

SUMMER
2023



JOURNAL OF MANAGEMENT & ENGINEERING INTEGRATION

AIEMS
VOL 16
NO 1

ISSN: 1939-7984

JOURNAL OF MANAGEMENT AND ENGINEERING INTEGRATION

Editor-in-Chief

Edwin Sawan, Ph.D., P.E.
Professor Emeritus
Wichita State University
Edwin.sawan@wichita.edu

Associate Editor

Abdulaziz G. Abdulaziz, Ph.D.
Wichita State University
abdulaziz.abdulaziz@wichita.edu

AIEMS President

Gamal Weheba, Ph.D.,
Professor and ASQ Fellow
Wichita State University
gamal.weheba@wichita.edu

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 emphasizes the integrative nature of business, management, computers, and engineering within a global context.

Editorial Review Board

Mohammed Ali, Ph.D.
The University of Texas at Tyler
mohammedali@uttyler.edu

Sue Abdinnour Ph.D.
Wichita State University
sue.abdinnour@wichita.edu

Gordon Arbogast, Ph.D.
Jacksonville University
garboga@ju.edu

Deborah Carstens, Ph.D.
Florida Institute of Technology
carstens@fit.edu

Hossein Cheraghi, Ph.D.
West New England University
cheraghi@wne.edu

Mohammad Kanan, Ph.D.
University of Business & Technology, KSA
m.kanan@ubt.edu.sa

Tamer Mohamed, Ph.D.
The British University in Egypt
tamer.mohamed@bue.edu.eg

Nabin Sapkota, Ph.D.
Northwestern State University, LA
sapkotan@nsula.edu

Nabin Sapkota, Ph.D.
Northwestern State University, LA
sapkotan@nsula.edu

Scott D. Swain, Ph.D.
Clemson University
sdswain@clemson.edu

Alexandra Schöning, Ph.D.
University of North Florida
aschonni@unf.edu

John Wang, Ph.D.
Montclair State University
wangi@montclair.edu

Gamal Weheba, Ph.D.
Wichita State University,
gamal.weheba@wichita.edu

Wei Zhan, D.Sc., PE
Texas A&M University
wei.zhan@tamu.edu

Reviewers

The Journal Editorial Team would like to thank the reviewers for their time and effort. The comments that we received were very constructive and detailed. They have been very helpful in our effort to continue to produce a top-quality journal. Your participation and timely response are very important for the success in providing a distinguished outlet for original articles. In this issue we continue to include Keywords, and the dates the publication was submitted and revised in an effort to achieve a higher standard for publication and increase the impact of the journal.

Edwin Sawan, Ph.D., P.E.
Editor-in-Chief

Abdulaziz Abdulaziz

Hemaid Alsulami

Abdelhakim Al Turk

Eylem Asmatulu

Ramazan Asmatulu

Ryan Atkins

Abdurrahman Basalan

LuAnn Bean

Paulo Cauchick

Dia Ali

Andrzej Gapinski

Ramkumar Harikrishnakumar

Hongsheng He

Abdelnasser Hussein

Bassam Jaradat

Steven Jiang

Krishna Krishnan

Adam Lynch

Lynn Matthews

Roger Merriman

Narasimha Nagaiah

Paul Nugent

R. Radharamanan

Clovis Ribas

Kaushik Sinha

Priyanka Thakur

Rick Wallace

Gamal Weheba

Bayram Yildirim

Mohammed Zwawi

Table of Contents

The US DOD Budget- Can it be Predicted?	1
Information Flow Theory: Circuit Network Approach	11
Large Scale Analytics for Workload Segmentation	19
Service Quality Between Tourism and Pilgrimage: A Literature Review	27
U.S. Wildlife Strikes by Phase of Flight	39
Unhelpful and Unaware of It: A Dyadic Analysis of Online Product Reviews	48
Internal and External Precursors to Favorable Adoption of Industry 4.0 Technologies	57

The US DOD Budget- Can it be Predicted?

Gordon W. Arbogast, Ph.D.¹

Arpita Jadav DBA¹

¹*Jacksonville University, FL*

garboqa@ju.edu; ajadav@ju.edu

Abstract

The Department of Defense (DOD) is part of the United States Federal Government which oversees the U.S. Military. This Department is one of the largest and most complex organizations in the world. The DOD mission is to protect and defend the United States (US) and provide national security. To achieve this, the DOD requires a major portion of the federal budget. Each year, the DOD portion is based on a variety of political and economic factors. The results of this study are noteworthy. A regression model was derived that explained 82.14% of the variation in the target ratio of the federal budget with a significance level of 0.05. Four variables were identified and listed in order of greatest impact, as determined, by their standardized coefficients. These variables may have a significant relationship with DOD's budget. The four variables are: (a) House Majority Political Party, (b) Doomsday Clock Value, (c) US President's Political Party Affiliation, and (d) US Gross Domestic Product Growth Rate. If corporations and other agencies that deal with the DOD were to be able to accurately predict year-by-year DOD budget levels, this would give them a unique, competitive advantage. The strong presence of political factors in the results may be a key indicator for DOD businesses to consider in ensuring the balance of an appropriate level of politically motivated drivers within their corporate strategy models. Further recommendations focused primarily on the factors and, ultimately, the variables that should be selected for future studies. Variables need to be selected to allow for a greater number of observations to increase the likelihood of producing accurate study results.

Keywords: US Federal Budget, DOD Budget, Doomsday Clock.

1. Background

A major part of the US Federal Budget is the DOD portion which is needed to support its mission to protect and defend the United States and provide security for its citizens. With a defense budget of more than \$611 billion in 2016, the US led the world in overall defense spending (McCarthy, 2017). A quick analysis of the U.S. 2016 Defense budget shows that it exceeded the combined annual defense spending budgets of the next eight countries by \$16 billion (The Peterson Foundation, 2017). These countries include China, Russia, and Saudi Arabia, India, France, the United Kingdom, Japan, and Germany, as depicted in Table 1. (U.S. Defense Spending Compared to Other Countries, 2017). A look at the 2022 US DOD Budget shows a major increase in spending for the DOD (up to \$801 billion), while the next nine countries are now spending a cumulative amount of \$777 billion. China is still second (spending in excess of \$200 billion) while India is third spending close to \$100 billion on defense. The

other countries are still relatively the same with the exception of the addition of South Korea; all of these countries spending less than \$80 billion apiece on Defense (Peterson Foundation, 2022).

US Defense spending represents the third largest budget item within the US federal budget following only Medicare and Social Security. The components of the total US Federal Budget in 2015 are as follows (Policy Basics, 2017): (a) Healthcare – 25%; (b) Social Security – 24% ; (c) DOD Discretionary – 16%; (d) Non-Defense Discretionary – 16%; (e) Mandatory Spending – 12%; and (f) Interest Payments – 6%. Looking at the structure of the DOD, it employs over 3.3 million employees, including 1.4 million active-duty military, 1.1 million reserve forces, and 861,000 civilians (U.S. Department of Defense Agency Strategic Plan Fiscal Years 2015-2018, 2018). Due to the sheer scope of this organization, millions of people are either directly or indirectly dependent on the DOD for their security and livelihood.

It is important to be aware of the reach of the DOD in the nation's industrial infrastructure. In the fiscal year 2015, the DOD contracted directly with more than 50,000 companies. The top five firms, commonly referred to as primes are Lockheed Martin, Boeing, Raytheon, General Dynamics, and Northrop Grumman. These five firms received 28% of the DOD's available contract dollars in that year. It is noteworthy that the number of 50,000 companies contracted by the DOD does not include the various contractors and subcontractors utilized by the primes (Rumbaugh & Peters, 2017). This speaks to the sheer scale of the organization's impact. Thus, understanding the scope and reach of the DOD is vital for analyzing the political, economic and social factors that may influence its budget levels. The ability to predict these budget levels may provide defense contractors and other supporting firms with additional means to assist in forecasting potential business opportunities and capital expenses.

