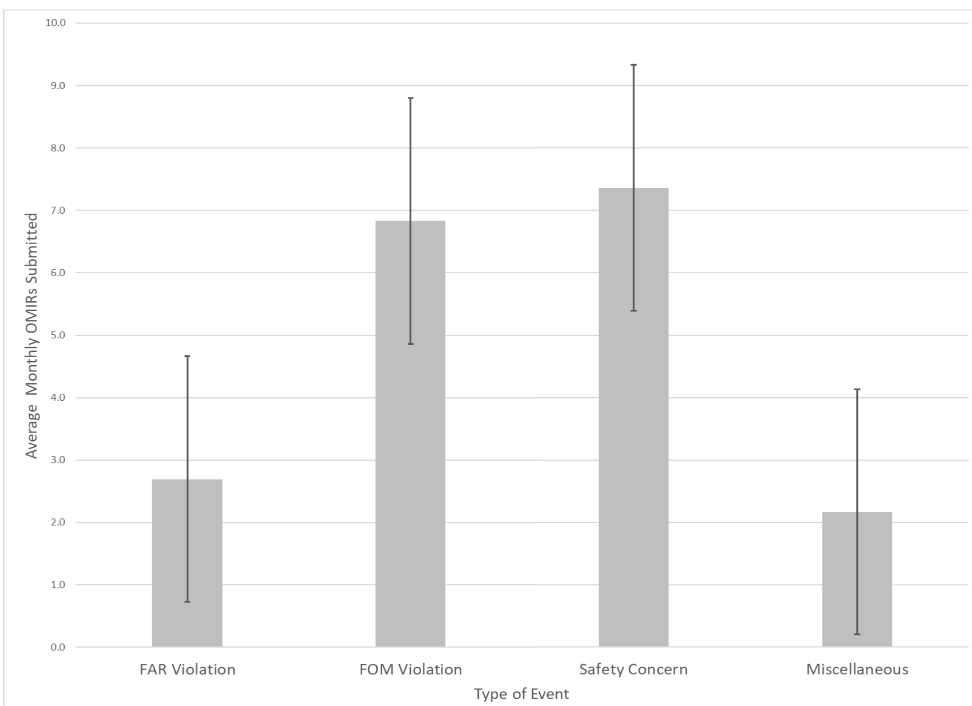


Figure 1. Monthly OMIRs by Type of Event

Note: OMIR categories are Federal Aviation Regulation (FAR), Flight Operations Manual (FOM), Safety concern (any general safety concern), and Miscellaneous (does not fit any other category).

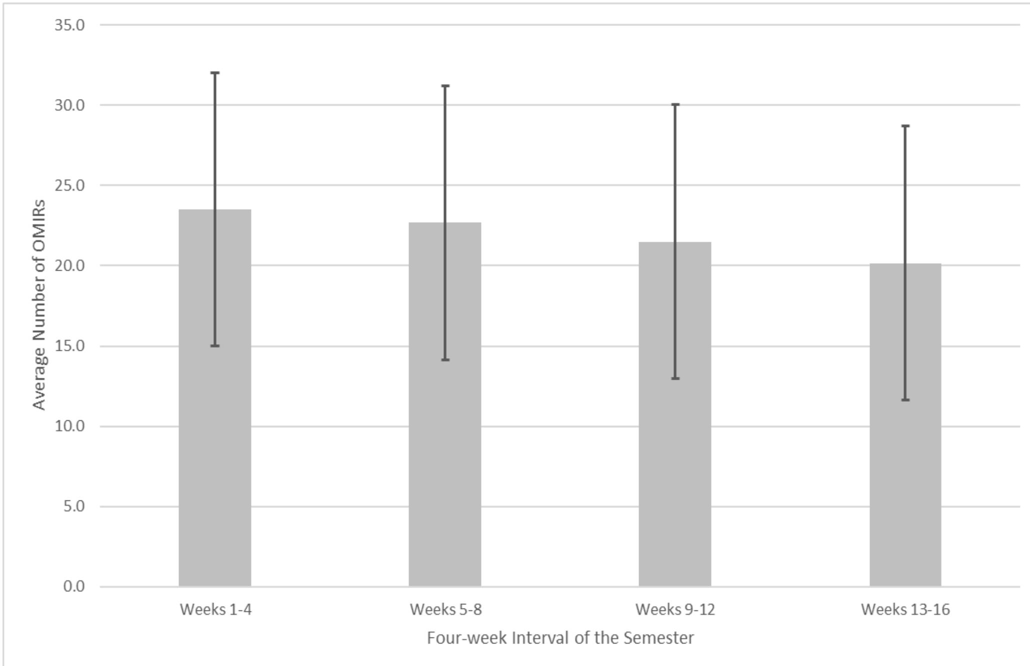


Figure 2. OMIR Frequency by 4-Week Interval of Semester

5. Discussion

The results supported our hypothesis that there would be a difference in the number of OMIRs by event type. Safety Concerns had the largest average number of OMIRs submitted (Table 1). Most OMIRs were in the groupings of FOM violation and safety concerns. These categories had significantly more reports than the FAR and Miscellaneous categories. The difference could be attributed to the fact that the FOM is more stringent than regulations placed on student pilots, which may lead to more student infringements of the FOM. From a safety perspective this is excellent because the goal is to catch unsafe behaviors before they are a true threat and retrain to avoid future issues.

The SMS at the flight school has been encouraging submission of OMIRs for any situation that is deemed unsafe. A key psychological factor in deciding to write an OMIR or any safety report is the safety culture at the organization. The flight school has implemented practices that introduced new flight students to the OMIR system and the SMS system at the very beginning, before flight training is initiated.

The results did not support our second hypothesis; there was no statistical difference in the number of OMIRs submitted by 4-week interval of the semester. The results indicated that OMIRs were submitted at a relatively constant rate across the semester intervals (Table 2). It was clear that there was not much difference between the means of each 4-week interval of the semester (Figure 2). It should be noted that there are no standard intervals during a semester, and these intervals were created by the authors. Future work should develop intervals based on surveys of student stress levels throughout the semester and actual exam, project, and paper due dates.

Studying OMIRs is critically important at Part 141 collegiate flight schools to identify and understand patterns that can be employed to improve safety and awareness. This study provided an initial assessment at a Part 141 flight school in the Southeastern US. While Part 141 flight training is highly standardized and overseen by the FAA, the student workloads may differ within a semester at different schools, and each school has its own SMS and safety culture.

As collegiate Part 141 flight students continue to take college courses concurrently with flight training, more research is needed to determine whether there are times during the semester that incidents and mishaps are more likely to occur. Using this research, Part 141 flight schools can better manage and mitigate risks, focusing on the relevant types of OMIRs to save time, money, resources, and most importantly, save lives if these mistakes were made with no recovery action due to lack of knowledge or stress. This work should be extended to examine the relationship between college course workload and student pilots' performance in flight training. Future work should also examine multiple flight schools for comparison and confirmation of trends in Part 141 training.

6. References

Beaudin-Seiler, B. (2013). Private Pilot Progress: Where Do We Fall Down? Journal of Aviation/Aerospace Education & Research, 22(2), 31-39, doi:10.15394/jaaer.2013.1398

Desai, V. M., Roberts, K. H., & Ciavarelli, A. P. (2006). The Relationship Between Safety Climate and Recent Accidents: Behavioral Learning and Cognitive Attributions. Human Factors: The Journal of the Human Factors and Ergonomics Society,48(4), 639-650. doi:10.1518/001872006779166361

Cusick, S. K., Cortes, A. I., & Rodrigues, C. C. (2017). Commercial Aviation Safety (Sixth ed., pp. 244-256). McGraw-Hill Education.

F.I.T. Aviation, LLC. (2018). Flight Operations Manual. Melbourne, FL: Florida Institute of Technology.

Freiwald, D., Lenz-Anderson, C., & Baker, E. (2013). Assessing safety culture within a flight training organization. Journal of Aviation/Aerospace Education & Research, 22(2), 41-57.

González, R., Walter, R., & Morote, E. (2011). A contrast of ethical attitudes and practices between aviation students at schools with and without an ethics course for pilots. Collegiate Aviation Review, 29(2), 21-46.

Kerr, S., Johnson, V. K., Gans, S. E., & Krumrine, J. (2004). Predicting Adjustment During the Transition to College: Alexithymia, Perceived Stress, and Psychological Symptoms. Journal of College Student Development,45(6), 593-611. doi:10.1353/csd.2004.0068

Lawton, R., & Parker, D. (1998). Individual Differences in Accident Liability: A Review and Integrative Approach. Human Factors: The Journal of the Human Factors and Ergonomics Society,40(4), 655-671. doi:10.1518/001872098779649292

Lindseth, P. (1994). Flight Anxiety: Predictors and Implications for Learning. Journal of Aviation/Aerospace Education & Research, 4(3) doi:10.15394/jaaer.1994.1131

Robertson, M. F., & Ruiz, L. E. (2010). Perceptions of stress among collegiate aviation flight students. Collegiate Aviation Review, 28(1), 115-126.

Stanford, L., & Homan, W. (1999). A model of "applied ethics" in aviation safety: The aviation safety reporting system. Journal of Aviation/Aerospace Education & Research, 9(1), 39-44.

Stokes, A. & Kite, K. (1994). Flight stress: Stress, fatigue and performance in aviation. Burlington, VT: Ashgate Publishing Company.

Wendell, A., Frederick-Recascino, C., Deluca, J., & Kirton, T. (2003). The Effects of an Induced Negative Mood State on Student Pilot Learning. Journal of Aviation/Aerospace Education & Research, 13(1), 43-53, doi:10.15394/jaaer.2003.1566

THE EFFECTIVENESS OF SMART COMPOSE: AN ARTIFICIAL INTELLIGENT SYSTEM

Matt Gnacek[1]
Eric Doran[1]
Sharon Bommer[1]
Philip Appiah-Kubi[1]
[1]University of Dayton
gnacekm1@udayton.edu; dorane2@udayton.edu; bommers1@udayton.edu;
pappiahkubi1@udayton.edu

Abstract

The growth of artificial intelligence (AI) technologies in everyday life and manufacturing are expected to reduce the mental workload of a user or human operator and increase their efficiency. In industrial systems, such as additive manufacturing (AM), as AM transitions from a technology of manufacturing prototypes to rapid manufacturing, it is important that these added technologies reduce an operator's mental workload, have high user satisfaction, and are easily implemented and incorporated into the operator's tasks. One growing AI technology is Smart Compose, an artificially intelligent system that provides writing suggestions when composing an email through Gmail. Like other AI technologies, the goal of Smart Compose is to enhance the performance of a user; in this case, typing an email. It is hypothesized that Smart Compose increases a user's performance when typing an email. Thus, the objective of this study is to test the capability of Smart Compose and whether it increases human performance and decreases mental workload for college students. The study found that Smart Compose does not significantly increase human performance and does not significantly decrease mental workload for college students

1. Background

Artificial intelligence (AI) has become a relevant topic in many fields, including engineering and computer science. In 1951, Alan Turing was the first to explore the question of whether humans could develop a machine that could think, which led to the term "artificial intelligence" being coined by John McCarthy in 1956 (Tugui, Danciulescu, & Subtirelu, 2019). McCarthy pictured AI as "the science and engineering that tries to make machines intelligent, trying to get them to understand human language and to reach problems and goals as well as a human being" (Garcia, Nunez-Valdez, Garcia-Diaz, Garcia-Bustelo, & Lovelle, 2019). Many computer scientists of the late 1950s and 1960s predicted that machines would be able to work, think, and reason and become intelligent like humans within the following years or decades (Atkinson, 2018). While that outlook was certainly overly optimistic, engineers and scientists have developed and grown several different forms of AI since then, with 2,000 new patents created for AI-based technologies between 2012 and 2016 alone (Fujii & Managi, 2018).

AI-based technologies can be broken down into two general subcategories: weak or narrow AI and artificial general intelligence or strong AI (Lu, Li, Chen, Kim, & Serikawa, 2018). Weak AI or narrow AI is a specialized form of AI that only performs specific tasks such as only recognizing a person through facial recognition but not verbal recognition (Lu, Li, Chen, Kim, & Serikawa, 2018). While its scope is limited, weak AI is personalized for a specialized task and unable to solve other tasks without being modified (Wirth, 2018). Weak AI is capable and efficient at solving *complex* problems, but it is unable and struggles with solving *difficult* problems. Complex problems are quantitative and can be traced by rules or units such as playing a game of chess (Lu, Li, Chen, Kim, & Serikawa, 2018). On the contrary, difficult problems require some sort of fine or gross motor skills, such as tying one's shoes (Floridi, 2019). When given complex problems and quantitative data, weak AI can be effective and efficient at the given tasks.

Artificial general intelligence (AGI) is vastly different than weak AI and has not been successfully developed by humans. While weak AI does not have motor skills, AGI is achieved when a software program "controls itself autonomously, with its own thoughts, worries, feelings, strengths, weaknesses and predispositions" (Goertzel & Pennachin, 2007)**.** Currently, AI systems are efficient at following patterns but there is no AI that can act as a human brain or form a relationship between the human mind and the human body (Lu, Li, Chen, Kim, & Serikawa, 2018). Until general AI is successfully developed, scientists and engineers will continue to use weak AI technologies such as machine learning and deep learning. One weak AI technology that uses machine learning and deep learning is Smart Compose. Smart Compose is designed to increase the performance of a user typing an email on Gmail and reduce their mental workload like other weak AI technologies. Thus, it was hypothesized that Smart Compose significantly affects the performance of a user typing an email on Gmail and mental workload. For this study on Smart Compose, the focus was on a college student's capability with Smart Compose due to a large portion of college students using Gmail on a regular basis.

1.1. Machine Learning

Machine learning (ML) is one of the most common forms of AI that "is based on the development of learning techniques to allow machines to learn, based on the analysis of data and applying some algorithms for this processing and others for decision making from the data already analyzed" (Garcia, Nunez-Valdez, Garcia-Diaz, Garcia-Bustelo, & Lovelle, 2019). Essentially, a machine is presented with some data, and it analyzes said data via a decision-making algorithm in order to *think*. When presented with this data, ML networks use one of three primary modes of learning to process the data, which are supervised, unsupervised, and reinforcement learning (Stinis, 2019). For this paper, the focus will be on supervised learning and the two major types of networks that use supervised learning - classification oriented networks and regression-oriented networks (Welsh, 2019).

Classification oriented networks match a classification to data in hopes that, over time, a machine can recognize similarities in the classes and analyze data correctly into those classes (Welsh, 2019). An example of this would be recognizing different types of trees where a machine could be given thousands of images of palm trees, fern trees, and oak trees with the respective class tied to each image. Over time, the machine would develop an algorithm to determine the type of tree that is shown. The major downside to this network is that if the class is not predetermined by a human, the network will not be able to classify them correctly (Welsh, 2019). If an image of a cactus would be given to the machine described above, it would misclassify the cactus due to its lack of prior knowledge on the cactus in the machine.

Regression networks use inputs and weighted nodes to pass the data onto another set of nodes, which eventually leads to an output (Welsh, 2019). Regression networks use a decision tree, which is set up as a binary tree. The data traverses the tree until an output is reached (Garcia, Nunez-Valdez, Garcia-Diaz, Garcia-Bustelo, & Lovelle, 2019). It is common for multiple regression networks to be connected, and the trees can be combined to create a forest, which allows for nodes to be passed back into the same tree, often resulting in better data (Welsh, 2019). Regression networks can be used in conjunction with deep learning techniques, which are described more in the next section of this paper.

1.2. Deep Learning

Artificial neural networks (ANNs) are non-linear classification systems that attempt to model the human brain (Prieto, et al., 2016). ANNs can "learn" about data, and they feed into a neural network (NN), which contains weighted nodes or neurons that are used to create a data path to an output. The weighted nodes in neural networks are commonly used in deep learning (DL), the most common subcategory of ML. DL was first explored in its current form around 2006, when multi-layer NNs and deep learning models were created (Yu, 11). However, the term "neural network" was coined by Warren McCulloch and Walter Pitts in 1943, and D.O. Hebb was the first to weight different neurons and connect them (Prieto, et al., 2016). The goal of DL is to map data from a bottom layer to an upper layer in a hierarchical manner, like that of the human brain (Lu Y. , 2019). There are many layers in DL networks, and "deeper" layers in the network tend to be more abstract (LeCun, Bengio, & Hinton, 2015). It is also important to note that the layers are not designed by human engineers; rather, the machine generates the layers by "learning" from data received (LeCun, Bengio, & Hinton, 2015). The key to deep learning is assigning weights to each neuron in a network correctly and then doing so for several layers (Schmidhuber, 2015). In a neuron tree, one weighs a node to provide a certain magnitude to that node to increase the magnitude between the node initial. If a node has a higher weight than another node, the former node will have a greater influence than the latter node.