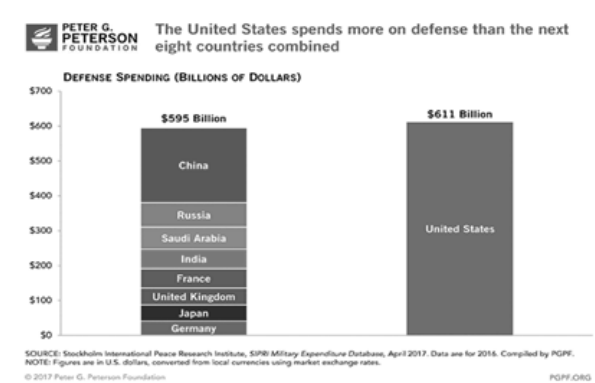


Table 1. Defense Spending by Count

2. Literature review

The Federal Budget Process is set forth each year in a document published by Capitol.Net. The document explains the overall process that is used in arriving at the final budget. It describes the steps that begin with each part of the Federal Government sending its requirements to a centralized location; These inputs then follow a detailed process in order to create the US Federal Budget (Capitol.Net, 2012). DOD budgets are an integral part of the final budget. These budget levels are an obvious concern for those parties interested in the maintenance and operations of the US Military, as well as the security and defense of the United States and its global interests. The federal budgeting process is complex. However there are books that describe more in detail how the budgeting process works. One such book

is Budget Tools: Financial Methods in the Public Sector. It provides a thorough discussion of public budgeting, by organizing and processing data through thorough analysis and presentation (Chen, Weikart, and Williams, 2015). The DOD portion of the budgeting process came under great scrutiny during Kennedy's presidency. Secretary of Defense, Robert McNamara undertook a major overhaul of the DOD process. What came out of careful due diligence was his famous PPBS System. PPBS stands for planning, programming and budgeting systems, and this process has become an integral part of the Pentagon's annual budgeting system (Enthoven, Smith 2005). Over the years, a few research groups have published studies on the makeup, and use of DOD budget dollars. Such efforts have been described in a number of books, including one by Jones and McCaffrey. Their book specifically focuses on financial management and business process challenges and issues. It describes the processes used to acquire resources for defense war fighting assets, including reforms in acquisition and linkages between the PPBS and the defense acquisition process (Jones and McCaffrey, 2008).

Another prominent institution among these researchers is the independent and nonprofit Center for Strategic and International Studies (CSIS). It was founded during the height of the Cold War, with a mission to find "ways to sustain American prominence and prosperity as a force for good in the world (Harrison & Daniels, 2018)." Since 1962, CSIS has focused a significant portion of its resources on issues related to defense and security including topics concerning the state and direction of the US DOD's budget. The CSIS's 2018 report on the US defense budget points to a number of items that affect the DOD budget, including Continuing Resolutions (CR), the Budget Control Act (BCA), and ongoing defense related strategy reviews commissioned by the US Government. The report described Continuing Resolutions as a temporary funding measure used in the absence of an approved federal budget. These resolutions limit the ability of government agencies, including the DOD, to plan their budgets effectively and match their available resources with assigned responsibilities. Unpredictability in the funding of the US federal budget may also negatively affect the US economy, which in turn can have an adverse effect on the level of discretionary funds available to the DOD. The CSIS report further explains that the Budget Control Act and similar acts enacted in previous years place limits on the DOD's discretionary budget. However, certain defense related items such as the DOD budget allocation in support of war-related efforts, were excluded from the act. Additionally, the CSIS report explains how the ongoing Defense Strategy Review, the Nuclear Posture Review, and the Ballistic Missile Defense Review can lead to significant changes in future DOD plans and budget allocations (Ibid, 2018).

Several other Defense Analysts have attempted to describe the process by which the DOD budget is determined. In 2009, O'Hanlon described how the military budget works, how the military assesses and deploys new technology, develops strategy and fights wars, handles the logistics of stationing and moving troops and equipment around the world, and models and evaluates battlefield outcomes (O'Hanlon, 2009). However, O'Hanlon and other analysts have been unable to produce a definitive model to help predict DOD budgeting. The Strategic Studies Institute (at the Army War College) conducted similar studies pertaining to the DOD's budget. The Institute's 2005 report describes a number of external factors that apply pressure to the defense budget. These items include budget and trade deficits, anticipated growth in nondiscretionary federal expenditures, the reform of Social Security and other entitlement programs, and the increased demand for discretionary spending including that in support of homeland security. The Institute's report also noted a number of internal factors that exert

pressure on the defense budget. These items include the increase in operations and maintenance expenditures incurred per soldier, the impact of military operations in the Middle East, and the need for continued replacement, reconstitution, and modernization efforts to support the military's infrastructure (White, 2005). The John F. Kennedy School of Government at Harvard University has also provided research for this study to the Government. A report from March 2013 details how the operations in Iraq and Afghanistan will have a lasting impact on the DOD's budget due to "long-term medical care and disability compensation for service members, veterans and families, military replenishment and social and economic costs encountered as a byproduct of Middle East operations" (Bilmes and Stiglitz, 2013). Additionally, they describe how the US has experienced unprecedented growth since 2001 in budgets by supporting the Department of Veterans' Affairs (Ibid, 2013)".

3. Problem statement and hypothesis

The DOD is charged with protecting and defending the United States and providing security for its citizens. (Jonas, Tina W. Under Secretary of Defense (Comptroller), 2008). To continue its mission, the DOD requires a budget apportionment commensurate with its needs and in a proper proportion to the overall US federal budget.

The Problem Statement in this study aims to analyze the relationship between several political, economic, and social factors to determine the degree to which these factors affect the ratio of the DOD annual budget to the total US federal budget.

The independent variables were: (a) Geopolitical Consolidated Risk Index; (b) Doomsday Clock Value (Timeline, 2018); (c) US Senate Majority Political Party; (d) US House Majority Political Party; (e) US President Political Party Affiliation; (f) US Consumer Confidence Index; (g) US Economic Situation Index; (h) US Gross Domestic Product Growth Rate; and (i) US Federal Budget Sequestration

The Null Hypotheses is: There **is no** relationship between the ratio of the DOD annual budget to the total US Federal Budget and any of the nine independent variables cited above.

The Alternative Hypotheses is that: There may be a relationship between the ratio of the DOD annual budget to the total US federal budget and the independent variables cited above.

4. Research design and methodology

Data were gathered from several sources to acquire the information necessary to complete the analysis. The data included information regarding the performance of the Republican and Democratic Parties in the US House, Senate, and Presidential elections, as well as the levels of US Federal and DOD budgets from 1978 to 2016. Additional details for the study were gathered relating to the perceived levels of geopolitical threats quantified by the Geopolitical Consolidated Risk Index and the Doomsday Clock, as well as economic factors such as the rate of growth of the US Domestic Product, US Consumer Confidence Index, and economic sentiment indicators representing US consumers' confidence in the overall economy again from 1978 to 2016.

To test the hypotheses defined for this study, the following research methodology was established:

- Study data were collected from January 1st, 1978, through December 31st, 2016. An analysis was conducted to determine which independent variables (X) should remain in the model and which should be removed. A One-Way Analysis of Variance (ANOVA) test on the categorical factors was performed to determine if there were any statistically significant differences between the mean of each group and the mean of the ratio relating the DOD annual budget to the total US federal budget. The variables with P-values that exceeded 0.05 (α) were removed from the regression model.
- Correlation analysis was performed on continuous independent variables. A Correlation Coefficient threshold of 0.5 and a p-value threshold of 0.05 was used to determine which additional variables were candidates to be removed from the model to reduce the effects of multicollinearity. To validate the correlation analysis results, a full multiple regression analysis was performed using continuous and categorical independent variables. The Variation Inflation Factors (VIF) of each of the independent variables were compared with the Correlation Coefficients noted in the ANOVA Test. This formed the basis for removing independent variables from the model.
- A regression analysis using Stepwise Backward Elimination of Terms was then conducted. The test started with a full model including all remaining variables and successively removed each predictor with a P-value greater than 0.05 until all remaining predictors met the criterion to remain in the model. A final multiple regression analysis of the remaining terms was conducted to validate the resulting regression equation against the sample data.

The regression model incorporated 39 observations, that met the established criteria for a statistically relevant population. Additionally, the residual scatterplot confirmed the assumption of homoscedasticity between the predicted dependent variable scores and the prediction errors, showing a reduced risk of making Type I and Type II errors during the analysis.

The dependent and independent variables referenced in the study are defined below, with all data values used for the analysis included in the Appendix section of this report (See Appendix A: **Research Data**). The details provided for each variable include, the origin of the data, a description of the information provided by the variable, and the variable type (continuous or categorical). The dependent variable was as follows:

(Y) Ratio of DOD Annual Budget to the total US Federal Budget.

a) Description: A derived value based on the ratio of the DOD annual budget to the total US federal budget. (Office of Management and Budget, 2018)

b) Variable Type: Continuous

The independent variables are shown as follows:

- ***(X₁) Geopolitical Consolidated Risk Index***
Description: An index developed by Dario Caldara and Matteo Iacoviello to measure the level of geopolitical risk based on a tally of newspaper articles covering geopolitical tensions (Caldara & Iacoviello, 2018). **Variable Type:** Continuous

- **(X₂) Doomsday Clock Value**
Description: “The value symbolizes the urgency of the nuclear dangers that the magazine's founders — and the broader scientific community — are trying to convey to the public and political leaders around the world.” (Timeline, 2018). **Variable Type:** Continuous
- **(X₃) US Senate Majority Party**
Description: The value represents one of the following three conditions regarding the majority party in the US Senate: (Composition of Congress, by Political Party, 1855-2017, 2018)
Variable Type: Categorical. 0 Democrat Majority; 1 Republican Majority; 2 Equal members/party
- **(X₄) US House Majority Party**
Description: The value represents one of the following three conditions regarding the majority party in the US Senate: (Composition of Congress, by Political Party, 1855-2017, 2018)
Variable Type: Categorical. 0: Democrat; 1: Republican
- **(X₅) US President Party Affiliation**
Description: This value represents one of the following two conditions regarding the political party associated with the US President: (Chart of Presidents and Vice Presidents, 2018)
Variable Type: Categorical. 0: Democrat; 1: Republican
- **(X₆) US Consumer Confidence Index**
Description: A measure of consumer optimism based on household plans for spending and saving in both the short-term and immediate timeframes. (Consumer Confidence Index (CCI), 2018)
Variable Type: Continuous
- **(X₇) US Economic Situation Index**
Description: A measure of US citizen’s optimism in the country’s general economic situation over the next 12-month period). (Ni, Baker, Bloom, & Davis, 2018); **Variable Type:** Continuous
- **(X₈) US Gross Domestic Product Growth Rate**
Description: A measure of the rate of growth in US GDP from the prior year as a percentage.: (US GDP by Year Compared to Recessions and Events, 2018); **Variable Type:** Continuous
- **(X₉) Sequestration Events within the US Federal Budget**
Description: The value represents one of the following two conditions regarding the level of the US Federal Budget compared to the prior year: (Congressional Budget Office, 2018), (US Defense Spending History, 2018); 0: Decrease in US federal budget compared to previous year ; 1: Increase in US federal budget compared to previous year; **Variable Type:** Categorical

5. Results

The results of the one-way ANOVA, correlation analysis, and backwards stepwise regression tests removed the following predictors as they did not meet the stated criterion to remain in the model which included a P-value < 0.05 and a Correlation Coefficient < 0.5 (See Appendix B: **Descriptive Statistics**)

(X₁) Geopolitical Risk; (X₃) US Senate Majority Political Party; (X₆) US Consumer Confidence Index
(X₇) US Economic Situation Index; (X₉) US Federal Budget Sequestration

Multiple Regression Analysis performed on the remaining factors produced the following (Table 2):

Full Regression Test Using Remaining Continuous and Categorical Factors					
Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	0.22196	0.00929	23.88	0.000	-
(X ₂) Doomsday Clock Value	-0.003538	0.000728	-4.86	0.000	1.09
(X ₄) US House Majority Party	-0.05088	0.00640	-7.95	0.000	1.15
(X ₅) US President Political Party Affiliation	0.02766	0.00652	4.24	0.000	1.20
(X ₈) US Gross Domestic Product Growth Rate	0.595	0.164	3.63	0.001	1.04

Table 2. Multiple Regression Output

These results provided the general regression equation: $Y = \beta_0 + \beta_4 X_4 + \beta_6 X_6 + \beta_7 X_7 + \beta_{10} X_{10}$. By substituting the actual variable names and associated coefficients in the general regression equation, the final regression equation is as follows:

$$\begin{aligned} \text{Ratio of DOD Budget to US Federal Budget} = & 0.22196 - 0.003538 (X_2) \text{ Doomsday Clock Value} \\ & - 0.05088 (X_4) \text{ US House Majority Party} + 0.02766 (X_5) \text{ US President Party Affiliation} \\ & + 0.595 (X_8) \text{ US Gross Domestic Product Growth Rate} \end{aligned}$$

6. Discussion

The results of this analysis will allow for: (1) a failure to reject one or more of the null hypotheses, or (2) the ability to reject one or more of the null hypotheses in favor of one or more of the stated alternate hypotheses. Based on the results of the regression test, 82.14% of the variation in the ratio of the DOD annual budget to the total US federal budget can be explained by the derived model. The sample outcome derived in the analysis would be unlikely if the null hypotheses for the remaining variables were true.

6.1. Doomsday Clock Value:

If all other variables are held constant, a one-minute increase in the Doomsday Clock's value will decrease the ratio of the DOD's annual budget to the total US federal budget by 0.3538%. This outcome is consistent with the expectation that the DOD would afford a greater percentage of the US Federal Budget in times of heightened nuclear uncertainty, and a lower ratio in times of reduced nuclear concerns.

6.2. US House Majority Party:

If all other variables are held constant, the presence of a Republican majority in the US House of Representative would decrease the ratio of the DOD annual budget to the total US federal budget by 5.088%. A more thorough review of the Republican and Democratic Party Platforms is needed to better understand this result.

6.3. US president party affiliation:

If all other variables were held constant, the presence of a Republican President in the White House would increase the ratio of the DOD annual budget to the total US federal budget by 2.766%. A more thorough review of the evolution of the Republican and Democratic Party Platforms is needed to better understand this result.

6.4. US gross domestic product growth rate:

If all other variables are held constant, a 1% increase in the US GDP growth rate increases the ratio of the DOD annual budget to the total US federal budget by .00595%. This outcome is consistent with the expectation that an increase in GDP would allow for an overall increase in the US federal budget with an increase in the DOD's budget.

The model was applied to the sample data and produced the projected ratio of the DOD annual budget to the total US federal budget (see Appendix: Projected Ratios using the Final Regression Equation). The model produced an average error rate of 0.75% when calculating the DOD annual budget to the US federal budget ratio when applied to the 39 samples included in the analysis. An additional review of the predictions uncovered the following:

- For 2009, the model over-estimated the ratio of the DOD's budget to the US federal budget by 33.86%. This year, the actual ratio was 19.29% less than that in 2008, representing the sharpest marginal ratio decline observed in the study. At face value, this observation is misleading as the assumption is that defense spending was lower in 2009 compared to spending levels observed in 2008. However, the opposite was true. The DOD's budget in 2009 was one of the highest levels observed during the sample period; however, so was the overall US federal budget. This combination produced an uncharacteristic decline in the ratio of the DOD's budget to the US federal budget.
- In 2011, the model underestimated the actual budget ratio by 17.29%. This appears to be the case where relatively larger swings in the US GDP growth rate cause a greater risk of error in the model. The impact of this condition can most likely be addressed by increasing the sample size in the analysis.