One common network using DL concepts is the convolution neural network (CNN). CNNs use multiple layers to sort data that comes as an input in multiple arrays (LeCun, Bengio, & Hinton, 2015). The inputted data is passed through several filter banks, which is a set of weights that helps to process the data in a non-linear fashion. Each layer of the CNN has a separate data bank, and these data banks help to determine patterns and produce an output (LeCun, Bengio, & Hinton, 2015). CNNs are very good at processing images and sounds and are used in devices such as autonomous vehicles (LeCun, Bengio, & Hinton, 2015).

1.3. Current and Future Artificial Intelligence in Industry

Several industries have begun to flourish by using one or more forms of AI technologies to enhance products or technology that they are producing. The most noticeable trend is the consistent use of DL techniques with nearly every industry examining the uses of DL concepts, which makes sense with the recent advancements in neural networks. It is no secret that automating processes will allow for industries to be more efficient; however, should the trend continue, it will be interesting to see how it affects industries beyond efficiency. Another trend in AI technology is the growth of DL when it comes to AI-based technologies due to the success rates of multilayered neural networks. This is obvious in literature, where the number of publications related to DL skyrocketed from just over 650 in 2010 to over 7,400 in 2017 (Bostan, Ekin, Sengul, Karakaya, & Tirkes, 2018). A large part of this growth can be attributed to the success of convolutional neural networks (CNN) that have found to be effective with object recognition and scene analysis (Ibrahim & Qaisar, 2019).

Finally, a trend that is growing and particularly relevant for this research is the combination of DL techniques and natural language processing (NLP). NLP attempts to allow machines to analyze human language; however, in the past, most NLP research had been done using shallow models (Young, Hazarika, & Poria, 2018). In the present, NLP concepts are being combined with DL techniques in order to create multilayered networks seeking to improve results and cut out previously used hand-crafted features that were time-consuming (Young, Hazarika, & Poria, 2018). The results have been positive for techniques such as word embeddings, which seek to sort similar words into vectors, and character embeddings, which seek to do the same, but using characters rather than words (Young, Hazarika, & Poria, 2018).

With these current trends in AI-based technology, it's important that AI technologies being implemented into various systems truly make them efficient and reduce the mental workload of the system's operators. This can be done through experimentations, by comparing the system with and without the AI technology and see if the AI technology significantly improves the efficiency of the system and reduces the mental workload on the operator. Two examples of industrial technologies that continue to grow in popularity and are designed to increase the efficiency of the system and reduce mental workload of the user or operator are autocomplete and autocorrect, both features in word processing.

1.4. Autocorrect

Autocorrect stems from predictive text systems, which have their roots as far back as the 1970s (Kidney and Anderson, 103). While autocorrection is widely accepted to be helpful in correcting obvious errors, it also is not easy to implement (Hasan et al., 451). In modern autocorrect systems, a dictionary is "learned," and then words that are not in the dictionary but commonly typed are added over time while common corrections that the user continuously makes are also added (Kidney and Anderson, 103).

Autocorrect in the modern-day uses a combination of ML techniques and natural language processing (NLP) techniques, as they tend to adapt to a user's word choice over time (Kidney and Anderson, 103). As a system learns the word choices that a user tends to choose, it can begin to predict when words are incorrect and can be changed (Kidney and Anderson, 103). There are currently systems that use neural networks, language models, and NLP, as well as systems that use neural network learning and database changes in order to implement autocorrect (Ouazzane, 5961). While there do not appear to be any insights that relate specifically to Google's version of autocorrect, it can be presumed that it uses these techniques as well. Seeing that much of the leading research on the topic tends to focus on similar techniques, it would make sense for a technological powerhouse to use the leading information that is available.

1.5. Smart Compose

Autocomplete is a similar technology to autocorrect. However, autocomplete is used to complete phrases and sentences, while autocorrect is typically used to complete a single word. Regardless, the deep learning methods that go into each technology is very similar. Smart Compose is Gmail's version of autocomplete. Using neural networks, it can provide users with suggestions for how to complete sentences, phrases, or words. Smart Compose was launched in 2018 (Wu, 2018). Smart Composes has three major challenges according to Wu, a principal engineer on Google's Brain Team. They are latency, scale, and fairness and privacy (Wu, 2018). Latency requires a quick response from the system. Scale requires that the system caters to a diverse body of users. Fairness and privacy are essential for preserving the user's personal information (Wu, 2018).

2. Methodology

A within-subject design experiment was conducted to test the capability of college students using Smart Compose, which is comprised of four major parts. The results and data collected from these four parts will then be analyzed using IBM SPSS Statistics software. The first part consisting of collecting demographic information on a participant through a 5-question multiple-choice Pre-Questionnaire and a Typing Test. Each participant was asked to answer four questions, and a fifth question was posed to those who responded in a certain manner to another question. Out of the 21 participants, only two did not answer the fifth question, while all 21 participants answered the first four questions. The first question enquires on the gender of the participant while the second question request the age of the participant. The third question requests the participant to rate their typing skills as either Novice, Intermediate, or Advanced. The fourth question inquires on the participant's familiarity with the autocomplete function used in Gmail. If the answer is yes, the participant is questioned whether they believe autocomplete is a useful tool or not.

Following the completion of the Pre-Questionnaire, an online typing test sponsored by LiveChat was administered to each participant to test a participant's typing ability with the metrics of words per minute and accuracy being provided at the end of each test. Each participant had one minute in order to type as many words as possible with the words being randomly selected from the 1,000 most common words in the English language according to LiveChat.

The second part of the experiment gave a participant a 100-word sample email and had the participant to type the email for four iterations under different conditions that are outlined in Table 1. Smart Compose was enabled twice and disabled twice while a 90-second timer was enabled twice and disabled twice to add a stress variable in order to see how the participant performed and felt when time was limited to emulate an industrial setting where operators must adhere to a tight schedule. Table 1 shows the four trials of the test below.

Table 1. Typing test trials

Trial 1 Smart Compose: OFF 90 Second Timer: OFF	Trial 2 Smart Compose: OFF 90 Second Timer: ON
Trial 3 Smart Compose: ON 90 Second Timer: OFF	Trial 4 Smart Compose: ON 90 Second Timer: ON

Each participant had the trials randomized using a MATLAB script and to avoid the data from being skewed due to participants becoming proficient with the email by the fourth trial. When the 90 second timer was enabled, the participant could not see the timer with the participant with the experiment administrator telling the participant that they had 90 seconds to complete the trial, and once the timer expired, participants were instructed to stop typing by the experiment administrator. Though Trials 2 and 3 had the Smart Compose feature on, the user was not required to use Smart Compose when it was enabled.

The third part of the experiment had each individual fill out a NASA Task Load (NASA-TLX) form following the completion of each trial. NASA-TLX assesses six subjective mental workload measures based on an individual's perceived mental workload of the task. The form measures mental demand, physical demand, temporal demand, performance, effort, and frustration so it could be analyzed later which trials had the highest and lowest demands. Mental demand enquires on how mentally demanding the task on the participant was while physical demand tests how physically demand the task was for the participant. The participant is also asked about their temporal demand which is the pressure felt by the participant based on the pace of the task given. Next, the participant is questioned about their performance with the task regarding how successful they were in accomplishing the given task and about how much effort they had to give to get to their level of success with the task. Finally, the amount of frustration the participant had during the task was examined.

Lastly, the fourth and final part of the experiment had a System Usability Scale administered after all four iterations and all four NASA-TLX forms were completed. The form had ten different statements, and the user was able to select an answer from 1-5, with 1 being “Strongly Disagree” to 5 being “Strongly Agree” regarding the Smart Compose technology. The purpose of the scale was to collect data from each participant regarding the usability of Smart Compose in a qualitative manner.

3. Results

The impact of Smart Compose on college students’ typing efficiency was the focus of this study, hence, data was collected from a sample of twenty-one different college students, all of whom were 18-25 years old. Each participant voluntarily completed the experiment with the option to stop at any point throughout the experiment. The Pre-Questionnaire provided data on the participants. From the data obtained, the gender, age range, and typing skills of participants were collected. The results of the Pre-Questionnaire (N = 21) had 11 females and 10 males. Of the 21 participants, 13 participants rated themselves as intermediate typists with 5 suggesting that they were novice the other 3 suggesting they were advanced typists. The results from question #4 in the Pre-Questionnaire were that 19 participants stated they there were familiar with Autocomplete and the other 2 participants saying they were not. The fifth and final question in the Pre-Questionnaire asked the participants if they thought autocomplete was a useful technology. Out of the 19 participants who were familiar with Smart Compose, 17 of them believed that it was a useful technology while the other two believing otherwise.

As previously stated, each participant took the LiveChat typing test with their word per minute and accuracy results being recorded with accuracy being the measure of human performance. A word was deemed incorrect if it was spelled incorrectly, capitalized incorrectly, punctuated incorrectly, or missed. There is an exception to a missed word, which was if there were extra spaces at the end of a paragraph, the accuracy was not affected. The words per minute were the main statistic that was examined; the chart below (Figure 1) shows the words per minute each user typed as well as their self-classification of their typing skills. The lines shown in the graph are 25th percentile (43 words/min) and 75th percentile (61 words/min) lines. A statistical test was done to test if there is a significant difference in the words per minute between the three groups of typists (Novice, Intermediate, and Advanced). With all sets of data being normally distributed and with homogeneous variances ($p > 0.05$), an analysis of variance (ANOVA) was conducted. It was found that there was a significant difference between the words/minutes between Advanced and Novice typists ($p < 0.05$) while there were no significant differences between both Advanced and Intermediate typists and Novice and Intermediate typists ($p > 0.05$). Additional summary statistics can be seen in Table 2.

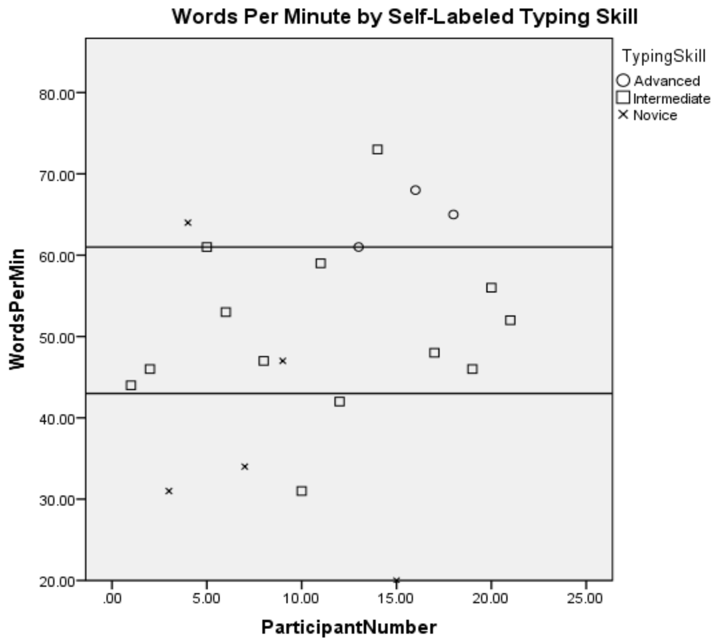


Figure 1. Words per minute by self-labeled typing skill

The average accuracy for each trial is shown below in Table 3. The accuracy means for when Smart Compose was on (Trial 1 and 2) appear to be comparable to when Smart Compose was off (Trial 3 and 4). On the other hand, Trial 2 and 4, which had the timer active seem to be at least .10 or 10% less than the accuracy averages of Trial 1 and 3, which did not use the timer. Thus, two statistical tests were completed: one tested if Smart Compose significantly affects a user's performance while the other tested if the 90-second timer significantly affects a user's performance.

Table 2. ANOVA test results on words per minute

(I) TypingSkill	(J) TypingSkill	N	Mean	Standard Deviation	Sig.[b]
Advanced	Intermediate	13	50.615	10.332	0.228
	Novice	5	39.200	16.873	0.023
Intermediate	Advanced	3	64.667	3.512	0.228
	Novice	5	39.200	16.873	0.237
Novice	Advanced	3	64.667	3.512	0.023
	Intermediate	13	50.615	10.332	0.237

To test the significance of Smart Compose, the sample data of Trial 1 and Trial 2 when Smart Compose was off and Trial 3 and Trial 4 when Smart Compose was on were compared to see if there was a statistically significant difference in the means. The data was not normally distributed ($p < 0.05$). Thus, the Related-Samples Wilcoxon Signed Rank Test was conducted, and found human performance in typing an email through Gmail with the use of Smart Compose did not significantly affect a user's performance since it's p-value is more than 0.05 ($p = 0.173$).

To test the impact of the 90-second timer, the sample data of Trial 1 and 3 when the timer was off, and the sample data of Trial 2 and 4 when the timer was on were compared. Since the data was not normally distributed ($p < 0.05$), a non-parametric test revealed that human performance in typing an email with a 90-second timer does significantly affect a user's performance in typing an email through Gmail since it's p-value is less than 0.05 ($p = 0.000$).

Lastly, a multivariate analysis of variance (MANOVA) test was done on the accuracy averages of the three groups of typists to see if there were any significant differences. It was observed that an Advanced typist's accuracy average on Trial 4 was significantly different than a Novice typist's accuracy average on Trial 4 ($p < 0.05$). All the results and p-values from the MANOVA test can be seen in Table 4 while the F-statistic for typing skill can be seen in Table 3.

Table 3. Accuracy averages

Source	Dependent Variable	N	Mean	Standard Deviation	F	Sig.
TypingSkill	Trial_1	21	0.971	0.049	0.775	0.475
	Trial_2	21	0.823	0.166	3.428	0.055
	Trial_3	21	0.962	0.055	0.446	0.647
	Trial_4	21	0.816	0.181	5.592	0.013

Next, the NASA-TLX forms were examined and averaged in order to compare subjective mental workload results, which can be seen under Table 5. Then all six of subjective mental workload measures of the NASA-TLX were added together and divided by 6 to get an overall average mental workload for each trial. Those overall averages were then used to conduct a pair of non-parametric t-tests that tested if there is a significant difference in the overall mental workload when the timer is off (Trial 1 and Trial 3) and when the timer is on (Trial 2 and Trial 4) and if there was a significant difference between the overall workload when Smart Compose is off (Trial 1 and Trial 2) compared to when Smart Compose is on (Trial 3 and Trial 4).

Table 4. MANOVA test on typist's accuracy averages

Dependent Variable	(I) TypingSkill	(J) TypingSkill	N	Mean	Standard Deviation	Sig.[b]
Trial_4	Advanced	Intermediate	13	0.842	0.137	0.393
		Novice	5	0.644	0.210	0.015
	Intermediate	Advanced	3	0.993	0.006	0.393
		Novice	5	0.644	0.210	0.066
	Novice	Advanced	3	0.993	0.006	0.015
		Intermediate	13	0.842	0.137	0.066

The non-parametric t-tests between the NASA-TLX variables when the timer is off versus when the timer is on found that mental demand ($p = 0.001$), physical demand ($p = 0.000$), temporal demand ($p = 0.000$), performance ($p = 0.000$), effort ($p = 0.000$), and frustration ($p = 0.000$) were all significantly different from one another.

For the NASA-TLX performance averages significance test on Smart Compose, a test of normality of the data was done and found that most of the data was not normal. Since the normality assumption was not met, a Wilcoxon Signed-Rank test was done to see if there was a significant difference any of the NASA-TLX performance averages when Smart Compose is off versus when Smart Compose is on. The significance tests between NASA-TLX performance averages when Smart Compose is off versus when Smart Compose is on found that there were no significant differences between any NASA-TLX performance average when Smart Compose off compared to when it was on.

Table 5. NASA-TLX performance averages

	N	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration	Overall Average
Trial 1	21	23.810	20.714	20.000	19.048	31.667	18.810	22.341
Trial 2	21	37.619	28.333	61.190	42.857	56.905	35.238	43.690
Trial 3	21	29.286	23.571	25.238	21.667	30.714	27.381	26.310
Trial 4	21	40.238	29.762	63.810	43.571	54.762	37.857	45.000

Lastly, the results of the Systems Usability Survey mirrored the responses from the Pre-Questionnaire. The results found that the 21 participants strongly disagreed that they needed to learn a lot about Smart Compose before using it (1.38), strongly disagreeing that they needed a technical person to operate Smart Compose (1.10), strongly agreed that Smart Compose was easy to use (4.33), and strongly agreed that most people would learn Smart Compose quickly (4.52). The perceived usefulness of Smart Compose as a program by the participants contrasts Smart Compose's actual effectiveness as a system for reducing a user's mental workload as well as increasing the user's performance.

To see if the perceived usefulness of Smart Compose as a program may vary based on one's typing skill, a significance test was done to test the ten Systems Usability Survey questions and if they are significantly different based on the type of the typist. A test of normality was done to see if the data met the normality assumption for a MANOVA test and found the data is not normal. Though the data was not normal, a MANOVA was still done. The MANOVA test results which can be seen in Table 7 found that the Advanced typist's and Novice typist's answers from Question #5 (p = 0.029), the Intermediate's typist and Novice typist's answers from Question #5 (p = 0.017), and the Intermediate's typist and Novice typist's answers from Question #10 (p = 0.043) were all significantly different while Table 6 provides the F-statistic.

Table 6. Between-Subjects effects

Source	Dependent Variable	N	Mean	Standard Deviation	F	Sig.
TypingSkill	Systems Usability Q1	21	3.571	1.076	0.202	0.819
	Systems Usability Q2	21	2.048	0.921	0.326	0.726
	Systems Usability Q3	21	4.333	0.856	1.337	0.287
	Systems Usability Q4	21	1.095	0.436	1.714	0.208
	Systems Usability Q5	21	3.762	0.889	6.030	0.010
	Systems Usability Q6	21	2.476	0.928	1.290	0.300
	Systems Usability Q7	21	4.524	0.602	0.382	0.688
	Systems Usability Q8	21	2.095	1.221	0.319	0.731
	Systems Usability Q9	21	3.762	0.944	2.449	0.115
	Systems Usability Q10	21	1.381	0.669	3.671	0.046

Table 7. MANOVA test on typist's systems usability survey answers

Dependent Variable	(I) TypingSkill	(J) TypingSkill	N	Mean	Standard Deviation	Sig.[b]
SystemsUsabilityQ5	Advanced	Intermediate	13	4.000	0.707	1.000
		Novice	5	2.800	0.837	0.029
	Intermediate	Advanced	3	4.333	0.577	1.000
		Novice	5	2.800	0.837	0.017
	Novice	Advanced	3	4.333	0.577	0.029
		Intermediate	13	4.000	0.707	0.017
SystemsUsabilityQ10	Advanced	Intermediate	13	1.154	0.376	1.000
		Novice	5	2.000	1.000	0.426
	Intermediate	Advanced	3	1.333	0.577	1.000
		Novice	5	2.000	1.000	0.043
	Novice	Advanced	3	1.333	0.577	0.426
		Intermediate	13	1.154	0.376	0.043

4. Conclusion

Smart Compose is an AI technology designed to increase a user's performance when typing an email through Gmail. Based on this pilot study, a operator's performance does not increase when Smart Compose is used ($p > 0.05$). As seen in the significance test on the NASA-TLX data, the operator's mental workload does not decrease when Smart Compose is active as well (see Table 7). On the contrary, the time constraint affects the operator's performance ($p < 0.05$) and mental workload (see Table 7). Though the Smart Compose did not improve a user's performance or reduce their mental workload, the Systems Usability Survey's results showed that users perceive that Smart Compose does help one's performance and is easily able to be used. Tests on the results also found that a typist's ability significantly affects certain measurables in this experiment. The results from the Systems Usability survey found that most of the groups perceived Smart Compose to be a useful technology that is easy to learn and use. The statistical test on the Systems Usability survey results found that a Novice typist's answer to Question #5 (questions how well the system's functions were integrated) was significantly different than an answer from an Advanced and Intermediate as well as a Novice typist's answer to Question #10 (questions if a participant needs to learn a lot about the system before using it) was significantly different than the answer from an Intermediate typist.

5. Limitations of Study

Smart Compose is generally perceived as a useful and easy to use tool for enhancing operator performance and reducing an operator's mental workload. This study observed that there was no difference in operator performance and workload with and without Smart Compose. However, some limitations of the study should be noted. The small sample size precludes the generalization of the findings. Also, the participants in the study were all college students at one institution, causing a lack of diversity in the demographics. Faculty members, administrators, and industrial workforce who write and respond to varieties of emails as part of their daily routine should be included in any future study for a more comprehensive review of the impact of Smart Compose on operator performance and workload. Hence, further studies with a stratified and larger sample size will be needed to validate the findings of this study. Moreover, additional studies about the efficiency of industrial AI applications such as Smart Compose will help businesses to make informed decisions and assure AI applications meet the goals of works systems design before implementing them.

6. References

Atkinson, R. D. (2018, October). "It Is Going to Kill Us!" and Other Myths About the Future of Artificial Intelligence. IUP Journal of Computer Sciences, 12(4), 7-56.

Bostan, A., Ekin, C., Sengul, G., Karakaya, M., & Tirkes, G. (2018). Deep Learning and Current Trends in Machine Learning. 2018 3rd International Conference on Computer Science and Engineering (UBMK) (pp. 467-470). Sarajevo: IEEE.

Floridi, L. (2019). What the Near Future of Artificial Intelligence Could Be. Philosophy & Technology, 32(1), 1-15.

Fujii, H., & Managi, S. (2018, June). Trends and priority shifts in artificial intelligence technology invention: A global patent analysis. Economic Analysis and Policy, 58, 60-69.

Garcia, C. G., Nunez-Valdez, E. R., Garcia-Diaz, V., Garcia-Bustelo, B. C., & Lovelle, J. M. (2019, March). A Review of Artificial Intelligence in the Internet of Things. International Journal of Interactive Multimedia and Artificial Intelligence, 5(4), 9-20.

Goertzel, B., & Pennachin, C. (2007). Artificial General Intelligence. Berlin: Springer.

Ibrahim, M. E., & Qaisar, A. (2019, January). Current and Future Trends of Deep Learning based Visual. International Journal of Computer Science and Network Security, 19(1), 155-160.

LeCun, Y., Bengio, Y., & Hinton, G. (2015, May 27). Deep learning. Nature, 521(1), 436-444.

Lu, H., Li, Y., Chen, M., Kim, H., & Serikawa, S. (2018). Brain Intelligence: Go Beyond Artificial Intelligence. Mobile Networks and Applications, 23, 368-375.

Lu, Y. (2019, March). Artificial intelligence: a survey on evolution, models, applications and future trends. Journal of Management Analytics, 6(1), 1-29.

Prieto, A., Prieto, B., Ortigosa, E. M., Ros, E., Pelayo, F., Ortega, J., & Rojas, I. (2016, November 19). Neural networks: An overview of early research, current frameworks and new challenges. Neurocomputing, 214, 242-268.

Schmidhuber, J. (2015, January). Deep learning in neural networks: An overview. Neural Networks, 61, 85-117.

Stinis, P. (2019). Enforcing constraints for time series prediction in supervised, unsupervised and reinforcement learning. 1-30. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=edsarx&AN=edsarx.1905.07501&site=eds-live

Tugui, A., Danciulescu, D., & Subtirelu, M.-S. (2019). The Biological as a Double Limit for Artificial Intelligence: Review and Futuristic Debate. International Journal of Computers, Communications & Control, 14(2), 253-271.

Welsh, R. (2019, January). Defining Artificial Intelligence. SMPTE Motion Imaging Journal, 128(1), 26-32.

Wirth, N. (2018, September). Hello marketing, what can artificial intelligence help you with? International Journal of Market Research, 60(5), 435-438.

Wu, Y. (2018, May 16). Smart Compose: Using Neural Networks to Help Write Emails. Retrieved from Google AI Blog: https://ai.googleblog.com/2018/05/smart-compose-using-neural-networks-to.html

Young, T., Hazarika, D., & Poria, S. (2018, August). Recent Trends in Deep Learning Based Natural Language Processing. IEEE Computational Intelligence Magazine, 13(3), 55-75.

DRONE DELAY: CONSUMER WILLINGNESS TO FLY AFTER AWARENESS OF RECENT UAS EVENT

Brooke Wheeler[1]
Nicolas Duboc[1]
Zachary St. Amand[1]
[1]College of Aeronautics, Florida Institute of Technology
bwheeler@fit.edu; nduboc2016@my.fit.edu; zstamand2015@my.fit.edu

Abstract

Unmanned Aerial Systems (UASs) operated in the National Airspace system increase the potential for hazardous operations by remote pilots as well as intentional disruption of commercial aviation. Clothier et al. (2015) indicated that the perceived risk of UASs was equivalent to manned aircraft; however, recent UAS events at airports (e.g. Lee, 2019) illustrated possibilities for disruptive operation of UASs, making it important to understand the impacts on aviation. This study examined whether there is a change in consumer willingness to fly (Rice et al., 2015) before and after being made aware of a UAS event by reading a news article. The study was a repeated measures design; a survey measured participants' willingness to fly, presented a news article (Lee, 2019) describing a UAS event, and then measured willingness to fly again. Sixty-one undergraduate students in core Aeronautic, Communication, and Business courses at the Florida Institute of Technology participated. The results indicated that consumer willingness to fly decreased after being made aware of the UAS event. The large effect size also suggested that consumers are affected by a news article describing a UAS event, and more research should be done in this area. As UASs become more common, their exposure to the public will also increase as the media begins to focus on events similar to those described in the news article. This can lead to a decrease in consumer willingness to fly that could have detrimental effects on aviation businesses as consumers will opt for different modes of transportation.

1. Introduction

Commercial Unmanned Aerial Systems (UASs) are now in common use and continue to gain popularity with both hobbyists and industry. As UASs integrate into the existing airspace system, they share airspace with thousands of manned, commercial operations. Consequently, it is important to determine how this will affect consumer willingness to fly. Perceptions of UASs may be intertwined with views of technology and automation, and these perceptions are likely to be influenced by new information on these systems, such as news articles on the impacts of UASs on aviation. The purpose of this study was to examine consumer willingness to fly before and after being presented with a news article about an UAS event. Participants rated their will rate their willingness to fly (Rice, et al., 2015), read a news article about a recent event (Lee, 2019), and then rated their willingness to fly again.

Recent UAS events, resulting in disruptions at major airports such as Newark, NJ, and Gatwick in London (Lee, 2019), underscored the need to understand the impact of news coverage on the aviation industry. This study provided baseline data on how consumer willingness to fly (Rice et al., 2015) differed after a treatment, reading news that increases awareness of a current UAS event. This study contributed to our understanding of the public's opinion on UAS, specifically as it relates to the impact of UAS on commercial aviation. Understanding the impact of UAS events and related news coverage is critical as UASs are currently active throughout the national airspace system. Consumer willingness to fly after exposure to news of UAS events has broad implications for commercial aviation, affecting revenue and profitability.

1.1. Regulation and integration into the National Airspace System

Though UASs are a relatively new integration into the National Airspace System (NAS), there are already several regulatory practices in place. In 2012, the U.S. Department of Transportation enacted the FAA Modernization and Reform Act. This act allowed for the integration of UAS to allow them to operate “harmoniously, side-by-side with manned aircraft, occupying the same airspace and using many of the same air traffic management (ATM) systems and procedures” (FAA, 2018, p.4). An official roadmap by the FAA described the appropriate advances and future integration of UASs into the NAS (FAA, 2018). To allow visual line of sight (VLOS) operations without having an exemption process, 14 Code of Federal Regulations (CFR) Part 107 was established in 2016. This has enabled thousands of UAS operators to function in controlled airspace with certain permissions. Safety continues to be the driving force behind regulation for the FAA, and this will continue with UAS regulations.

For all the progress that has been made in UAS integration, there are still several factors that have yet to be addressed. Examples include the development of Detect and Avoid systems (DAA), which have yet to be perfected enough in order to allow for the autonomous operation of UASs in the NAS. Because of this, specific legislation has been put in place to allow operation of a small UAS, categorized as weighing less than 0.55 pounds (FAA, 2018), only when it is in the operator’s VLOS. Though many details are yet to be determined, UASs have been introduced in the NAS, but the integration will not be complete without the proper interaction of the key players in the NAS. Brooker (2013) described the importance of Air Traffic controllers as a High Reliability Organization (HRO). Currently, there is some coordination between Air Traffic Controllers (ATC) and UASs, including airspace, timeframe, and weather. As commercial use of UASs increases, so too must the coordination between UAS pilots and ATC. Booker (2013) emphasized the importance of ATC as an HRO, and the key role ATC plays. Thus, as UAS operations increase so will the responsibility of ATC.

Zwickle, Farber, and Hamm (2018) gathered information about public perceptions of current UAS legislation and which areas were most important to the respondents. Participants were most concerned with privacy and keeping unwanted drones away from their private property. The respondents wanted more regulations to protect privacy and fewer regulations stopping the application of UASs for public safety (Zwickle et al. 2018). The public perception is that UAS regulations should be focused on applications and privacy rather than integration into the NAS.

1.2. UASs and safety

Plioutsias, Karanikas, and Chatzimihailidou (2018) conducted research about the safety systems that are programmed into 19 UAS devices available today. While automation does have a drawback from increased weight, the UASs used in Plioutsias et al. (2018) showed a level of safety that related to their total cost; if a drone has a higher price, then the inherent safety systems built in to the UAS are higher than a drone with a lower price tag. In the end, it is the user’s responsibility to prevent any hazardous scenarios. Manufacturers of these devices should consider adding more layers of safety. Overall, the consumer level UASs that are on the market today have some safety layers, but technology could be implemented to improve safety of UAS operation in the NAS.

1.3. Public perception of UASs

UASs are relatively new technology, and the public's perception of this technology continues to be of research interest. Clothier, Greer, Greer, and Mehta (2015) performed study with two parts, the first of which was used to gauge the public's idea of risk level between UAS and manned aircraft through a mixed methods survey. They sampled 200 Australians and found a neutral reaction to the use of a “drone”. Clothier et al. (2015) also found that the participants believed drone technology to be neutrally safe and neutrally threatening. The second part of the study had an experimental design with 510 participants from the same age group. This study determined that people found the perceived risk of UAS was no more or less risky than manned aircraft. It went on to show that the respondents believed that UAS and manned aircraft were beneficial to society. The acceptable risk of manned aircraft was about the same as UAS (Clothier et al., 2015). This indicated that the public has approximately the same risk perception for UASs and manned aircraft.

Humans consistently try to innovate and automate basic tasks. This is incredibly important as the perception that consumers have of UASs could alter how willing they are to fly on a commercial flight. Human perception of autonomy can vary depending on the level of automation and how the automation interacts with a human being. Leung, Paolacci, and Puntoni (2018) found that most people who drove automatic transmission vehicles believed was more efficient and less likely to break down; this preference for automation extended to other products as well. Automation is something that consumers are looking for in products, and therefore, may impact consumers’ view of UASs.

Conversely, Liang and Lee (2017) described how fear of autonomous robots and artificial intelligence (FARAI) was still prevalent within their sample population. They found that individuals did not concern themselves with the difference between a robot and artificial intelligence. Therefore, it did not matter which one was mentioned; a fear for one was a fear for the other. The study concluded that 1 out of 4 people reported having FARAI (Liang & Lee 2017). This was true even though most had not ever interacted with a robot before, enforcing the idea that many people have a negative view of artificial intelligence and robots through preconceived notions (Liang & Lee, 2017). Although humans viewed autonomy in a positive light when used in consumer products (Leung et al., 2018), they also reported a fear of autonomy in a more general sense. The context in which automation is applied may be the deciding factor for how consumers perceive automation. Therefore, it is imperative that we know where UAS fits within these two studies. Leung et al. (2018) depicted positive consumer feedback on several forms of automation, but all examples depicted automation on a very small scale and with no adverse effect on the safety on the consumer. Liang and Lee’s (2017) assessment of human fear of automation was much better suited for human perception of an UAS as it describes a system in the form of a robot or artificial intelligence that is commonly viewed outside of the consumer’s control. It highlights the fact that UASs could be viewed as a threat.

1.4. Influences on consumers

In aviation, factors that affect the decision to use commercial aviation over other means of transportation are of interest, as are all variables that influence consumer choice to utilize commercial aviation. Yimga (2017) conducted research on on-time performance (OTP) and consumer choice. The study showed that customers are aware of OTP as a metric used to decide on the purchase of air fare. They are also willing to pay a premium to get a higher level of OTP as it relates to their flights because this represents quality (Yimga, 2017). OTP is one of the several components that make airline travel attractive to consumers over other means of transportation. Consumers find it effective and even comforting in some cases. However, if a UAS event occurs near an airport, one of the immediate impacts would be interruption of commercial flights and delays; this indicates that UAS events may negatively influence willingness to fly.

News can shape public perceptions and impact how people understand a situation and answer questions. Kapuściński and Richards (2016) tested how news framing affects the decision to go to a vacation destination. The study used various situations paired with a news article about a destination. Respondents rated their perceived risk of going on the vacation; the results indicated that the news media presented affected the decision to go on the vacation and the risk rating (Kapuściński & Richards, 2016). The way the news was presented influenced the risk level. Article A in the study was written with a higher threat level and had a higher perceived risk. Article B was written with less threatening language but with the same overall message; however, it had lower perceived risk. This data supports the idea that news impacts risk perception, and the way news is written and presented can influence risk perception as well.