Null Hypotheses Evaluation			
Null Hypotheses	Accept Null Hyp	Reject Null Hyp	Alternative Hypotheses Accepted
There is no relationship between the ratio of the US Department of Defense annual budget to the total US federal budget and the:			There may well be a relationship between the ratio of the US Department of Defense annual budget to the total US federal budget and the:
1. Geopolitical Consolidated Risk Index	X		1. N/A
2. Value of the Doomsday Clock		X	2. Value of the Doomsday Clock
3. US Senate Majority Political Party	X		3. N/A
4. US House Majority Political Party		X	4. US House Majority Political Party
5. US President Political Party Affiliation		X	5. US President Political Party Affiliation
6. US Consumer Confidence Index	X		6. N/A
7. US Economic Situation Index	X		7. N/A
8. Growth rate of the US GDP		X	8. Growth rate of the US GDP
9. US Federal Budget Sequestration	X		9. N/A

Table 3. Null Hypotheses Evaluations

Based on the results of the analysis, the evaluations in Table 3 are presented for the stated hypotheses.

7. Conclusions

The DOD requires budget an apportionment commensurate with its needs and in proper proportion to the overall US federal budget. A well-funded DOD is vital for ensuring the safety and security of the United States and its citizens. This study considers the relationship between several political, economic, and social factors to determine the degree to which these factors affect the ratio of the DOD annual budget to the total US federal budget. The results of the study produced a regression model that explained 82.14% of the variation in the target ratio, with a significance level of 0.05. The study concluded that the following four variables, listed in order of greatest impact as determined by their standardized coefficients (see Appendix Section: Standardized Coefficients), may well have a relationship with the DOD's budget represented by the ratio comparing the overall DOD annual budget to the total US federal budget:

- US House Majority Political Party
- Doomsday Clock Value
- US President Political Party Affiliation
- US Gross Domestic Product Growth Rate

The ability to predict DOD budget levels may provide defense contractors and other supporting firms with additional means to assist in forecasting potential business opportunities and capital expenses. The strong presence of political factors in the final regression model may be a key indicator for professionals dealing with DOD to ensure that they are balancing an appropriate level of politically motivated points within their corporate strategy models.

8. Recommendations

The appropriate selection of variables for such a study is critical. Nine variables were selected for this study from a variety of social, political and economic areas. A variable from each category was retained in the final regression equation. Therefore, it is recommended that future research continue to include these areas of study. By continuing the review in this manner, researchers will be able to determine the most specific variables from each category that may well have a relationship with the value of the ratio of the DOD budget to the U.S. Federal Budget. Second, in this study the selection of variables served to limit the number of observations because of the lack of data for some of the variables prior to 1978. Future studies need to select a set of variables that will allow for a greater number of observations to increase the likelihood of producing accurate results. Finally, an effort needs to be made to improve the development of new indices to secure additional metrics for study. While the Geopolitical Risk and Economic Situation Indices did not remain in the final regression equation, it is plausible that parts of these indices may combine with additional data to create new and more meaningful measurements.

9. References

- Bilmes, Linda and Stiglitz, Joseph E. The Three Trillion Dollar War: The True Cost of the Iraq Conflict. John F. Kennedy Center Reports (Public Policy), March, 2013.
- Caldara, D., & Iacoviello, M. (2022). Measuring Geopolitical Risk. In American Economic Review (Vol. 112, Issue 4, pp. 1194-1225), American Economic Association. <https://doi.org/10.1257/aer.20191823>
- Capital Net. The Federal Budget Process, Congressional Research Service, 2012.

- Martin, K. Chart of the presidents and vice presidents. ThoughtCo., Chicago
<https://www.thoughtco.com/presidents-and-vice-presidents-chart-4051729>
- Chen, G. G., Weikart, L. A., & Williams, D. W. (2014). Budget tools: financial methods in the public sector. CQ Press.
- Composition of Congress, by Political Party, 1855-2017. (2018, March 10). Retrieved from infoplease:
<https://www.infoplease.com/history-and-government/us-government/composition-congress-political-party-1855-2017>
- Congressional Budget Office. (2018, March 11). Defense and National Security. Retrieved from Congressional Budget Office: <https://www.cbo.gov/topics/defense-and-national-security>
- Consumer Confidence Index (CCI). (2018, March 10). Retrieved from OECD Data: <https://data.oecd.org/leadind/consumer-confidence-index-cci.htm>
- Enthoven, A. C. and Smith, K. (1999) How Much is Enough?: Shaping the Defense Program, 1961-1969. The Rand Corporation, 2005.
- Harrison, T., & Daniels, S. (2021). Analysis of the FY 2018 Defense budget. Washington: Center for Strategic & International Studies.
- Jonas, Tina W. Under Secretary of Defense (Comptroller). (2008). US Department of Defense Fiscal Year 2009 Budget Request. Washington: US Department of Defense.
- Jones, Lawrence and McCaffrey, Jerry. Budgeting, Financial Management and Acquisition Reform in the U.S. Department of Defense, Information Age Publishing, March 2008.
- McCarthy, N. (2017, April 24). The Top 15 Countries for Military Expenditure in 2016. Retrieved from Forbes: <https://www.forbes.com/sites/niallmccarthy/2017/04/24/the-top-15-countries-for-military-expenditure-in-2016-infographic/#1b1d4a2843f3>
- Ni, H., Baker, S., Bloom, N., & Davis, S. (2018, March 10). Global Economic Policy Uncertainty Index. Retrieved from Economic Policy Uncertainty: http://www.policyuncertainty.com/global_monthly.html
- Office of Management and Budget. (2018, March 10). Fiscal Year 2017 - Historical Tables - Budget of the U.S. Government. Retrieved from Government Publishing Office (GPO): <https://www.gpo.gov/fdsys/pkg/BUDGET-2017-TAB/pdf/BUDGET-2017-TAB.pdf>
- O'Hanlon, Michael E. The Science of War: Defense Budgeting, Military Technology, Logistics, and Combat Outcomes, Princeton University Press, 2009.
- Office of Management and Budget. (2018, March 03). Historical Tables. Retrieved from whitehouse.gov: <https://www.whitehouse.gov/omb/historical-tables/>
- The Peterson Foundation, US Defense Spending Compared to Other Countries, Source: Stockholm International Peace Research Institute, SIPRI Military Expenditures Data Base, April 2017.
- The Peterson Foundation, US Defense Spending Compared to Other Countries, Source: Stockholm International Peace Research Institute, SIPRI Military Expenditures Data Base, April 2022.
- Policy Basics: Where Do Our Federal Tax Dollars Go? (2017, October 11). Retrieved March 25, 2018, from <https://www.cbpp.org/research/federal-budget/policy-basics-where-do-our-federal-tax-dollars-go>
- Rumbaugh, R., & Peters. H. (2017, February 10). Defense Primer: DOD Contractors. Retrieved March 21, 2018, from <https://fas.org/sgp/crs/natsec/>
- Timeline. (2018, March 03). Retrieved from Bulletin of the Atomic Scientists: <https://thebulletin.org/timeline>
- U.S. Defense Spending Compared to Other Countries. (2017, June 01). Retrieved March 23, 2018, from https://www.pgpf.org/chart-archive/0053_defense-comparison
- U.S. GDP by Year Compared to Recessions and Events. (2018, March 03). Retrieved from theBalance.com: <https://www.thebalance.com/us-gdp-by-year-3305543>
- US Defense Spending History. (2018, March 11). Retrieved from usGovernmentSpending.com: https://www.usgovernmentspending.com/defense_spending
- White, J. P. (2005). Transformation for What? Carlisle: Strategic Studies Institute United States Army War College.

Appendices:

Available on Request (Statistics and Data)

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Information Flow Theory: Circuit Network Approach

Abdurrezzak Sener¹

Mehmet Barut²

M. Bayram Yildirim²

¹*Penn State University*

²*Wichita State University*

aks6310@psu.edu; mehmet.barut@wichita.edu; bayram.yildirim@wichita.edu

Abstract

With the rapid development of the idea of integration and cooperation and the proliferation of information systems and technology in supply chains, the importance of information, and its management are becoming more and more evident. In this chapter, we explore how the instruments used in a circuit network be applicable to a supply chain network. We borrow the laws developed in electric circuit theory and conceptually explain the relationship and open discussion on how to apply it for future researchers.