1.5. Willingness to Fly

Previous work on willingness to fly provides a framework for this concept. Bergstrom and McCaul (2004) conducted a survey with 115 undergraduates at North Dakota State University. They used a scale from 1 to 100 where 1 was least willing and 100 is most willing to fly. The survey was administered 34 days after the terrorist attacks on 9/11. The results showed that worry affected willingness to fly. It also implied that concerns and risk perceptions of the consumer, in this case about the threat of another terrorist attack, would affect their willingness to fly. To better access what else affects consumers, the willingness to fly scale was developed by Rice et al. (2015). This self-report scale of seven questions uses words like comfort, fear, and confidence in rating consumer willingness to fly (Rice et al., 2015).

Rice and Winter (2015) gauged consumer willingness to fly based on pilot configuration and determined which emotions controlled the decision. The study used two scenarios: one where the flight had two human pilots at the controls, and one which was controlled by autopilot. Respondents were asked to answer the same willingness to fly scale (Rice et al. 2015). Like Bergstrom et al. (2004), Rice and Winter (2015) confirmed that feelings played a part in willingness to fly. In addition to worry, the study determined that anger, fear, and happiness also impacted willingness to fly. Thus, we may expect to see a decrease in willingness to fly with fear, concern, or other negative emotions.

There are studies available on consumer willingness to fly in various aspects of commercial aviation. The pilot configuration study (Rice & Winter,2015) indicated that when an aspect of the flight is changed, in this case pilot configuration between autopilot and a human pilot, there is a change in willingness to fly. The research involving 9/11 and willingness to fly (Bergstrom et al., 2004) also indicated that real world events can have some influence on consumer answers after the event has occurred. There has not been research on how a real-world event involving UAS specifically would impact willingness to fly. This study aims to bridge the gap between UAS risk perceptions, real world events, and consumer willingness to fly.

2. Methodology

A single group pre-test post-test design was administered via a survey. This research addressed the question, how will awareness of a UAS event change a consumer’s willingness to fly on a commercial flight?

The target population was consumers of commercial aviation, and the accessible population was undergraduate students at the Florida Institute of Technology. The study collected responses from students, who were at least 18 years old and enrolled as an undergraduate.

We intentionally recruited from sections of classes that cover a wide range of academic disciplines and undergraduate standings (i.e. freshman, sophomore, junior, senior). This provided a sample including multiple disciplines and the entire range of undergraduate standings within the available population. Recruitment emails were sent to professors in Aeronautics, Communications, and Business courses, asking the professors were to post the survey in the online learning management system, as well as emailing the students the recruitment materials.

In keeping with procedures to protect participants and ethical research, we applied for and were granted an Institutional Review Board exemption (19-024), and no personal information was collected. The survey was divided into three sections: 1) instructions for taking the survey and an acknowledgement of consent, 2) pre-test asking that the participants rate their willingness to fly (Rice et al. 2015) on a standard commercial flight, and 3) reading an article on a recent UAS event, and the post-test to rate their willingness to fly on a standard commercial flight again. The UAS article was a recent BBC news article published on the air traffic disruption at Newark airport in New Jersey because of two UASs flying near the airport in January (Lee, 2019). Google forms was used to collect data.

The responses collected in the Google form were exported to Excel for descriptive analysis. The Likert scale responses were coded from -2 to 2 where “strongly agree” was equal to 2, “agree” was equal to 1, “neither disagree nor agree” was equal to 0, “disagree” was equal to -1, and “strongly disagree” was equal to -2. R Studio (v1.1.463) was used to conduct a paired *t*-test, Cronbach’s alpha to test for internal consistency, and Cohen’s d to measure the effect size.

3. Results

Of the 960 students enrolled in the sections recruited, 61 responses were collected for a response rate of 15.73%. The survey also indicated that 58 respondents (95%) had been on a commercial flight: fourteen respondents (23%) fly on commercial aviation one to two times a year, 22 (36%) fly three to four times a year, and 21 (34%) fly more than five times a year. Nine respondents rated their familiarity with unmanned aircraft systems as “very familiar”, and eight respondents (13%) indicated their knowledge of UAS was “not at all familiar”. Of the respondents, fifty-one (83%) said that they owned a UAS, while ten (17%) indicated they did not.

Responses to both the pre-test and post-test yielded a Cronbach’s alpha of 0.97, indicating a very high internal consistency. Therefore, for all further statistics, the average of the willingness to fly scale responses were used.

Table 1 shows the descriptive statistics for the average willingness to fly before and after a participant was made aware of the UAS event. Willingness to fly averaged 1.34 (agree) on the pre-test, and after being made aware of a UAS event, the average was 0.42 (neutral). Figure 1 illustrates the average willingness to fly before and after the intervention.

The paired *t*-test was significant ($t(60) = 6.00$, $p < 0.0001$). This indicated a significant difference in average willingness to fly between the pre-test and post-test. Cohen’s d was 0.88, which is a large effect size.

Table 1: Descriptive Statistics for Willingness to Fly before and after UAS news treatment.

	Pre	Post	Overall
Mean	1.34	0.42	0.88
Median	1.86	0.29	1
Mode	2	2	2
Range	-1.57 to 2	-2 to 2	-1.57 to 2
Standard Deviation	0.89	1.18	0.86

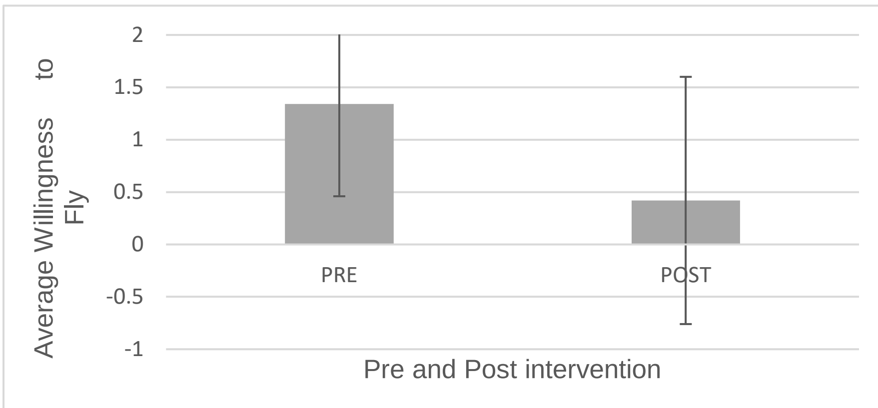


Figure 1: Average willingness to fly before and after UAS event news intervention.

4. Discussion

This study examined consumer willingness to fly before and after reading a news article about a UAS event. The response rate of 15.73% is within the expected response rates for email recruitment. The respondents closely matched the target population in that 95% of participants flew commercially, and 70% reported flying three or more times a year.

Our hypothesis stated that consumer willingness to fly would decrease after participants were made aware of a UAS event. The data supported this hypothesis with a significant decrease in willingness to fly after being made aware of a UAS event (Figure 1, Table 1). Initial willingness to fly (1.34) was slightly higher than "Agree" on the Likert scale, and after reading about the UAS event, willingness to fly decreased to slightly above neutral (0.42). Though this does not indicate that the participants were unwilling to fly after being made aware of a UAS event, it does suggest that the original hypothesis was correct. The initial willingness to fly measurement had a standard deviation of 0.89, while there was more variability in the post-test measure (standard deviation of 1.18). It should also be noted that all responses were positive on the pre-test, but the range on the post-test included strongly disagree (-1.57). Cohen's *d* (0.88) indicated a large effect size and suggested that the difference in willingness to fly is substantial. Respondents were also replying with extremely high levels of consistency (Cronbach's alpha= 0.97), supporting the internal reliability of the scale.

Recruitment was targeted to undergraduate students in Aeronautic, Business, and Communication courses to capture a representative sample of the accessible population. Although the sample matched the target population well in terms of frequent commercial travel and diversity, the undergraduate student population is not representative of the age range of typical aviation consumers. Therefore, these findings should be replicated in an older sample of air travelers. The study included one news article; however, it was a current article, selected from a reputable news source. The article was deemed a reasonable representation of recent UAS events by aeronautics faculty members to provide both a measure of face validity and lack of bias in the news presentation.

Building on the results of this study, aviation consumer travel choices should be examined to determine how negative UAS event press could negatively impact the aviation industry. Future studies should include other demographic variables and test for relationships to willingness to fly: age, fixed-wing piloting experience, UAS piloting experience, frequency of flight. Prior UAS knowledge could also be analyzed to gauge how this may impact the decision to fly commercially after a UAS event. Future work should also examine several, different samples of UAS news coverage to compare the effects on consumer willingness to fly.

This study found a significant decrease in willingness to fly after consumers read a news article about a UAS event. The data suggested that awareness of UAS interfering with airport operations affects consumer willingness to fly. This means that news of UAS events has the potential to impact the aviation industry. As UAS commercial operations continue to increase as they have within recent decades, the chances for an aviation consumer to see a news article similar to the one presented in this study becomes more likely. A lower willingness to fly among consumers can have detrimental effects on aviation businesses. This in turn increases the need for UAS regulation and safety measures, particularly around airports.

5. References

Bergstrom, R. L., & McCaul, K. D. (2004). Perceived risk and worry: The effects of 9/11 on willingness to fly. Journal of Applied Social Psychology, 34(9), 1846-1856.

Brooker, P. (2013). Introducing unmanned aircraft systems into a high reliability ATC system. The Journal of Navigation, 66(5), 719-735.

Clothier, R.A., Greer, D.A., Greer, D.G., & Mehta, A.M. (2015). Risk Perception and the Public Acceptance of Drones. Risk Analysis: An International Journal, 35(6):1167-1183.

Comstock, J. R., McAdaragh, R., Ghatas, R. W., Burdette, D. W., & Trujillo, A.C. (2014). UAS in the NAS: Survey Responses by ATC, Manned Aircraft Pilots, and UAS Pilots. NASA Langley Research Center. NASA/TM–2014-218250. Retrieved from https://ntrs.nasa.gov/search.jsp?R=20140004575.

Federal Aviation Administration. (2018, July). Integration of Civil Unmanned Aircraft Systems (UAS) in the National Airspace System (NAS) Roadmap. Second Edition.

Kapuściński, G., & Richards, B. (2016). News framing effects on destination risk perception. Tourism Management, 57, 234-244.

Lee, D. (2019, January 23). Drone sighting disrupts major US airport. BBC News Technology. Retrieved from https://www.bbc.com/news/technology-46968419.

Leung, E., Paolacci, G., & Puntoni, S. (2018). Man Versus Machine: Resisting Automation in Identity-Based Consumer Behavior. Journal of Marketing Research, 55(6), 818–831.

Liang, Y., & Lee, A. A. (2017). Fear of Autonomous Robots and Artificial Intelligence: Evidence from National Representative Data with Probability Sampling. International Journal of Social Robotics, 9(3), 379-384.

Plioutsias, A., Karanikas, N., & Chatzimihailidou, M. M. (2018). Hazard Analysis and Safety Requirements for Small Drone Operations: To What Extent Do Popular Drones Embed Safety? Journal of Risk Analysis, 38: 562-584.

PytlikZillig, L. M., Duncan, B.,Elbaum, S., & Detweiler, C. (2018, March). A Drone by Any Other Name: Purposes, End-User Trustworthiness, and Framing, but Not Terminology, Affect Public Support for Drones. IEEE Technology and Society Magazine, 37(1), 80-91.

Rice, S., & Winter, S. R. (2015). Which passenger emotions mediate the relationship between type of pilot configuration and willingness to fly in commercial aviation? Aviation Psychology and Applied Human Factors, 5(2), 83-92.

Rice, S., Mehta, R., Dunbar, V., Oyman, K., Ghosal, S., Oni, M. D., & Oni, M. A. (2015). A valid and reliable scale for consumer willingness to fly. In Proceedings of the 2015 Aviation, Aeronautics, and Aerospace International Research Conference.

Yang, L., Tjiptono, F., & Poon, W. C. (2018). Will you fly with this airline in the future? An empirical study of airline avoidance after accidents. Journal of Travel & Tourism Marketing, 35(9), 1145-1159.

Yimga, J. (2017). Airline on-time performance and its effects on consumer choice behavior. Research in Transportation Economics, 66, 12-25.

Zwickle, A., Farber, H.B., Hamm, J.A. (2019). Comparing public concern and support for drone regulation to the current legal framework. Behavioral Sciences & the Law. 37, 109–124.

USING TOLERANCE INTERVAL METHOD AS AN ALTERNATE APPROACH FOR MONITORING PROCESS PERFORMANCE (PROCESS CAPABILITY) OF SURFACE ROUGHNESS OF GEAR TOOTH FLANKS TO AVOID GRINDING BURNS

Akaff S. Diam[1]
Julie Z. Zhang, Ph.D., Professor[1]
Ali E. Kashef, Ph.D., Professor[1]
[1]University of Northern Iowa
akaffd@gmail.com; julie.zhang@uni.edu; kashef@uni.edu

Abstract

The process capability measures most commonly referred to as process capability indices (PCIs), C_p, C_{pk}, C_{pm} evaluate process yield, process consistency and statistical variation compared to the specification. This paper explores an alternate approach for monitoring the process performance of tooth flank surface roughness in gear tooth grinding to avoid grinding burn. The gear tooth flanks are susceptible to microstructural damage due to high thermal energy generated during the process of grind operation. The thermal damage in ground gears can cause catastrophic failure in gears' intended life. To avoid grinding burns the process parameters (feed, speed and depth of cut) are selected such that thermal damage should not occur, however, setting up such parameters could affect the process capability of surface roughness (finish) at a desired level. The surface roughness is commonly specified in design as unilateral tolerance and considered as a critical or key characteristic, therefore, most of the automotive or component manufacturing industry requires it to exhibit a process capability index (C_{pk}) of 1.33 or above. Achieving and maintaining process capability of surface roughness is difficult due to the nature of process, moreover, the thermal damage of the ground feature can impair the product life. In this paper, for gears that are made with ferrous materials (plain carbon steels or alloy steels), tolerance interval method is proposed for monitoring the process performance of roughness in gear tooth grinding process as compared to the conventional process capability.

1. Introduction

All processes show variations and the variations in manufacturing processes are due to common causes and special causes. Process capability was introduced in industry to monitor if the common or chance causes of variations as compared to the specification are at a certain level of acceptance. Process capability is evaluated when the process is in statistical control. A process that is operating with only chance causes of variation present is said to be in statistical control. (Montgomery, 1997, p. 130). Several process capability indices (PCIs), C_p, C_{pk}, C_{pm} are available to quantify if the process meets the requirements set by the designers or customers. Mostly C_{pk} index is used which is defined as the minimum of CPU or CPL, where CPU and CPL referred to as upper capability index and lower capability index (AIAG, 2009). Other definition is $C_p = (U - L) / (6\ \sigma)$ and $C_{pK} = \min \{(U - \mu)/(3\sigma), (\mu - L)/(3\sigma)\}$, where U and L are upper and lower specifications respectivley (Samual &

Norman, 2002). For unilateral tolerances only upper capability index or lower capability index is used whichever is applicable. If the design print calls out maximum tolerance only, CPU is measured. It considers process average and evaluates the process spread with respect to where the process is actually located. The magnitude of C_{pk} relative to C_p is a direct measurement of how away it is from the target.

The assumption for capability index evaluation is that process is approximately normally distributed. If the process variation is centered or targeted between its specification limits, the calculated value for C_{pk} is equal to the calculated value for C_p. But with the process variation away from the specification center, the C_{pk} index degrades. Generally, a $C_{pk} \geq 1.33$ indicates that a process is capable. Values less than 1.33 tell that the variation is either too wide compared to the specification or that the location of the variation is off from the center of the specification. It could be the combination of both spread and location measures that how far the process mean is from the nearer specification limit in terms of 3σ distances. C_{pk} works well only for the bell-shaped "normal" (Gaussian) distribution. For others it is an approximation. McCormick (2007) mentioned that knowledge of the mean and variance allows one to calculate with certainty where any value or range of values lies within the distribution. This is not necessarily the case for other distributions. The shapes of non-Normal distributions are usually affected by moments not expressible in terms of the mean and variance alone.

Generally, gears are made of plain carbon steels or alloy steels of various chemical compositions depending upon the application. The choice depends upon number of factors including size, service and design (Dudley, 1984, pp.4.6). There is a wide range of international standards available in industry for ferrous materials used for gear manufacturing. The gears included in this study are subjected to carburizing, quenching and tempering processes to achieve the required surface hardness and relatively softer core hardness for desired mechanical properties during the application and service. The microstructure of heat-treated gears is therefore very important for desired properties.

In gear dimensioning schemes the surface roughness is normally specified as maximum tolerance instead of bilateral tolerance. The ground gears have specification between 0.8 Ra (max) to 1.5 Ra (max) depending upon the type of application of gears. Surface roughness of many precision gears are held at 0.8 Ra max (Dudley, 1984). Since the gears in their application are meshed with the mating gears and rub constantly, so the tooth flank roughness (finish) really matters along with other gears geometrical features and thus regarded as a critical or key characteristic. Gear load capacity is also affected by the gear grades and the surface roughness. The experimental investigations and service experience indicate that a relationship exists between grades of surface texture and aspects of gear load capacity (ISO, 2006).