Keywords: Information flow; supply chain management; circuit networks; information usage

1. Introduction

A supply chain is a sequence of processes that consist of the flow of information and materials between members of the supply chain network. This study focuses on the flow of information between supply chain members. In the literature, the flow of information, as well as its intensity and extent, are investigated in different ways. Some focused on the state of flow such as internal or external (Flynn et al., 2010; Yu et al., 2013; Huo et al., 2014). Others have examined categorical levels such as operational, tactical, and strategic (Mentzer et al., 2001; Benton and Maloni, 2005; Swink et al., 2007). While some focus on the extent that information technologies are utilized as enablers (Gilmour, 1999; Vickery et al., 2003), others investigate the barriers (security, infrastructure, traditions, and perceptions) (Harland et al., 2007; Govindan et al., 2014; Mathiyazhagan et al., 2013). Some, however, are interested in the direction of the flow such as upstream and downstream (Ganesh et al., 2014; Prajogo and Olhager 2012). The main reason behind the investigation of flow is that it is an essential factor for achieving and improving the efficiency and effectiveness of the supply chain (Zhao et al., 2013; Vickery et al., 2003).

In a supply chain network, information always exists. The most important matter is how easy/hard acquiring and using this information. Sener et al. (2019) focuses the separation of information sharing from information usage is valid, and the mediating role of usage is significant in improving operational effectiveness and efficiency. The increase in global competition and search for areas for performance improvement through integration and cooperation in order to meet customer requirements have led supply chain members to acquire as much information as possible and create a competitive advantage. Information is used towards the utilization of resources in a more efficient and effective way based on its characteristics such as strategical, tactical, or operational. When customer information is acquired,

the supplier receives an advantage and by using this information, an advantage is realized in terms of performance improvement.

Moreover, an electric circuit network and Kirchhoff's voltage and current laws, offer an alternative approach in determining the level of information so that each member of a supply chain has the same information power. Thus, the objective of this paper is to show how electric circuit dynamics can be used to illustrate supply chain information flow and equilibrium where each supply chain member actively utilizes the optimum level of scale and scope of information. The purpose is to introduce the elements of a conceptual framework that will be a base for information flow theory and to give direction for future studies.

In the next section, we compare and contrast supply chain networks with electric circuit networks followed by the investigation of relationships among the instruments of an electric circuit and supply chain network in section 3. Section 4 defines power in electric circuits and its implications for supply chain networks. Finally, section 5 provides a conclusion and future directions.

2. Type of Network Arrangement

In series circuits, components of a circuit are arranged in one path for current to flow. Following are the features of such an arrangement: The current flow is one-directional; Voltage flows in a single path; The magnitude of current that flows through each device is the same; Resistance level can vary for each device on the circuit; The circuit has to be in a continuous loop; Current stops flowing if any disruption occurs. On the other hand, components of parallel circuits are connected in a parallel way so that the current is separated through junction points. Following are the features of such an arrangement: The current could flow in more than one direction; Voltage flows in more than one path, and if any branch breaks down, other branches continue working without disruption; The magnitude of the current that flows depends on the resistance in each branch; Voltage is the same across all branches; Each branch works self-reliantly (independently). Thus, if the current stops flowing in a branch, the rest of the branches will continue working.

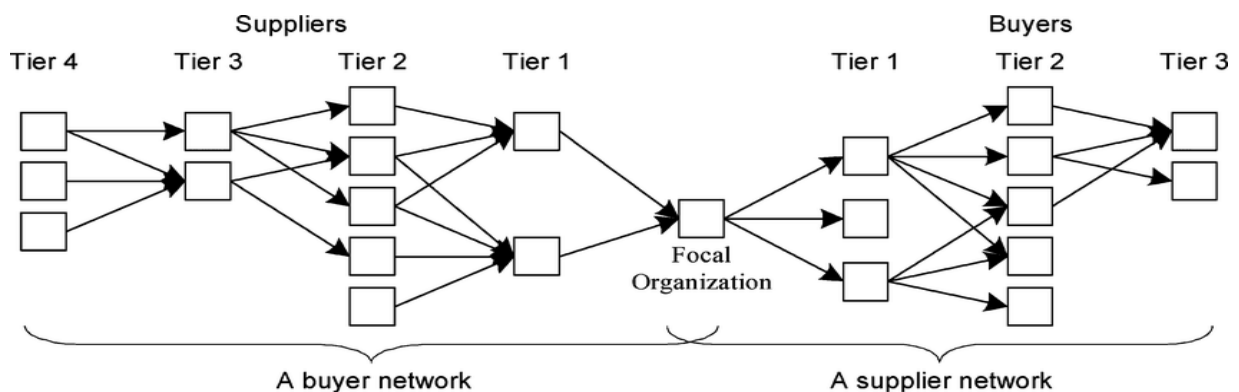


Figure 1: Sample Supply Chain Network Adopted from (Bi and Lin 2008)

When we consider supply chains, there could be many different network arrangements. However, these network designs can be categorized into series or parallel which are very similar to the circuit networks. An example of a series circuit network in supply chain management would be connecting

players from the ultimate supplier tier to the ultimate customer tier, with one player in each tier. The direction of information flow is one way, and if any disruption happens to a member of the network, the flow stops, and all other members are affected by this event. Considering parallel circuits, a supply chain network can be imagined with multiple suppliers, manufacturers, retailers, and customers that are connected with each other. Members of the chain can be clustered into branches. In this type of supply chain network information could flow more than one direction. In the case of a disruption, the supply chain won't be affected as a whole, and the remaining branches will remain operational.

3. Supply Chain Network versus Electric Circuit Network

We observe similar use of instruments between supply chain networks and electric circuit networks. First, information flow is very similar to electric current. Information flow is denoted by the amount of information shared within the supply chain network in a given period of time. Electric current is represented by the rate at which charge passes by a point on the circuit network in a given period of time. Second, an electric voltage is defined as the electric potential difference between two points in the circuit. Similarly, we can define information integration as the potential information between two nodes of a supply chain network. Third, electric resistance is depicted as the measure of the difficulty to pass an electric current through electric current nodes. Similarly, information intensity can be depicted as the friction against a motion of information in terms of the amount of flow and content among nodes in a supply chain network. On the other hand, we can also define information intensity as the inertia against a motion of information in terms of fluctuation, and time-based opposition. This is represented by an electric reactance in a circuit network that is inertia against the motion of electrons in an electric circuit network. Fourth and lastly, we define a load of information as the agility towards future unknowns, preventing information overload. Equivalence in a circuit network is an electric capacitor described as the opposition to changes in voltage by drawing or supplying current as they charge or discharge to the new voltage level. Table 1 summarizes these similarities and denotes each with corresponding notations.

Electric Circuit Network		Supply Chain Network	
Current, I	The rate at which charge passes by a point on the circuit network in a period of time.	→ Information Flow, F	Amount of shared and used information within the supply chain network in a period of time.
Voltage, V	The electric potential difference between two points in the circuit.	→ Knowledge Integration, I	Potential knowledge integration between two nodes in the supply chain network.
Resistance, R	The measure of the difficulty to pass an electric current through	→ Information Intensity, II	Friction against the motion of information. (In terms of amount, and content)

Reactance, X	Inertia against the motion of electrons	→	Information Intensity, II	Inertia against the motion of information. (in terms of fluctuation, time-based opposition)
Capacitor	Oppose changes in voltage by drawing or supplying current as they charge or discharge to the new voltage level.	→	A load of information, LI	Agility towards future unknowns. Prevents information overload.

Table 2. Information Types and Flows in a Supply Chain

Type of Information shared	Supplier	Manufacturer	Distributor	Customer
Demand	→	→	→	→
Capacity	→	→	→	→
Inventory	→	→	→	→
Logistics	→	→	→	→
Quality	→	→	→	→
Sales	→	→	→	→
Production	→	→	→	→
Shipping Errors	→	→	→	→
Cost	→	→	→	→
Cycle Time	→	→	→	→
Supply Disruptions	→	→	→	→
Performance Evaluation	→	→	→	→
Lead Time	→	→	→	→
Customer Satisfaction	→	→	→	→
Forecast	→	→	→	→
Availability	→	→	→	→
Conformance	→	→	→	→
Price	→	→	→	→
New Product	→	→	→	→
Customer Complaints	→	→	→	→
The direction of information flow:		→	Towards customer	
		←	Towards Supplier	
		↔	Both direction	

The main dissimilarity, however, is the bi-directional flow of information in the supply chain versus the one-directional flow of electrons in the circuit network. The direction of information can be identified according to the location of the members of the supply chain. Moreover, in a supply chain network, various types of information are moved or shared among supply chain members. The setting of the flow could be in different formats based on the supply chain network design a sample design is selected as Supplier/ Manufacturer/ Distributor/ Customer as shown in Table 2.

4. Relationship among the instruments

4.1. Ohm's Law

Ohm's law defines the relationships between current flow and voltage considering the magnitude of resistance. If there is a very high resistance for a given voltage, the current flow will be small. However,

in the presence of low resistance, the current flow will be higher. Thus, we can use the following formula: $I = \frac{V}{R}$ Where I is an electric current, V is Voltage, and R is resistance. We can further make the following assumptions about the relationship among resistance, current, and voltage: If we hold resistance fixed, then an increase or decrease in voltage means an increase or decrease in current respectively; If we hold current fixed, then an increase in resistance means an increase in voltage; If we hold voltage fixed, then an increase or decrease in current means a decrease or increase in resistance respectively. It is suggested that resistance is the easiest variable to control in a circuit network (Kuphaldt, 2006). This is also true for supply chain management information flow control. Companies may easily introduce a resistance barrier to sharing their financial information, for example.

4.2. Kirchhoff's Current Law

This law indicates that currents entering into any node in an electric circuit are equal to the currents flowing out. In other words, sum of all the currents entering and exiting a node must be equal to zero. For a given section of a circuit network, as seen in Figure 2, there are two entering currents into a node and two exiting the node. Thus, according to this law, the relationship can be shown as

$$\sum i = 0$$

Let's assume that currents entering a node as positive and currents exiting the node as negative.

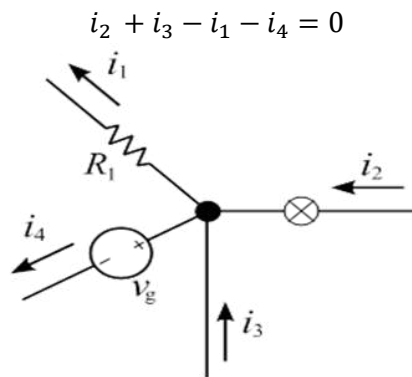


Figure 2. Representation of Kirchhoff's Current Law

Source: Kuphaldt, T. R. (2006)

Likewise, in the ideal case of a fully integrated supply chain network, the flow of information entering into any junction in a supply chain network is equal to the flow of information leaving the junction, so that sum of all the information flow entering and leaving a junction is equal to zero. In Figure 3, while information coming from the customer (c4) and from the manufacturer (m1) represents information flowing into the junction node, s2 and r3 represent information going towards the supplier, and retailer leaving the junction node.

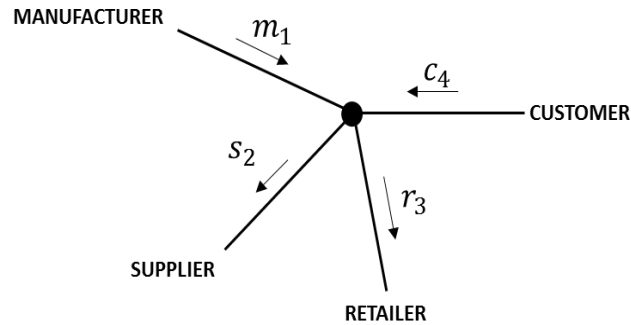


Figure 3. Representation of a simple node in a supply chain network

Recall from Table 2 that the direction of information flow can be in different types (demand, capacity, inventory, etc.) and move towards the customer, supplier, or either direction. In order to apply Kirchhoff's current law, different types of information in the same direction to a junction may be consolidated into a composite information flow.

4.3. Kirchhoff's Voltage Law

This law indicates that the voltage variations around any loop should equal zero. In other words, the sum of all the voltages in a loop must sum to zero. Regardless of the path, you choose in an electric circuit, if you come back to your starting point the voltage change in the loop must be zero. Thus, according to this law, based on the circuit network in Figure 4, the voltage law can be shown as

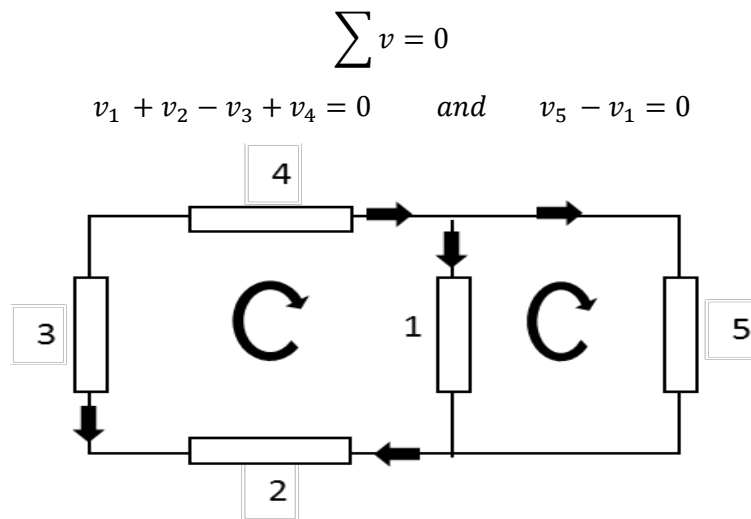


Figure 4. Kirchhoff's voltage law example

Source: Kuphaldt, T. R. (2006)

Similarly, in a fully integrated supply chain network, the potential knowledge variations around any tier (or loop) should equal zero. Information flows from high potential knowledge nodes to lower potential knowledge points. Equilibrium or stabilization is achieved when potential knowledge across the junction nodes is minimized. Let's consider a supply chain network, as in Figure 4, with two suppliers (S), a manufacturer (M), a distributor (D), and a customer (C). Sum of knowledge variations in each loop located on left and right side of the Figure 5 equal to zero.

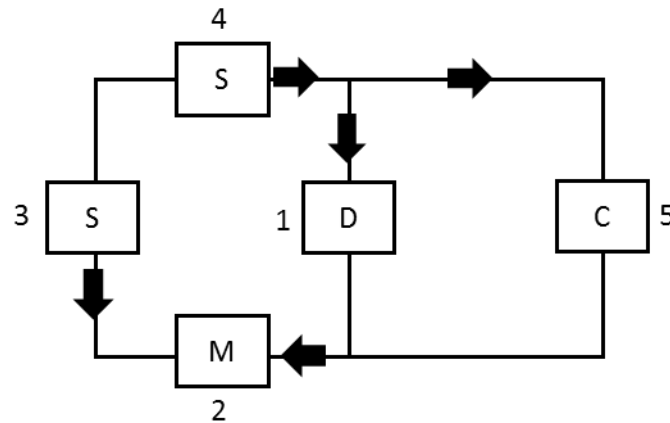


Figure 5. Representation of simple information flow in a supply chain network

4.4. Power

Power (P) in a circuit network is defined by the product of voltage and current, which is $P = V \cdot I$. In a supply chain sense, power is a function of information flow (sharing) and knowledge difference among the nodes and represents the rate of converting the information into other measures such as efficiency and effectiveness. Thus, sharing information alone does not constitute a power unless the knowledge of a player is enhanced by the use of information shared. This power measurement can be one of the important tools along with the other laws in analyzing supply chain networks from an information integration perspective. Determining the knowledge (voltage) and information flow rate (current) associated with a particular supply chain network for different information types and computing the power may provide a tremendous understanding of the integration state of the supply chain network.