2. Gear Tooth Flanks Grinding and Grinding Burns

The grinding is a process that shapes the surface by passes with a rotating abrasive wheel (Dudley, 1984, pp. 5.2). Gear grinding is a complex machining process and is affected by several factors. Wang, Wang and Zhou (2005) reported that grinding is most used process for obtaining high level of surface quality, it remains as one of the most difficult and least understood processes. Apart from achieving surface roughness, the contact between the grinding wheel grain and workpiece material generate a thermo-mechanical load on the workpiece. The most significant disadvantage of increasing the circumferential speed is the rising thermal load acting on the surface layer of the workpiece (Fritz, Sebastian, Patrick, 2016). Some of the important factors for machining process are infeed, wheel speed, depth of cut, types of grinding wheel used and the dressing frequency. The gear grind operation is usually conducted after the heat treatment of semi-finished gears to obtain required sizes and tolerances. A grind stock is left in pre-heat treatment operations of gears' tooth flanks, so that finished operation makes it to required geometrical sizes after grinding. Grinding process generates high thermal energy and if the parameters are not set correctly, could damage the microstructure of tooth flanks. The inspection of thermal damage is done with an acid etch process. Grinder burn is a broad term and encompasses varying degrees of thermal damages. To visually detect, any burn must exhibit enough contrast between grey and dark grey (Crow & Pershing, 2018). Gears having thermal damages could adversely affect the product life and cause premature failure. Grinding speed, and depth of cut could create thermal damage. According to Augier, Cruz, Paula and Bianchi (2008), the measurements of surface roughness show an increase in magnitude as the depth of cut was increased, mainly after the test with 35µm depth of cut where grinding burn on the workpiece surface took place. Controlling the grinding parameters is a key to avoid grinding burns, which however could impact surface roughness process mean that could potentially degrade the process capability (≥1.33) of surface roughness. Some of the factors causing grinding burns are:

- Grind stock amount on the tooth flanks – overstock / understock
- Pitch diameter runout in the semi-finished gears
- Gear geometric features' (gear lead, gear profile and tooth thickness) variations
- Heat treatment distortion
- Material variation
- Oil nozzle placement
- Machine set up

The grinding process must be optimized to a level to avoid thermal damages. This situation, however, causes the roughness measurements skewed and most of the data concentrate close to the upper / maximum tolerance level, creating a non-centered process for surface roughness values. The underlying assumption for process capability analysis is that data should exhibit a normal distribution beside being statistically predictable and under control (AIAG, 2005). Hence, achieving capability of e.g. C_{pk} of 1.33 becomes difficult. Moreover, in batch processes, batch to batch variation is inherent and mostly dependent on how well the process is controlled in heat treatment and pre-heat treatment manufacturing processes of steel grades used . Post-heat treatment processes are greatly affected by the consistency of carburized area of the gear tooth flanks. During heat treatment process various factors affect the gears distortion. According to Davis (2005), in gears two types of distortion occurs, one is body distortion, which includes run-out, out-of-roundness, out-of-flatness. The second is the distortion in gear tooth geometry. Body distortion also influences the tooth geometry distortion a great deal. The grinding stock is thus ensured and is determined on the basis of cleaning of all the surfaces of teeth considering distortion and growth of gears after carburizing and hardening (Davis, 2005). Also, when tool wear and adjustments are frequently made, typically maintaining process mean for surface roughness at certain level is difficult. So even if a batch displayed surface roughness process capability of 1.33 or above, there will be no surety that process will always be running at or above 1.33 C_{pk} index in the next batch. Achieving and monitoring capability index of 1.33 and above is therefore has ever been challenging for surface roughness in gear grinding while ensuring no grinder burn should happen.

3. Statistical Tolerance Interval, an Alternate Approach

A statistical tolerance interval is an interval that one can claim to contain at least a specified proportion, *β (also P is used in some texts)*, of the distribution with a specified degree of confidence, 100 (1-α) %. Such an interval would be of particular interest in setting limits on the process capability of a product manufactured in large quantities (Hahn, Meeker & Escobar, 2017). According to the international standard a statistical tolerance interval is an estimate interval, based on a sample, which can be asserted with confidence level (1- α), for example 0.95, to contain at least specified proportion *p* of the items in the populations (ISO, 2014). The international standard discusses two methods for determining the statistical tolerance intervals, a parametric method for the case where the characteristic being studied has normal distribution and a distribution-free method for the case where nothing is known about the distribution except that it is continuous. Before using tolerance interval that depends heavily on the normality assumption, one should assess the adequacy of the normal distribution as a model, (Hahn & Meeker, 1991). As reported by Hahn and Meeker, two-sided distribution-free statistical intervals from a random sample from a specified population one generally proceeds as follows:

- Specify the desired confidence level for the interval
- Determine (from tabulations and calculations) the order statistics (ordered observations from smallest to largest are called ordered statistics) that provide the statistical interval with at least the desired confidence level for the sample size
- Use the selected order statistics as the end points of the distribution-free interval

One sided distribution -free bounds are obtained in a similar manner, except that only the order statistic is used as the desired lower or upper bound (Hahn & Meeker, 1991). Tolerance interval method is useful in one sided tolerance (upper bound or lower bound) where process centering (targeting) is an issue and is less important than feature's conformance. Also, non-parametric (distribution-free) tolerance intervals methods can be used if the measured data do not follow normal distribution to assess the conforming proportions.

4. Statistical Tolerance Interval Procedures

There are several procedures available in text to construct the statistical tolerance intervals depending upon the requirements. This paper will discuss following statistical tolerance interval procedures.

4.1 Upper-bound tolerance interval when the population is normal distribution (with known mean and known variance)

The international standard (ISO, 2014) provides procedure when the values of the mean, μ, and the variance, σ2, of normally distributed population are known, the distribution of the characteristic under investigation is fully determined. There is exactly a proportion P of the population to the left of of the one-sided interval equation below;

$$X\text{U} = \mu + \mu_{p} \text{ X } \sigma \qquad [1]$$

In this equation, μp, is the P-fractile of the standardized normal distribution, XU is the upper limit of statistical tolerance interval, σ is the population standard deviation and μ is the population mean.

4.2 Upper-bound tolerance interval when the population is normal distribution (with unknown mean and known variance)

If we have a normal population with unknown mean and known variance, find *k* such that $\bar{x}+k\ \sigma$ satisfies that *at least* a proportion *P* of the population is below $\bar{x}+k\ \sigma$. Note that $\mu + \mu_p \sigma$ is the population tolerance limit in the sense that exactly a proportion *P* of the population is below that limit. So if $\bar{x}+k\ \sigma \geq \mu+\mu_p \sigma$, where $\bar{x}$ is the sample mean and k is the factor used to determine the tolerance limits then the proportion of the population that is smaller than $\bar{x}+k\ \sigma$ is at least *P*. Thus the probability (confidence level) that a proportion of the population is at least P is (1-α), if

$$P\,[\bar{x}+k\,\sigma \geq \mu + \mu_p\,\sigma] = 1-\alpha \qquad [2]$$

The probability on the left side of [1] can be rewritten:

$$P\,[\bar{x}+k\,\sigma \geq \mu + \mu_p\,\sigma] = P\,[(\sqrt{n}(\bar{x} - \mu)/\sigma \geq \sqrt{n}\mu_p - \sqrt{n}k] = 1-\alpha \qquad [3]$$

The variable $\sqrt{n}(\bar{x} - \mu)\,/\sigma$ in formula [3] has a standard normal distribution, and follows from the last equality in [2] that:

$$\sqrt{n}\mu_p - \sqrt{n}k = \mu_\alpha,$$

Which can be rewritten as:

$$k = \frac{1}{\sqrt{n}}\,\mu_{1-\alpha} + \mu_p \qquad [4]$$

The value of k is based on the confidence level (1- α), minimum proportion of the population asserted to be lying in the statistical tolerance P and number of observations in the samples n.

4.3 The distribution-free statistical tolerance interval for any type of distribution

For establishing distribution-free tolerance intervals the order statistics (of sample is determined by solving binomial distribution function for smallest sample size n. International standard (ISO, 2014) describes the procedure for constructing distribution-free or for any type of distribution statistical tolerance bounds. Assume we have samples, $x_1, x_2, x_3,\ldots\ldots,x_n$, of independent random observations on a population (continuous, discrete, or mixed) and let its order statistics ((ordered observations from smallest to largest are called ordered statistics) is $x_{(1)} \leq x_{(2)} \leq x_{(3)} \leq x_{(4)} \leq \ldots\ldots\ldots \leq x_{(n)}$. The interval with 100(1-α) % confidence that at least 100 P % of the population lie between the v^{th} smallest observation and w^{th} largest observation is:

$$\sum_{x=0}^{v+w-1} \binom{n}{x} p^{n-k}(1-p)^{x} \leq \alpha, \qquad [5]$$

where v ≥ 0, w ≥ 0, v + w ≥ 1, 0 < α < 1.

Hanson and Owen discussed that continuity requirement on cumulative distribution function (c.d.f.)'s is unnecessary. They noted that many (c.d.f's) which occur in practice are not continuous, and in many cases where distribution-free tolerance limits are applicable they are not being used because of an uncertainty as to whether the underlying distribution is or is not continuous or because of the certainty that it is not (Hanson & Owen, 1963). So, when the (c.d.f.) of the population characteristic X is not continuous, the statement will be slightly modified such that there is at least 100(1-α) % confidence that at least 100 P % of the population is between $x(_v)$ and $x(_{n-w+1})$ or equal to $x(_v)$ and $x(_{n-w+1})$. When v + w = 1, formula [4] reduces to:

$$P^n \leq \alpha, \text{ or}$$

$$n = \log(1-\alpha) / \log(P) \qquad [6]$$

The equation [6] shows the minimum sample size required for certain confidence level and proportion to be in conformance. Thus, to be 95% confident that at least 95% percent population lies below the largest value of the sample, the sample size must be n = log (1-0.95)/log (0.95) = (-1.3) / (-0.022) = 59. There are several tables available which gives minimum samples required for constructing statistical interval at certain proportion p and certain confidence level (1-α)%. Hahn and Meeker have compiled and listed tables based on the selection of sample size n, confidence levels 100(1-α) % and at various levels of conformance proportion percentage P (Hahn & Meeker, 1991). Also, now-a-days with the help of latest statistical analysis softwares like Minitab® it is easily possible to calculate, and construct distribution based or distribution-free statistical intervals.

5. Data Analysis and Interpretation

In this paper, a real-life sample data of 7 different gears (total 2039) ran on the same gear grinder is used and analyzed to calculate one sided (upper bound) tolerance interval. All parts are machined from alloy steel forgings, for example grades 8620, 4120, 20MnCr5 (mainly carbon contents between 0.18%-0.2% along with varying composition of Ni, Cr, Mn, Si, P, S and Mo) and have print tolerance of roughness 1.0 RA max., however the sizes of parts are different in gear geometry, module size and tooth flank face width. Table 1 shows the quantity of each part and the minimum and maximum values recorded during 3 shift operations. Same instrument, Mitutoyo Surface Tester, used for measuring the roughness values for all parts.

Table 1 below shows none of the measured value exceeded the specification limit of 1.0 max, however, the distribution fit for normality failed for 5 out of 7 parts.

Table 1. Statistical Summary of Roughness (Ra) by Part Numbers

Part	Pieces Checked	Roughness Min (Ra)	Roughness Max (Ra)	Roughness Mean *(Ra)*	Std. Dev of Roughness *(Ra)*	Distribution Fit
A	210	0.386	0.977	0.703	0.1251	Normal
B	219	0.451	0.986	0.732	0.1147	Non-Normal
C	381	0.423	0.984	0.714	0.1145	Normal
D	175	0.442	0.998	0.765	0.1291	Non-Normal
E	581	0.410	0.998	0.780	0.1272	Non-Normal
F	306	0.425	0.997	0.778	0.1068	Non-Normal
G	167	0.334	0.998	0.722	0.1592	Non-Normal

There are several reasons that the data failed to be normally distributed. Some of the reasons are as follows:

- Shift to shift variations
- Batch to batch variations
- Operator to operator variations

Running analyses at 95% percent confidence level (1-α) for all seven different parts included in the study for estimating that 98% of population (P %) being in tolerance bound. Results are summarized in Table 2, since part A and part C follow the normal distribution the upper tolerance bound for part A and C is 0.988 and 0.968 respectively. This indicates that at 95% confidence, 98% of the population will have values equal to or less than 0.988 Ra and 0.968 Ra for part A and part C respectively, which is less than 1.0 max Ra as specified in the print.

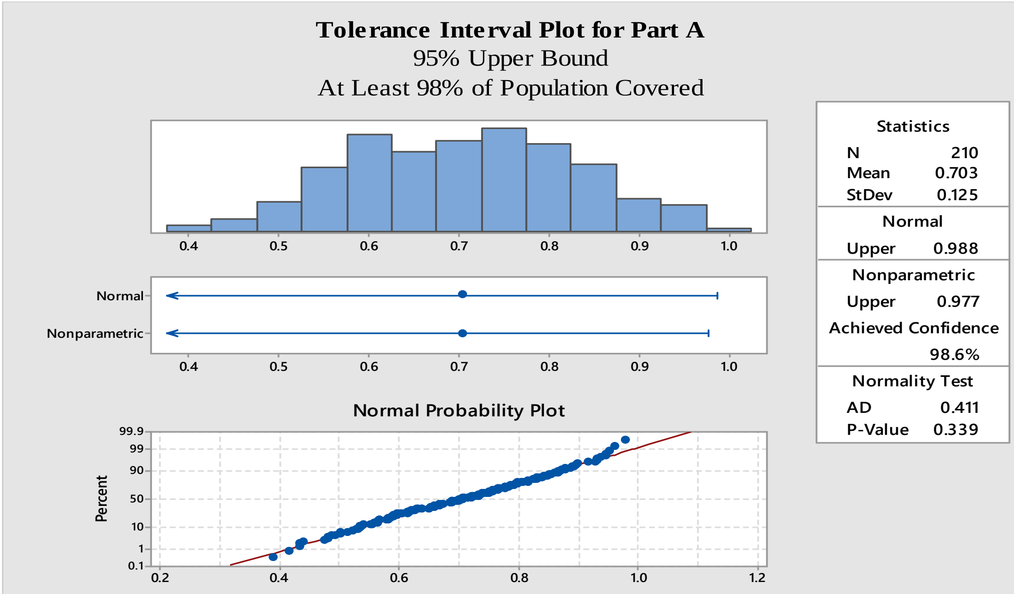


Figure 1. Tolerance Interval Plot for Part A Using Minitab®

Table 2. ***Summary of Calculated Tolerance Intervals of parts***

95% Upper Tolerance Bound			
Part	**Normal Method**	**Nonparametric Method**	**Achieved Confidence**
A	0.988	0.977	98.60%
B	0.992	0.986	98.80%
C	0.968	0.981	98.20%
D	1.062	0.998	97.10%
E	1.057	0.991	97.50%
F	1.017	0.993	98.50%
G	1.089	0.998	96.60%

Achieved confidence level applies only to nonparametric method.

Table 1 shows that the sample data for parts B, D, E, F and G do not follow the normal distribution, therefore, nonparametric method is used, the results indicated in the column "Achieved Confidence" in Table 2 are all greater than 95%.

Now we can analyze the process capability of the same sample and compare the results achieved in tolerance interval methods. Part A and C capability indices with upper and lower confidence interval bounds on the capability indices are given in the Table 3. The sigma used for calculating the process indices is the overall process sigma for each part. Table 3 showed that the capability indices for part A and C are less than 1.0, which means the process is not capable.

Table 3. ***Process Capability of Normally Distributed Parts***

		Confidence Interval on C_{PK}		Projected Performance
Part	C_{PK} (C_{PU})	**Lower CI**	**Upper CI**	**PPM (percent)**
A	0.79	0.70	0.88	8,933 (0.89%)
C	0.83	0.76	0.90	6,360 (0.63%)

To improve the process capability, process variation must be minimized. Figure 2 shows that reason for the part A process being not capable is that the mean is not centered.

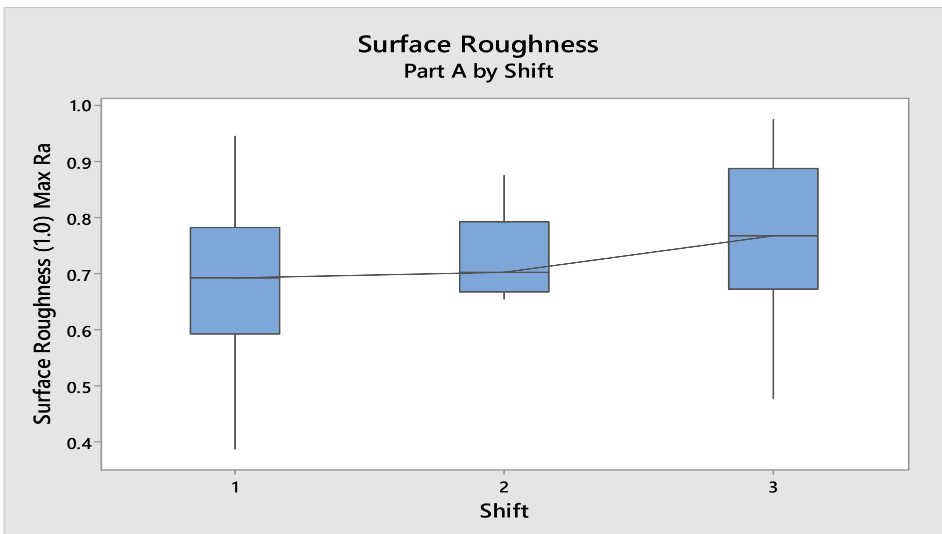


Figure 2

There is difference in means for each operator /shift and for each batch. Figure 2 shows the shift to shift difference of data in box plot for Part A. Actions must be taken to control the machine performance, consistency of input material, change the basic method by which the process operates or shift the process average closer to target. Frequent process adjustments, sorting and scrap costs will have to be increased to bring the process to an acceptable level of ≥ 1.33. Comparing with the process capability evaluation, tolerance interval method is more practical and economical to monitor the performance of process without sacrificing much for cost and it is providing greater than 95% confidence that 98.6% and 98.2% populations of these parts A and C respectively are conforming to the specification.

6. Conclusion and Further Recommendations

The above results for individual gears made of alloy steels suggest that estimating the tolerance interval has an advantage and provide a confidence that more than 98% population is within tolerance specification. Adjusting unreasonably the process to control the roughness measure and process mean could create issues of grinder burn and could increase the cost of process, cycle time and scrap cost. Companies can use this alternate approach of tolerance interval method where the process mean is not centered, and conventional process capability achievement is difficult. Parts are produced in short batches, parameters for grinding are set by individual operators, and the distribution of roughness measure is concentrated mostly between 0.65> Ra >0. 87. Selecting the gear grinding parameters like feed, speed and depth of cut to bring the mean of surface roughness measure close to the nominal or target will increase frequency of grinding parameter adjustments and hence will also increase the part processing cycle time and the risk of tooth flank grinding burn. Companies can formulate data collection schemes to collect surface roughness measurement data and keep monitoring and recording the process at the start and at the end of shift or with certain cadence between shift. Enough data can be collected over a period, it will become easier and economical way of monitoring the process performance of surface roughness in terms of conformance with a required degree of confidence.

For future studies, processes like reaming, hardness check (heat treatment) or broached diameters can be studied and evaluated if the process capability is mandated. These processes are also targeted where conventional process capability analysis could fail, for example, hardness is specified generally as unilateral tolerance (HRC 58Min) and making parts harder would affect finishing processes in terms of increased tooling costs. Similarly, broaching tools and reaming tools are always selected so that maximum tool life could be utilized, hence the process targeting is non-centered compared to the nominal when the tools are new in the early production of parts with these tools.

7. References

AIAG, 2005. Statistical Process Control (SPC), Second Ed. Automotive Industry Action Group, Detroit, MI.

Augier, P. R., Cruz, E. D. C., Paula, W. C. F., & Bianchi, C. E. (2008). Predicting Surface Roughness in Grinding using Neural Networks Advances in Robotics. Automation and Control. CC BY-NC-SA, 3.

Crow Jonathan R. & Pershing Michael A., Standard Samples for Grinder Burn Etch Testing, Gear Technology Magazine, June 2018, 54-55, Retrieved from https://www.geartechnology.com/articles/0618/Standard_Samples_for_Grinder_Burn_Etch_Testing/

Davis, J. R. (2005). Gear materials, properties, and manufacture. Retrieved from https://ebookcentral-proquest-com.proxy.lib.uni.edu

Dudley, Darle W. (1984). Handbook of Practical Gear Design (Rev. 1st Ed.). New York: McGraw-Hill.

D. W. McCormack, Jr., Ian R. Harris, Arnon M. Hurwitz & Patrick D. Spagon (2000) Capability indices for non-Normal data, Quality Engineering, 12:4, 489-495, DOI: 10.1080/08982110008962614

Fritz Klocke, Sebastian Barth & Patrick Mattfeld (2016) High Performance Grinding, Procedia CIRP, 4, 266-271, doi:10.1016/j.procir.2016.04.067

Hahn, G. J., Meeker, W. Q., & Escobar, L. A. (2017). Statistical intervals: A guide for practitioners and researchers. Retrieved from https://ebookcentral-proquest-com.proxy.lib.uni.edu

Hahn, G., & Meeker, W. (1991). Statistical intervals: A guide for practitioners (Wiley series in probability and mathematical statistics. Applied probability and statistics). New York: Wiley.

Hanson, D., & Owen, D. (1963). Distribution-Free Tolerance Limits Elimination of the Requirement that Cumulative Distribution Functions be Continuous. Technometrics, 5(4), 518-522.

International Organization for Standardization. (2014). Statistical interpretation of data – part 6: Determination of statistical tolerance intervals, ISO 16269-6:2014. Retrieved from https://asc.ansi.org

International Organization for Standardization. (2006). Technical Report, Cylindrical gears – code of inspection practice, ISO/TR 10064-4/:1998. Retrieved from https://asc.ansi.org

Montgomery, D. C. (1997). Introduction to statistical quality control (3rd Ed.). New York: John Wiley & Sons.

Samuel Kotz & Norman L. Johnson (2002) Process Capability Indices—A Review, 1992–2000, Journal of Quality Technology, 34:1, 2-19, DOI:10.1080/00224065.2002.11980119

Wang, J. Z., Wang, L. S., Li, G. F. & Zhou G. H. (2005). Prediction of surface roughness in cylindrical traverse grinding based on ALS algorithm, Proceedings of the Fourth International Conference on Machine Learning and Cybernetics, 18-21 August 2005, 549-554, Guangzhou.

A STUDY OF THE EFFECT OF VIBRATION ON ACCURACY OF 3D-PRINTED PARTS VIA VAT PHOTOPOLYMERIZATION

Odai M. Nassar[1]
Wilfredo Moscoso-Kingsley[1]
[1]Wichita State University
wilfredo.moscoso@wichita.edu; *omnassar@shockers.wichita.edu*

Abstract

Additive Manufacturing (AM) technology has emerged as a promising alternative for many industries, including aerospace, space, automobile, medical and biomedical. This is due to AM's advantages over traditional manufacturing processes in terms of material savings, the ability to produce customized microstructures and complex geometries. The result has been the manufacture of structures with optimum strength to weight ratios, especially when design for AM is optimized via finite element analysis. However, there are some challenges that impede AM. One challenge is the limited accuracy typically associated with all AM processes. Considerable investigations connect three-dimensional (3D) printing parameters, part scaling, and solid model discretization with AM printing accuracy. However, there appears to be a lack of knowledge regarding the effect of vibrations on the 3D printing accuracy. Additive manufacturing machines could be prone to vibration from adjacent machines in the workshop even when vibration isolation systems are deployed. Moreover, there are instances where intentional application of vibration during 3D printing has been found helpful to develop material structure. Therefore, it is of importance to investigate the effect of vibrations on 3D printing accuracy. In this study, controlled levels of vibrations were applied during the vat photopolymerization printing process, and a high-precision coordinate measuring machine (CMM) was used to correlate induced vibrations with dimensional and geometric accuracy of the printed parts.

1. Introduction

Additive Manufacturing (AM) technology has emerged as a promising alternative for many industries, including aerospace, space, automobile, medical and biomedical. This is due to AM's advantages over traditional manufacturing processes in terms of material savings, the ability to produce customized microstructures and complex geometries. The result has been the manufacture of structures with optimum strength to weight ratios, especially when design for AM is optimized via finite element analysis.

Because of the layer-by-layer build strategy, AM has more flexibility than traditional subtractive processes to produce very complex geometries, and it can produce customized microstructures for certain applications (Kamal & Rizza, 2019). Additive manufacturing also helps in producing parts with optimum volume to weight ratios, which is critical in aerospace and space applications. This is because AM enables fabrication of free forms, which may be analyzed using finite element analysis (FEA) to minimize material needed to accommodate expected stresses. It is difficult and costly to make these optimized free forms in traditional manufacturing (Kamal & Rizza, 2019).

Additive manufacturing is also promising for repair processes, not only for manufacturing as the name implies. The work of Gasser et al. showed such applications as seals repair, using laser metal deposition. It also discussed some highly funded projects in Europe for repairing aero engines using AM technologies. For instance, Gasser et al. discussed the FANTASIA repair project, which was launched in 2016 with a budget of 6.5 million Euros. The project had as objective to decrease both the repair cost and time of these engines by about 40 % (Gasser et al., 2010).

Rapid prototyping is another important advantage of AM, since AM helps print prototypes of the real components with true or scaled dimensions, for performance-related testing. The use of AM for rapid prototyping started in 1990s using the stereolithography apparatus (SLA). Then, other AM technologies like fused deposition modeling (FDM), ink jet printing (IJP), and selective laser sintering (SLS) were introduced to the market for rapid prototyping (Kruth et al., 1998).

From the above, it is obvious that AM is emerging as a viable manufacturing method, and this has encouraged development efforts within the recent years. However, there are some challenges that impede AM. One challenge is the limited accuracy typically associated with all AM processes, including polymer, metallic, composite and ceramic AM processes. Many researchers conducted studies to investigate the accuracy of AM processes. These studies have been complemented with many other efforts to minimize the drawback of limited accuracy typically associated with AM. There are several causes of dimensional and geometric deviations in AM parts. The first cause comes as a result of approximating the CAD file. The CAD file cannot be used by AM machines' software to produce the part, but it is instead approximated by a triangular mesh by converting it to a stereolithographic (STL) file format (Moroni et al., 2014). The second cause is called the "stair-case" effect, which arises as a result of slicing the STL file into layers (Moroni et al., 2014). The third cause is the residual stresses stored in the AM parts, especially metallic parts after they cool down, which leads to warping and distortion of the AM part (Mukherjee et al., 2017). The fourth cause is the shrinkage that takes place after solidification of each melted or cured layer. There are many other reasons.

This study investigates the effect of an external vibration source near the AM machine on the quality of AM parts in terms of dimensional and geometric accuracy. Since in real applications, the machines are arranged next to each other in the machine shop, and running other machines near the AM machine causes vibrations that might propagate through the floor to the AM machine. This is particularly important for applications, usually military in nature, where the AM equipment must work on board of terrestrial, marine and aerospace vehicles. Vibrations are potential disturbances even when using vibration isolation systems, as these systems typically reduce up to 60% of the vibration amplitude. Therefore, it is of importance to investigate the potential effect of vibration on the dimensional and geometric accuracy of AM parts.

Another realistic case where vibration may be present during 3D printing, and potentially detrimental, is in the international space station (ISS). Three-dimensional printing in space is attractive to scientists, since it provides an opportunity to make parts in the orbital environment under circumstances that are not available and not possible on earth. Spacecraft floating in space have low-level acceleration conditions that make microgravity science experiments to be possible not only for additive manufacturing, but also in various disciplines such as materials science, physics, fluid mechanics, and life sciences (Whorton, 2001). Microgravity conditions are beneficial. However, Whorton has questioned the validity of the assumption of microgravity conditions in the ISS, since the station's laboratories are not completely vibration-free. Despite that the ISS provides near-zero acceleration, the environment in the ISS is expected to deviate significantly from the requirements of acceleration-sensitive experiments, and that is because of the various sources of vibro-acoustic disturbances on the ISS (Whorton, 2001). Efforts are being made to reduce the propagated disturbances, but the acceleration levels and thus the vibration levels do not meet the requirements. The work presented herein, and the methodologies applied to study the effects of vibrations on dimensional and geometric accuracy could be modified to suit exploration of the effect of vibrations of the microgravity assumption.

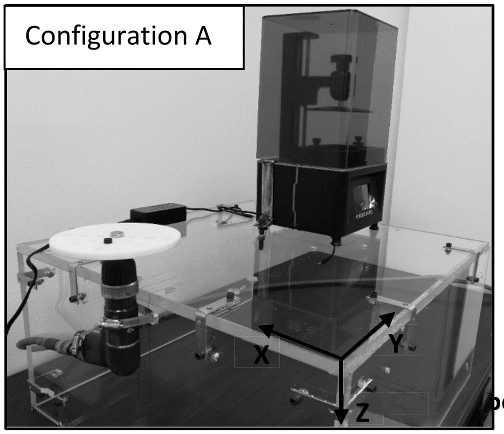
Additionally, vibration is intentionally applied in some 3D printing applications, seeking some favorable effects. For example, some extremely viscous materials (more than 1000 Pa·s) used in electronics and biomedical applications are not possible to be 3D printed using existing methods. A study conducted by Gunduz et al. showed that applying high-amplitudes of ultrasonic vibrations within a nozzle helped in reducing the wall friction and the flow stresses, and facilitated 3D printing of an aluminum-polymer composite, a commercial polymer clay, and a stiffened fondant with extremely high levels of viscosities up to 14000 Pa·s (Gunduz et al., 2018). Another study and registered patent by Koch and Kuklinski aimed at mixing metallic particles with the resin being printed via SLA and then applying acoustic field during the printing process. The study showed that the applied acoustic field prevented the metallic particles from agglomerating and helped in printing reinforced parts (Koch & Kuklinski, 2004). However, these studies have not investigated if these intentionally applied vibrations could influence the dimensional or geometric accuracy of the printed parts. The study presented herein aims at filling this knowledge gap.

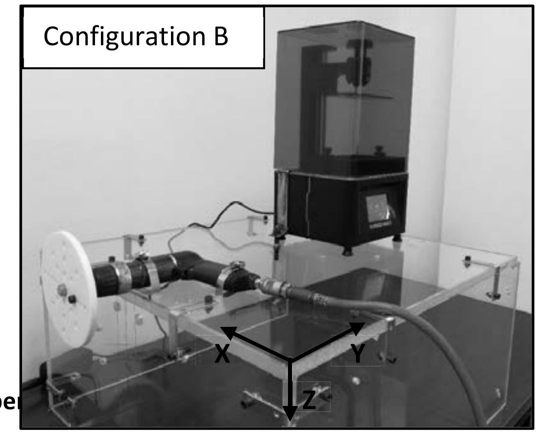
2. Experimental work

An ELEGOO MARS® vat photopolymerization machine (Figure 1) was used for all tests. The machine uses ultraviolet (UV) light to cure the resin layer by layer. The UV light is transmitted by a liquid-crystal display (LCD) screen built under the resin's vat. The vat has a transparent bottom through which the UV light passes in order to cure the layer being printed. The platform travels up by one-layer thickness, by actuation of a screw jack, after the previous layer is cured, in order to print the next layer. The part is printed upside down.

The machine has a special slicing software (ELEGOO ChiTu Box®). The CAD file must be saved as STL file and then opened by the machine's software. The parts and the support structures printed for testing were sliced into 548 layers of thickness of 0.03 mm. The exposure time was 8 seconds for each layer, except for the bottom layers, which had an exposure time of 60 seconds. The support structure was added to facilitate removing the part from the build platform, after finishing the printing process. Each printed part was soaked in isopropyl alcohol of concentration of 99 %, for 5 minutes, and then left at room condition for 18 to 20 hours, in order to allow further shrinkage to take place before measurements were performed to check for dimensional and geometric characteristics, using a high-precision coordinate measuring machine (CMM).

The vibration levels were controlled by clamping an Ingersoll Rand® 6AJST4 pneumatic motor to the table on which the ELEGOO® machine was installed (Figure 1). A predesigned disk with multiple holes was then installed on the pneumatic motor and a bolt with a nut of a total mass of 8 grams were used as an eccentric mass fixed to the rotating disk. The pneumatic motor has a constant rotational speed of 3600 rpm, and thus the magnitude of vibration was controlled by changing the distance from the center of rotation at which the eccentric mass was installed, which causes a change in the centrifugal force exerted on the table by the motor. Two vibration configurations were utilized (Figure 1 left and right). In configuration A, the force vibrated in the XY plane, whereas in configuration B, the force vibrated in the XZ plane. That is, in configuration A, a longitudinal vibration mode, propagating symmetrically along X and Y was produced; and in configuration B, a mixed transverse (oscillating along Z) and longitudinal (along X) vibration mode was produced. The YZ and XZ planes are symmetric, thus only the XZ plane was utilized. The authors opined not to add the two configurations as factors to the experimental design because they are categoric factors while the other two factors (vibration amplitude and wall thickness) are numeric. Categoric factors can be augmented with numeric factors in Design Expert® software that has been used for data analysis in this study, but the software will treat them as two configurations since categoric factors do not have center points (half-way value). The software will indeed test the levels of the numeric factors and will show the 3D plots once at each level of the third factor (the configuration) exactly as has been presented in the section of results in this study.