5. Conclusion and Future Directions

This study aims to borrow the laws developed for electric circuit networks and to conceptually explore how such instruments can be used for a supply chain network. The main takeaway is that it is possible to develop a single measurement to represent the power of a supply chain information coupling.

This study offers important implications for practicing managers. It may allow managers to understand the impact of using the information on efficiency and effectiveness performance and map their company's integration power across to that of competitors in an industry. They may also be able to identify the knowledge lacking nodes or weak spots and offer mitigation options or strategies in order to improve the overall performance of the chain. Our study supports (Sener et al., 2019) and may help managers to identify members who share or hold information and who use it or don't in a supply chain network.

One should note that supply chain networks are much complex than direct current circuit networks. In supply chain networks, information may flow into multi-direction compared to that of a one-directional circuit network. Moreover, some tiers in the supply chain (such as the ultimate customer tier) may require one-directional information flow (demand information, for example). These characteristics suggest visiting the alternate current circuit network computations and utilize the

instruments of both direct current and alternate current circuit networks. We will further this research paper and develop corresponding mathematical models and use them in a simulation environment.

6. References

- Benton, W. C., & Maloni, M. (2005). The influence of power driven buyer/seller relationships on supply chain satisfaction. *Journal of Operations Management*, 23(1), 1-22.
- Bi, H. H., & Lin, D. K. (2008). RFID-enabled discovery of supply networks. *IEEE transactions on engineering management*, 56(1), 129-141.
- Flynn, B. B., Huo, B., & Zhao, X. (2010). The impact of supply chain integration on performance: A contingency and configuration approach. *Journal of operations management*, 28(1), 58-71.
- Ganesh, M., Raghunathan, S., & Rajendran, C. (2014). The value of information sharing in a multi-product, multi-level supply chain: Impact of product substitution, demand correlation, and partial information sharing. *Decision Support Systems*, 58, 79-94.
- Gilmour, P. (1999). A strategic audit framework to improve supply chain performance. *Journal of business & industrial marketing*, 14(5/6), 355-366.
- Govindan, K., Kaliyan, M., Kannan, D., & Haq, A. N. (2014). Barriers analysis for green supply chain management implementation in Indian industries using analytic hierarchy process. *International Journal of Production Economics*, 147, 555-568.
- Harland, C. M., Caldwell, N. D., Powell, P., & Zheng, J. (2007). Barriers to supply chain information integration: SMEs adrift of eLands. *Journal of Operations Management*, 25(6), 1234-1254.
- Huo, B., Qi, Y., Wang, Z., & Zhao, X. (2014). The impact of supply chain integration on firm performance: The moderating role of competitive strategy. *Supply Chain Management: An International Journal*, 19(4), 369-384.
- Kuphaldt, T. R. (2006). Lessons In Electric Circuits, Volume I—DC. *Vol. Fifth Edition. Open Book Project*.
- Mathiyazhagan, K., Govindan, K., NoorulHaq, A., & Geng, Y. (2013). An ISM approach for the barrier analysis in implementing green supply chain management. *Journal of Cleaner Production*, 47, 283-297.
- Mentzer, J. T., DeWitt, W., Keebler, J. S., Min, S., Nix, N. W., Smith, C. D., & Zacharia, Z. G. (2001). Defining supply chain management. *Journal of Business logistics*, 22(2), 1-25.
- Prajogo, D., & Olhager, J. (2012). Supply chain integration and performance: The effects of long-term relationships, information technology and sharing, and logistics integration. *International Journal of Production Economics*, 135(1), 514-522.
- Sener, A., Barut, M., Oztekin, A., Avcilar, M. Y., & Yildirim, M. B. (2019). The role of information usage in a retail supply chain: A causal data mining and analytical modeling approach. *Journal of Business Research*, 99, 87-104.
- Swink, M., Narasimhan, R., & Wang, C. (2007). Managing beyond the factory walls: effects of four types of strategic integration on manufacturing plant performance. *Journal of Operations Management*, 25(1), 148-164.
- Vickery, S. K., Jayaram, J., Droge, C., & Calantone, R. (2003). The effects of an integrative supply chain strategy on customer service and financial performance: an analysis of direct versus indirect relationships. *Journal of operations management*, 21(5), 523-539.
- Yu, W., Jacobs, M. A., Salisbury, W. D., & Enns, H. (2013). The effects of supply chain integration on customer satisfaction and financial performance: An organizational learning perspective. *International Journal of Production Economics*, 146(1), 346-358.
- Zhao, L., Huo, B., Sun, L., & Zhao, X. (2013). The impact of supply chain risk on supply chain integration and company performance: a global investigation. *Supply Chain Management: An International Journal*, 18(2), 115-131.

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Large Scale Analytics for Workload Segmentation

Bing Hu¹

Nicholas Mason¹

¹Intel Corp.

bing.hu@intel.com; nicholas.mason@intel.com

Abstract

As the complexity and size of workload landscapes continue to evolve, traditional tools that are useful for characterizing workloads have gradually become inadequate. Approaches such as scaling studies or isolated analyses of a single workload are gradually becoming insufficient to understand the broad state of the workload universe. To address this problem, we follow a machine learning methodology to leverage large numbers of workload experiments that already exist on Intel. We advanced a workload analysis platform that stores, manages, and facilitates the analysis of workload information using two big data analysis software tools, a top-down workload decision classifier and a workload universe mapping tool, to characterize large-scale workloads. Both analysis tools present a novel way to look at workload data from a higher level and consider these new characteristics at the workload segmentation level. We seek to achieve this goal in alignment with the strategic goals of developing new benchmarks, informing market sizing, and detecting emerging workloads.

Keywords: Workload Characterization, Clustering, Interpretable Classifier, Time-Phased Analysis, Machine Learning, Statistics, Benchmark Workload, Big Data.

1. Introduction

In recent decades, the digital industry has witnessed rapid growth. Along with the high demand for computing power and the growth in software and hardware, the workload landscape continues to evolve. In general, a computer workload includes an application or program and related data or information. Over the years, the complexity and size of workloads have grown exponentially, and the diversity of workload applications has expanded.

This rapid expansion is most apparent in studies on workloads across broad categories and applications. In the early 70's people started to classify workloads to improve system performance and design better computer architectures (Calzarossa, 2016). Many techniques have been adopted for workload characterization, such as statistical analyses. In recent years, there has been a great interest in computer workloads and systems that are best suited for running them (Brooks, 2015).

Given the rapid expansion of the universe of known workloads, analytical methods and paradigms through which workloads have been studied in the past also fall short of Big Data. In the past, considerable effort has been devoted to understanding and optimizing specific computer workloads. We argue that understanding a set of workloads differs from understanding a single workload. For instance, running structured scaling studies in a well-established suite of benchmarks, such as SPEC CPU, will provide insights that are targeted and detailed but will fail to highlight emerging workloads.

Similarly, considerable time has been devoted to the development of workload benchmarks; however, this approach has shortcomings. For example, it does not guarantee that new benchmarks will provide the desired coverage. At the front and center of these shortcomings, there is one main problem: these methods are difficult to scale in a way that keeps pace with the growth in the workload field. There is an urgent need to characterize the workload universe from a broader perspective.

To address these shortcomings, we propose a machine-learning (ML) methodology for characterizing workloads at the segmentation level. With this approach, mapping the entirety of the workload universe and representing it with a suite of “relevant workload types” can be characterized, measured, and determined. To achieve this goal, we rely on two main large-scale analysis tools: 1) top – down workload decision classifier (TDWDC), and 2) the workload universe mapping tool (WUMT). These tools are built within a centralized workload analytics platform (WAP) that facilitates the storage, retrieval, and sharing of a large amount of workload data. There are numerous benefits of this ML approach, such as Big Data, potential insights, and the impact of the result scale on the size of the dataset.