Knowing the mass, the rotational speed, and the radial location of the eccentric mass, the centrifugal force *F* can be calculated by (Doherty, 1885):

$$F = mv^2/r \quad (1)$$

where, *F* is the centrifugal force in N, *m* is the mass in kg, *v* is the tangential velocity in m/s, and *r* is the radius of the path of the motion of the eccentric mass in m.

The tangential velocity can be calculated by (Myers & Myers, 2006):

$$v = \omega r \quad (2)$$

where ω is the rotational speed in rad/s.

The rotational speed of the pneumatic motor was determined using a capacitance-based displacement probe. A thin metallic shim was attached to the plastic rotary part of the motor and the motor was allowed to run. While the motor was running, the capacitance probe gave a reading every time it detected the metallic shim. The time of several rotations was thus measured. On the average, the rotation frequency of the motor was 60 Hz. As shown in Table 1, three levels of vibration were tested. One was zero (no vibration), and the other two were produced by rotating the eccentric mass at two different radii of rotation, namely, 30 mm and 60 mm. The corresponding levels of centrifugal force (force vibration amplitude) calculated using equations 1 and 2 are shown in Table 1.

Table 1. Levels of centrifugal force (force vibration amplitude) and frequencies

Rotational Speed (rad/s)	mass (g)	Radius (mm)	Centrifugal Force (N)	Frequency (Hz)
377	8	30	34	60
377	8	60	68	60

2.1 3D Printed Parts

ABS-like photopolymer resin was used for all tests. This resin has low shrinkage property and high smoothness after photo-curing. The properties of the used resin are shown in Table 2. A hollow cylinder with a square flange was printed for all tests but with three different values of wall thickness for the hollow cylinder (Figure 2). The cylindrical geometry was selected as it is a common engineering geometry that is highly susceptible to dimensional and geometric deviations. The square flange is necessary to facilitate clamping during the inspection process by the coordinate measuring machine (CMM), and to define and align the X, Y, and Z axes of the part with those of the CMM, and the 3D printing process/vibration directions. An arrow indicating the positive direction of the X axis was printed on all parts, as seen in Figure 2, to maintain consistent orientation during CMM measurement and orientation during 3D printing/application of vibration. That is, CMM data along this printed X was for printing/vibration along the X-axis in Figure 1. The inspection was carried out for both the outer and the inner cylinders. The CMM was a Hexagon® coordinate measuring machine, with a resolution of 2.2 micrometers.

Table 2. Properties of the used resin

Property	Value	Unit
Solidification wavelength	405	nm
Shore durometer hardness	84	-
Viscosity (25°C)	150 – 200	mPa·s
Density (liquid)	1.1	g/cm^3
Density (solid)	1.195	g/cm^3
Flexure strength	59 – 70	MPa
Extension strength	36 - 53	MPa

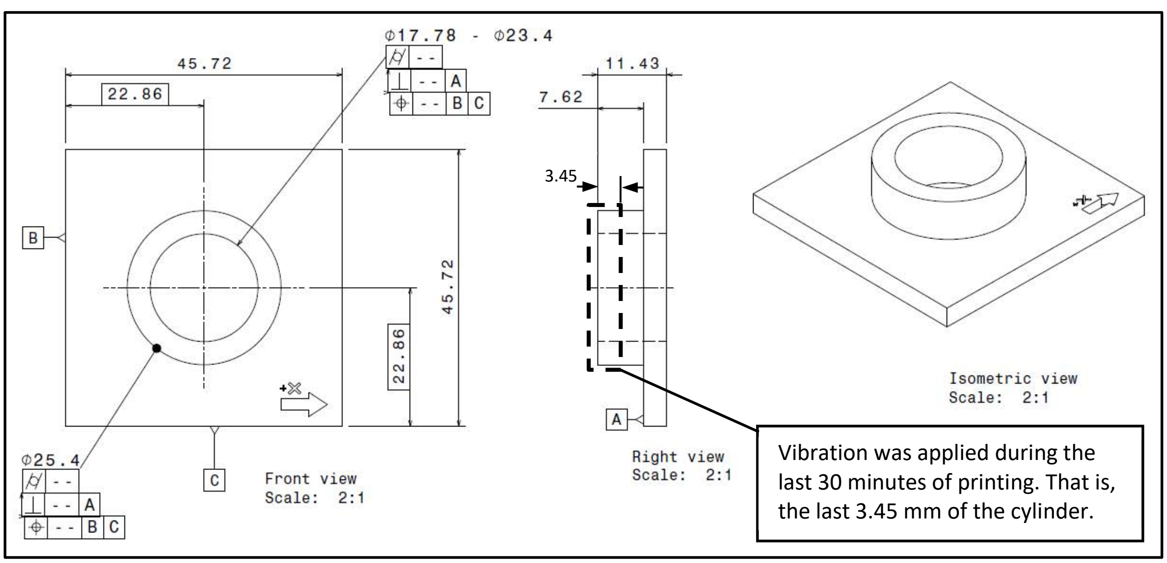


Figure 2. Drawing of the tested part (dimensions are in millimeters)

2.2 Experimental design

Design of experiments (DOE) methodologies were considered to objectively and reliably assess the effect of vibration and wall thickness on the dimensional and geometric accuracy of the 3D printed parts. Thus, the control factors were the force amplitude and the wall thickness of the 3D printed part. In the forthcoming, the force amplitude is factor A and the wall thickness is factor B. The response factors included a set of dimensional and geometric parameters of the 3D printed part. More details about the control and response factors are given below.

The initial design that was selected for this study was a 2^k factorial design with center points. In this design, two levels are tested for each control factor. In the forthcoming, these two levels are called the high level "+1" and the low level "-1". The center points are half-way values between the high and the low levels of each control factor. The center points are given the code "0". The center points are necessary to test for linearity within the tested range (from low to high). After the first round of experiments had been carried out, the curvature term turned out to be significant most of the time, and thus the tested range could not be assumed linear. Due to the significance of the curvature term, the 2^k factorial design had to be augmented with more experiments. Therefore, the final implemented augmented 2^k factorial design followed the response surface method (RSM) design (Antony, 2014).

2.3 Control factors

The control factors in this study are the force amplitude of vibration represented by the shaking force exerted on the work table by the rotating disk with an eccentric mass (Figure 1), and the wall thickness of a hollow cylinder printed by the machine (Figure 2). The tested levels are shown in Table 3.

Table 3. Control factors and their tested levels

Control Factor	**Levels** ("-1", "0", and "1", respectively)	**Unit**
Force amplitude (magnitude of shaking force)	0 / 34 / 68	N
Wall thickness	2.00 / 4.81 / 7.62	mm

2.4 Response factors

The response factors in this study reflect the dimensional and geometric accuracy of the printed part at the planned levels of vibration. They are determined by calculating the deviation of the printed part from the CAD model. The response factors investigated in this study are the perpendicularity of the printed cylinder relative to its base, and the diameter deviation, the cylindricity, and the position of the center of the cylinder. This is for both the outer cylinder and the inner cylinder of the part. The investigated response factors are illustrated by the geometric dimensioning and tolerancing (GD&T) notations in Figure 2. The measurements were performed for the last 30 minutes of printing. That is for the cylindrical band marked in Figure 2 at the top of the cylinder, of length = 3.45 mm.

3. Results

Analysis of variance (ANOVA) was performed in this study in order to test for statistical significance of the effect of each control factor and factors interactions on each response factor. If the p-value of a factor or factors interaction is less than 5%, the factor or the factors interaction is considered to have a statistically significant effect on the response. However, the ANOVA of one response factor is shown here as an example (table 4). In order to have a valid analysis of variance, the validity of the ANOVA assumptions like the normality of the experimental error and the constant variance must be checked (Mendenhall & Sincich, 2015). These diagnostics are also presented for one response factor in this paper (figure 3).

Table 4. CONFIGURATION A. ANOVA table of the cylindricity of the outer cylinder

Source	Sum of Squares	df	Mean Square	F-value	p-value	
Block	0.0001	1	0.0001			
Model	0.0005	3	0.0002	13.49	0.0027	significant
A-Force Amplitude	0.0001	1	0.0001	11.94	0.0106	
B-Wall Thickness	0.0002	1	0.0002	19.73	0.0030	
A^2	0.0001	1	0.0001	8.80	0.0209	
Residual	0.0001	7	0.0000			
Lack of Fit	0.0001	4	0.0000	2.64	0.2255	not significant
Pure Error	0.0000	3	6.229×10^{-6}			
Cor Total	0.0006	11				

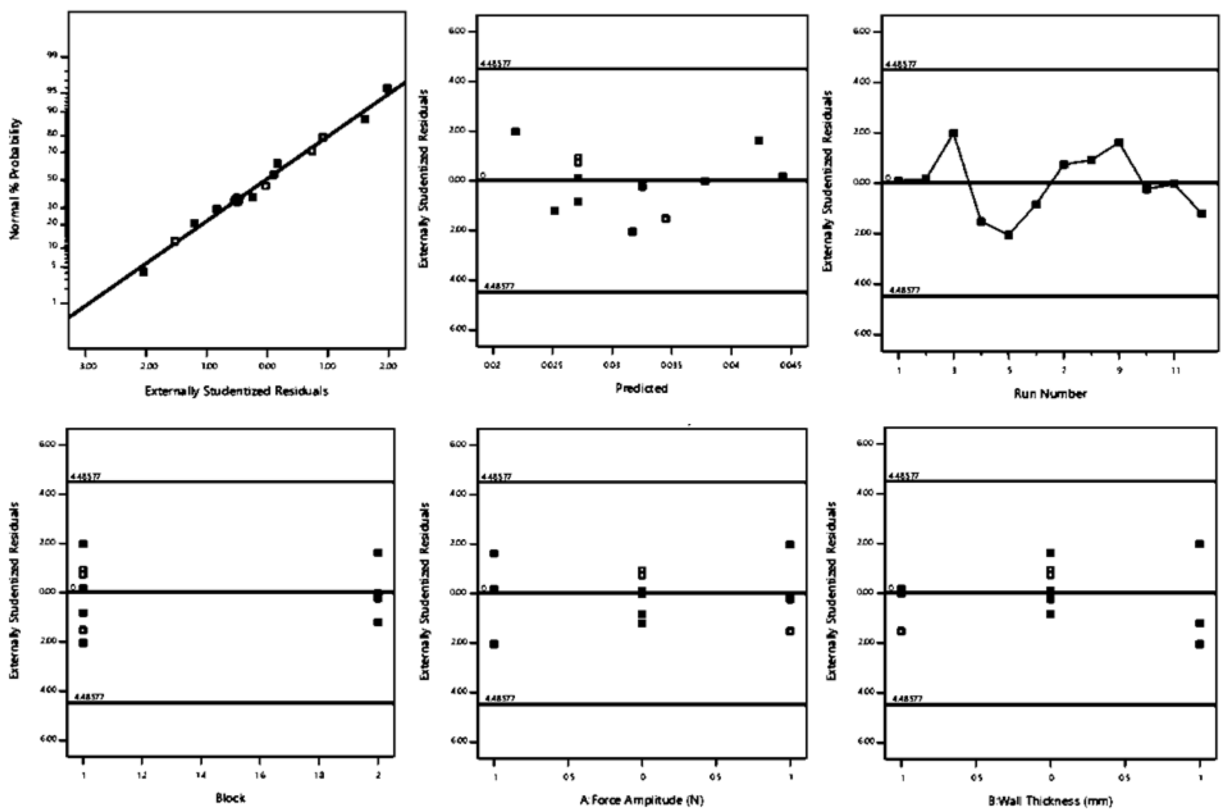


Figure 3. CONFIGURATION A. Residuals diagnostics of the cylindricity of the outer cylinder

Surface plots showing the relationship of the response factors as functions of the force amplitude (control factor A) and the wall thickness (control factor B) are shown in Figures 4 to 13. Figures 4 to 8 are for configuration A, where the force vibrates in the XY-plane (Figure 1). Figures 9 to 13 are for configuration B, where the force vibrates in the XZ-plane (Figure 1). In each figure, the left surface plot corresponds to measurements taken on the outer cylinder of the 3D printed part, and the right surface plot to measurements taken on the inner cylinder. The p-values from the analysis of variance (ANOVA) associated with control factor A, control factor B, and the interaction between these control factors are shown in the figure insets. The confidence level at which it can be concluded that the factor (A, B, or the interaction between A and B) significantly affected the response factor is 1 minus the corresponding p-value. If the confidence level exceeds 95 %, or the p-value is less than 5 %, then it is concluded that the control factor or factors interaction has a statistically significant effect on the response factor. If the factor did not affect the response significantly, the corresponding p-value is withdrawn from the ANOVA model. When this occurred, the p-value is shown as not applicable (N/A) in the figures.

For configuration A (force vibrating in the XY-plane, Figure 1):

An inspection of Figures 4 left and right shows that the cylindricity of the outer cylinder was affected by the force amplitude and the wall thickness, but these control factors did not interact; and that the cylindricity of the inner cylinder was only affected by the wall thickness. The tendency was for larger force amplitude and larger wall thickness to result in either improved cylindricity, or no change at all. An inspection of Figures 5 left and right shows that the perpendicularity of the outer and the inner cylinders was affected by the force amplitude and the wall thickness, but these control factors did not interact. The effect was slight as all the p-values were near or above 5 %. That is, the vibration and the wall thickness had a positive or null effect on cylindricity, and a nearly null effect on perpendicularity.

An inspection of Figures 6 left and right shows that the deviation in outer diameter was affected only by the wall thickness, and that the deviation in inner diameter was not affected by any of the factors. The deviation in outer diameter increased with increase in wall thickness. That is, an increase in all thickness caused the diameter of the outer cylinder to increase.

An inspection of Figures 7 and 8 left and right shows that none of the factors significantly affected the deviation from nominal position of the center of the outer and inner cylinders. This is, neither along direction X nor along direction Y (see Figure 1 for direction definitions).

For configuration B (force vibrating in the XZ-plane, Figure 1):

An inspection of Figures 9 left and right shows that the cylindricity of the outer cylinder was only affected by the wall thickness; and that the cylindricity of the inner cylinder was affected by the force amplitude and the wall thickness, and these control factors interacted. An inspection of Figures 10 left and right shows that the perpendicularity of outer cylinder was affected by the force amplitude and the wall thickness, and these control factors interacted; and that the perpendicularity of the inner cylinder was only affected by the force amplitude. In most cases where the control factor affected the response, or where the two control factors interacted, the larger the force amplitude and the smaller the wall thickness, the greater the cylindricity and/or the perpendicularity. That is, greater deviations from perfect cylindrical form and perfect perpendicular orientation between the printed cylinder and its base were promoted by larger force amplitudes and smaller wall thicknesses.

An inspection of Figures 11 left and right shows that the deviation in diameter was affected only by the wall thickness. The deviation in outer and inner diameter increased with increase in wall thickness. That is, an increase in wall thickness caused the diameter of the outer cylinder to increase but that of the inner cylinder to decrease.

An inspection of Figures 12 left and right shows that the deviation from nominal position of the center of the outer and inner cylinders, measured along X, was affected by the force amplitude and the interaction of force amplitude and wall thickness. As the wall thickness decreased, the effect of the force amplitude on this position error was more pronounced.

An inspection of Figures 13 left and right shows that none of the factors had a significant effect on the deviation from nominal position of the center of the outer and inner cylinders, measured along Y.

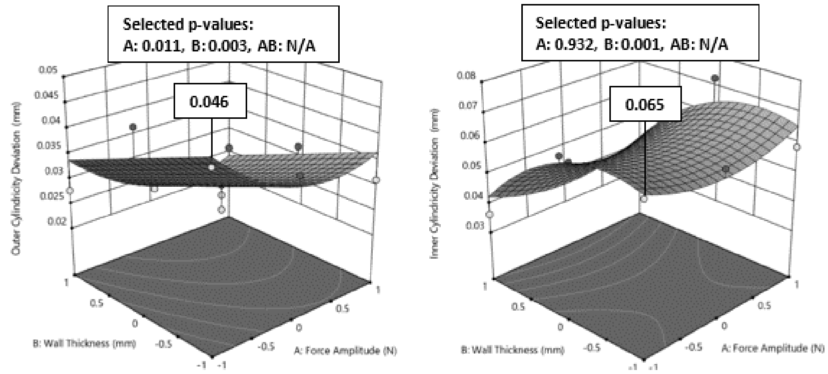


Figure 4. <u>CONFIGURATION A</u>. 3D plots of the cylindricity. Left: outer cylinder. Right: inner cylinder.

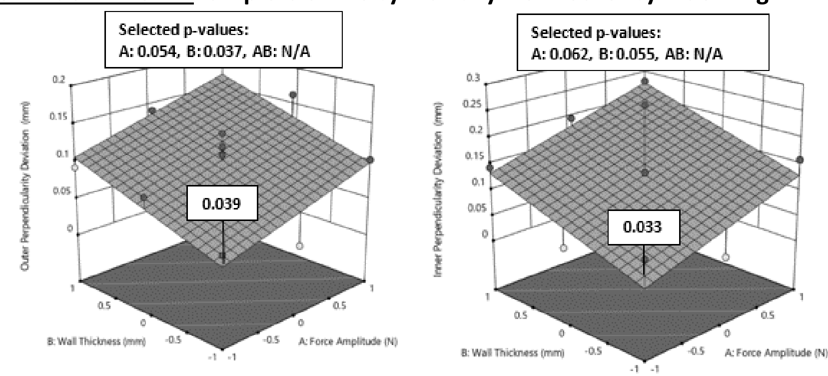


Figure 5. <u>CONFIGURATION A</u>. 3D plots of the perpendicularity. Left: outer cylinder. Right: inner cylinder.

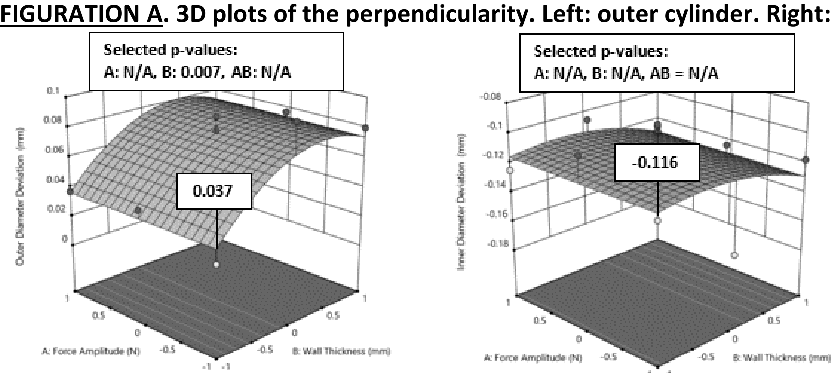


Figure 6. <u>CONFIGURATION A</u>. 3D plots of the diameter deviation. Left: outer cylinder. Right: inner cylinder.

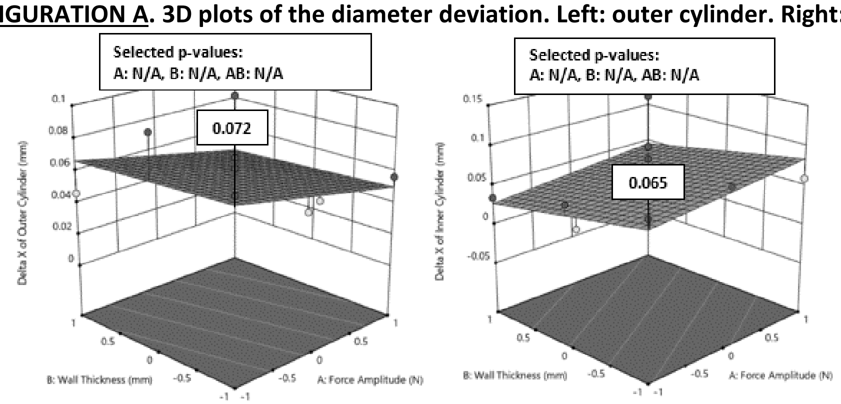


Figure 7. <u>CONFIGURATION A</u>. 3D plots of the position error along the X axis. Left: outer cylinder. Right: inner cylinder.

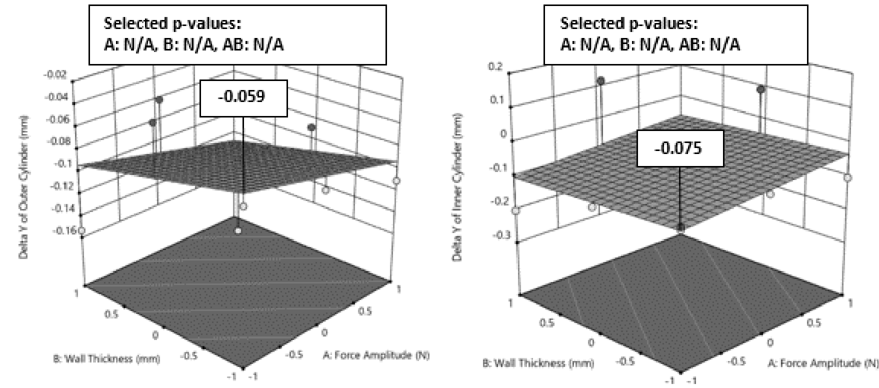


Figure 8. <u>CONFIGURATION A</u>. 3D plots of the position error along the Y axis. Left: outer cylinder. Right: inner cylinder.

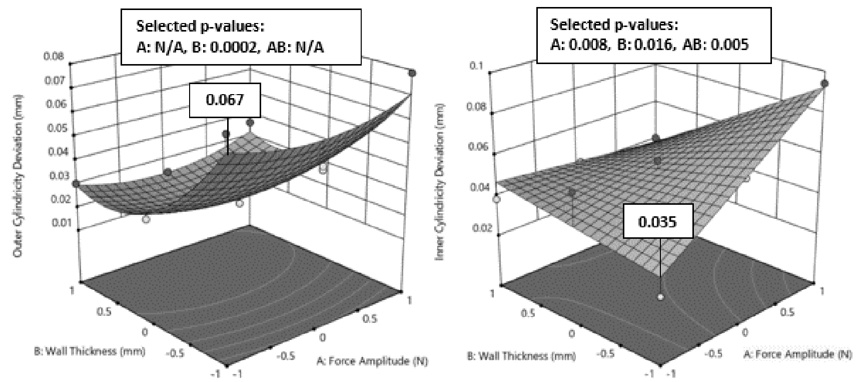


Figure 9. <u>CONFIGURATION B.</u> 3D plots of the cylindricity. Left: outer cylinder. Right: inner cylinder.

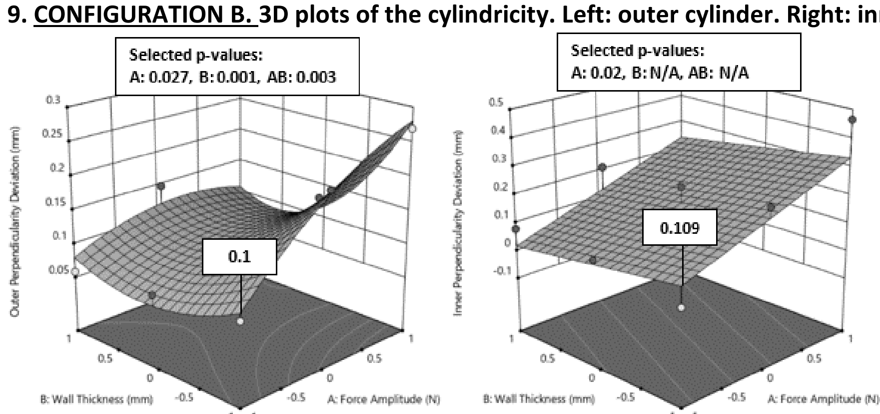


Figure 10. <u>CONFIGURATION B</u>. 3D plots of the perpendicularity. Left: outer cylinder. Right: inner cylinder.

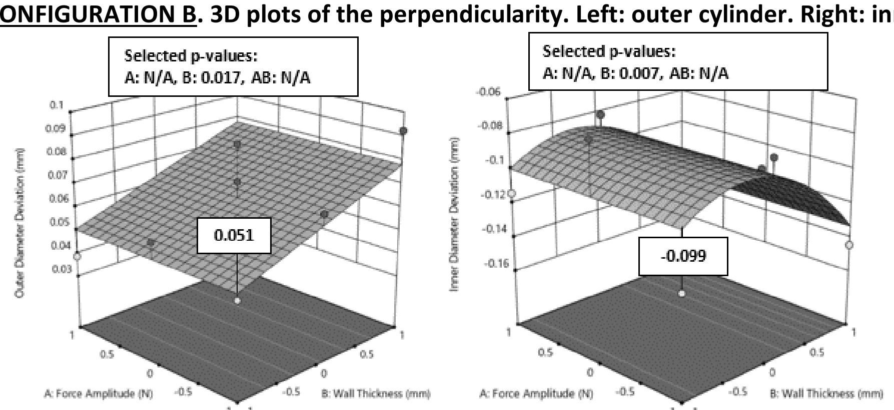


Figure 11. CONFIGURATION B. 3D plots of the diameter deviation. Left: outer cylinder. Right: inner cylinder.

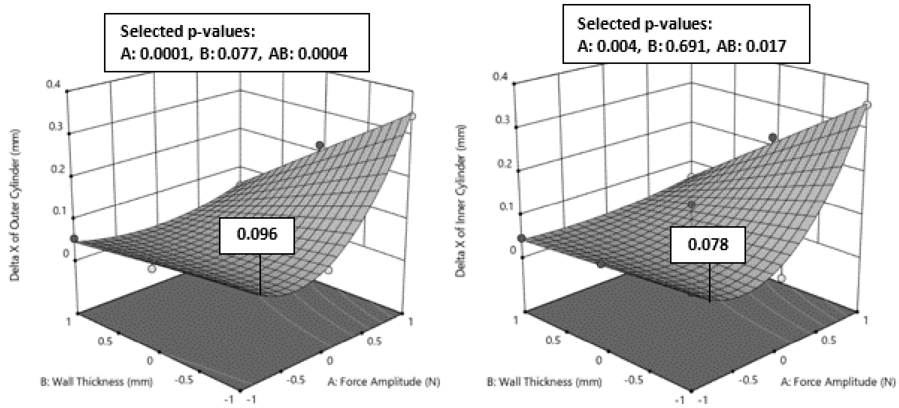


Figure 12. CONFIGURATION B. 3D plots of the position error along the X axis. Left: outer cylinder. Right: inner cylinder.

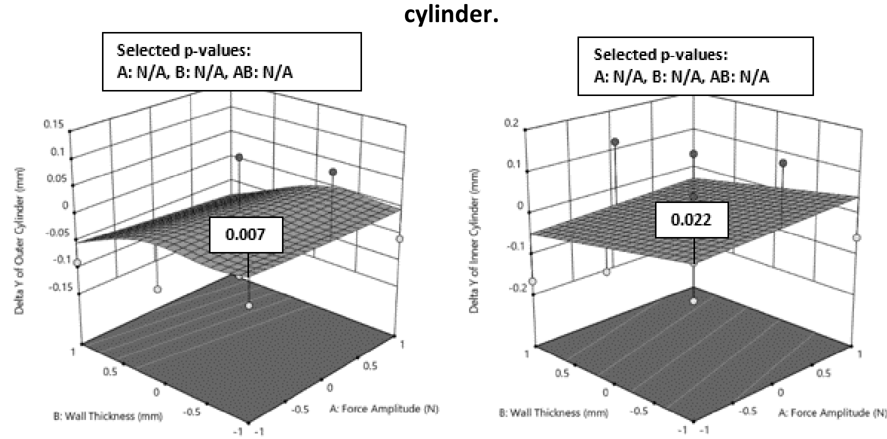


Figure 13. CONFIGURATION B. 3D plots of the position error along the Y axis. Left: outer cylinder. Right: inner cylinder.

4. Discussion

From the results presented above, geometric characteristics of form and orientation, such as cylindricity and perpendicularity, are potentially affected by vibrations of any mode, and that a given vibration amplitude is likely to cause more changes in these geometric characteristics as the feature's thickness decreases. The vibration appears to either improve these geometric characteristics, or leave it unaffected, if it is composed of longitudinal waves propagating symmetrically along directions X and Y, as in configuration A (Figure 1 left). The vibration appears to deteriorate the geometric characteristics, if it is composed of a mixed transverse (oscillating along Z) and longitudinal (along X) vibration mode as in configuration B (Figure 1 right). Although not directly observable from the data presented herein, it is expected that other size characteristics such as height of the 3D printed part and cross section shape (area moment of inertia) will affect the geometric characteristics in a way similar to wall thickness.

Also, from the results, the positions of the center of both the outer and the inner cylinders, but only when measured along X (Figure 1), were significantly affected by both the force amplitude and the wall thickness. However, the effects were significant only for configuration B, wherein the vibration is along directions Z and X. The largest deviation in the position of the cylinder's center, measured along X, occurred at large force amplitude and small wall thickness. The lack of effect of force amplitude and wall thickness on the position of these cylinders when configuration A was used is probably due to the symmetric propagation of vibration perpendicular to the build axis (Z-axis in

Figure 1).

From the above, it can be concluded that the part is more robust against vibrations as its thickness increase. That is, for vibration modes that deteriorate the geometric characteristics of the 3D printed part, there appears to exist a minimum wall thickness for which vibrations would not be a problem. For this 3D printing, vibration elimination would not be required.

The deviation in the diameter of the outer and the inner cylinders appeared to be independent of force amplitude. These diameters were affected, however, by the wall thickness. The deviation in diameter tended to increase with increase in wall thickness. A larger wall thickness caused the diameter of the outer cylinder to increase but that of the inner cylinder to decrease. The apparent independence between diameter and vibration is reassuring, since it suggests that size control during 3D printing would not require vibration elimination. The observed effect that the diameter deviation increased with increase in wall thickness is probably due to limitations of the 3D printer utilized here. However, further future study is needed to investigate the effect of vibration on other 3D printed geometries via different machines and using different printing materials.

5. Conclusion

It is concluded from the results presented in this study that vibrations propagated to the AM machine may have a positive or a negative impact on the dimensional and geometric accuracy of the printed parts, depending mainly on vibration mode. Symmetrically propagating vibrations in the build plane had a tendency to improve cylindricity and to leave perpendicularity and position unchanged. For vibrations propagating along one direction along the build plane and parallel to the build axis, the cylindricity of a model cylindrical part and the perpendicularity of this cylindrical part relative to its base had a tendency to display larger deviations from perfect cylindricity and perpendicularity as the applied force amplitude increased and the feature's wall thickness decreased. For these non-symmetrical vibrations, the deviation in the position of the cylinder's center also tended to increase as the applied force amplitude increased and the wall thicknesses decreased. However, this was observed only for measurements taken along the direction of the applied vibration along the build plane.

The deviation in the diameter was not significantly affected by the applied force amplitude, but it was significantly affected by the wall thickness. The deviation in the diameter increased as the wall thickness increased. The deviation always manifested as a greater spread in material radially out for outer features and radially inwards for inner features. That is, it resulted in an increase in the outer diameter and a decrease in the inner diameter.

6. References

Antony, J. (2014). Design of Experiments for Engineers and Scientists. Elsevier.

Doherty, J. J. J. (1885). An elementary text-book of mechanics (kinematics and dynamics).

Gasser, A., Backes, G., Kelbassa, I., Weisheit, A., & Wissenbach, K. (2010). Laser Additive Manufacturing. Laser Technik Journal, 7(2), 58–63. https://doi.org/10.1002/latj.201090029

Gunduz, I. E., McClain, M. S., Cattani, P., Chiu, G. T.-C., Rhoads, J. F., & Son, S. F. (2018). 3D printing of extremely viscous materials using ultrasonic vibrations. Additive Manufacturing, 22, 98–103. https://doi.org/10.1016/j.addma.2018.04.029

Kamal, M., & Rizza, G. (2019). Design for metal additive manufacturing for aerospace applications. In Additive Manufacturing for the Aerospace Industry (pp. 67–86). Elsevier. https://doi.org/10.1016/B978-0-12-

814062-8.00005-4

Koch, R. M., & Kuklinski, R. (2004). Method and system for production of fibrous composite prototypes using acoustic manipulation in stereolithography (United States Patent No. US6786711B2). https://patents.google.com/patent/US6786711B2/en

Kruth, J.-P., Leu, M. C., & Nakagawa, T. (1998). Progress in Additive Manufacturing and Rapid Prototyping. CIRP Annals, 47(2), 525–540. https://doi.org/10.1016/S0007-8506(07)63240-5

Mendenhall, W. M., & Sincich, T. L. (2015). Statistics for Engineering and the Sciences, Sixth Edition. Taylor & Francis.

Moroni, G., Syam, W. P., & Petrò, S. (2014). Towards Early Estimation of Part Accuracy in Additive Manufacturing. Procedia CIRP, 21, 300–305. https://doi.org/10.1016/j.procir.2014.03.194

Mukherjee, T., Zhang, W., & DebRoy, T. (2017). An improved prediction of residual stresses and distortion in additive manufacturing. Computational Materials Science, 126, 360–372. https://doi.org/10.1016/j.commatsci.2016.10.003

Myers, R. L., & Myers, R. L. (2006). The Basics of Physics. Greenwood Publishing Group.

Whorton, M. (2001, October 15). G-LIMIT - A microgravity vibration isolation system for the International Space Station. 2001 Conference and Exhibit on International Space Station Utilization. 2001 Conference and Exhibit on International Space Station Utilization, Cape Canaveral, FL, U.S.A. https://doi.org/10.2514/6.2001-5090