In the following sections, we describe the literature review and motivation for the adopted analytical approach: the workload analytics platform where the information resides, two proposed large-scale workload analysis software tools, and a description of the results and insights that could be derived from this breakthrough integration.

2. Literature review

Understanding the nature of workloads is crucial for workload characterization and system performance optimization (Brooks, 2015; Calzarossa, 2016). In the past, this meant that the main research stream focused on understanding workload characteristics to perform workload modeling and system architecture design. Many researchers have developed benchmarks based on the intrinsic properties of specialized workloads. HcBench is created for Hadoop clusters (Saletore 2013), vConsolidate (vCon) Benchmark is developed for virtual environments (El-Refaey, 2009), and SPEC CPU Benchmark is widely used for computer-intensive systems (Hassan, 2021; Henning, 2006).

However, these applications require surgical and specific views of workload. Few researchers have spent time studying workloads from a higher level seeking to understand and map the various types of workloads holistically. Some examples found in the literature include 3D workload subsetting research to characterize gaming workloads (George, 2015), regression-model-tree research to characterize the SPEC CPU2006 and SPEC OMP2001 Benchmark workloads (Ould-Ahmed-Vall, 2008), and a learning-based approach to predict the workload behavior of SPEC CPU benchmarks (Lozano, 2023). Unfortunately, these studies failed to simultaneously consider many workload types.

Within Intel Corporation, a key advancement in workload research is the Top-Down Microarchitecture Analysis (TMA) method (Yasin, 2014). This method leverages TMA metrics (TMAM) to identify performance bottlenecks in various applications (Yasin, 2019; Kanev, 2015; Roodi, 2018). Yasin (2014) sought to explain the characteristics of an individual workload at a given time, and allowed us to answer the following question: What percentage of this workload is constrained by factor X? However, this approach has not been used on a scale to answer the question of what percentage of a given collection of workloads is constrained by Factor X? Our research expands the TMA approach by

subjectively categorizing workloads to fulfil an expositional objective during the workload characterization process. We believe that by broadly using TMA metrics (TMAM) to define workload categories, we can obtain valuable insights into the behavior of workloads on a large scale.

We seek to address this research gap by leveraging thousands of workload experiments and the analytical tools developed by our team. Next, we describe the workload analysis platform (WAP) and two large-scale analysis tools built on it.

3. Workload analytics platform (WAP)

To effectively leverage the large amount of workload data generated by Intel, we must first provide a platform that is conducive to large-scale data sharing, storage, and analytics. The tool developed for this purpose is the Workload Analytics Platform (WAP). WAP 1) makes it easy for users to contribute and extract data under a common standard; 2) provides secure and scalable data management; and 3) provides analysis automation and support for large-scale studies.

We give special emphasis to the first point for two reasons: First, Intel workload data have historically been stored in non-standardized filesystems, which makes retrieval difficult. Second, there is no common method for describing and documenting workload experiment metadata such that data generated at different points in time (or by different teams) can be combined under a common framework.

Therefore, we worked extensively with experts in this field to understand the requirements for documenting workload metadata in a complete and extensible manner. We devised a data schema that allowed users to upload workload data in a semi-structured manner.

With this data standard, the platform will accept data from any team or collector that is compliant and then proceed to save all other relevant files, such as Event Monitor (EMON) files and run logs, into the WAP database. The data from workloads runs stored in WAP is denominated as an “experiment” and resides in the database until an analyst retrieves it to conduct detailed analysis on one or multiple workload-runs. At this point, the resulting analysis can be stored in the WAP database and denoted as a “study.” Moreover, the WAP will document relationships, settings, and results relevant to the study. With WAP, a single unified framework, we bring workload data at a scale within the reach of analysts, who may now leverage it at various scales and scopes (see Figure 1).

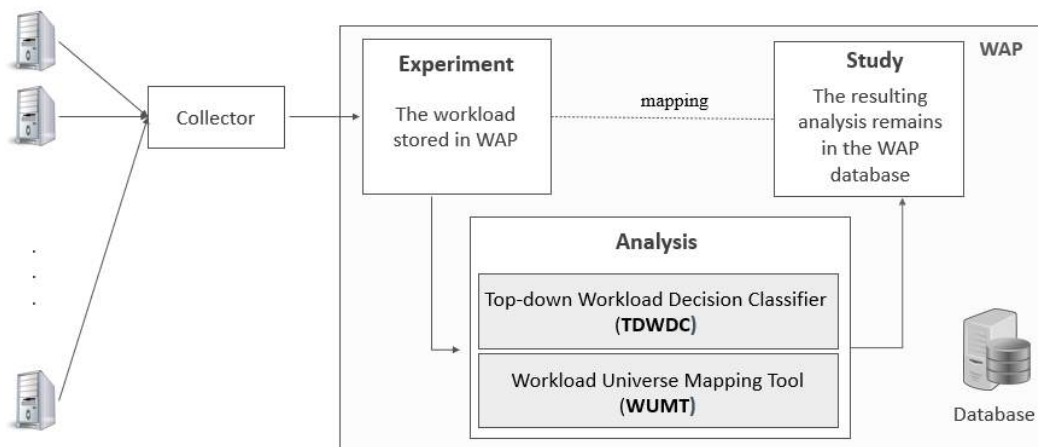


Figure 1. Workload analytics platform (WAP) architecture

4. Large scale analysis tools and results

4.1. Top-down workload decision classifier (TDWDC)

The top-down workload decision classifier (TDWDC) is an interpretable decision-tree classifier written in Python. It categorizes workloads in terms of the TMA performance categories. TMA categories are normalized percentages such that the sum of the subgroups (in percentage) from the same parent category is 100% (Yasin, 2014). For example, at level 1 (L1), the TMA metrics (TMAM) had four child groups: frontend, backend, bad speculation, and retiring categories. The sum of these four TMAM groups was equal to one.

TDWDC uses statistical analysis results for decision-making through the workload characterization process. It identifies a workload's corresponding performance category at each TMA metric level and uses only TMA metrics that are related to the performance category for the next level. The pseudocode is shown in Figure 2. The TDWDC algorithm first splits at level 1 (L1), based on the median value of the TMA metric instead of the average value. For example, if a workload has an outstanding front-end median for a given set of experiments at L1, then the workload is labeled with a front-end footprint. The overall performance of the classifier is secured by using the median, because the median is a popular analysis strategy that is not affected by outliers.

```

For Level 1 TMAM:
    max_bottleneck( $w$ ) = max { Frontend( $w$ ), Backend( $w$ ), Bad Speculation( $w$ ) }
    if max_bottleneck( $w$ ) ≥ median{Frontend} or median{Backend} or median{Bad speculation} then:
        classify to TMA category which is the largest one
    else:
        the workload  $w$  belongs to Retiring category

For Level 2, 3, and 4 TMAM:
    For all sub – groups from  $S_{L(t)}(w)$ 
        the workload  $w$  belongs to TMAM category  $t = \arg \max_t \{ S_{L(l,m)} \}$ 

Notation:
 $w$  : a workload, where  $w \in W$  and  $W$  is a set of workloads
 $t$  : a TMAM category
 $m$  : a TMAM metric, where  $m \in M$  and  $M$  is a set of all TMAM metrics
 $S_{L(l,m)}$  : a set of TMAM metrics at level  $l$  whose parent is  $l - 1$ ; where  $l \in \{2, 3, 4\}$ ,  $m \in M$ 

```

Figure 2. Pseudo code for top-down workload decision classifier (TDWDC)

We tested our classifier using an entire set of 29 experiments from the SPEC CPU2006 Benchmark suite. These workload experiments were collected from a system with 32 Cores, 384GB RAM, and a clock frequency of 2500 MHz. The results are illustrated in Figure 3 and 4. We only displayed Level 1 (L1), Level 2 (L2), and Level 3 (L3) TMA metrics for better visualization. From this figure, we can see that our computational method automatically segments the workloads into subcategories according to the importance of each TMA metric for each group and TMA metric level. This result is consistent with the manual categorization used in Yasin's study (Yasin, 2014).

In Figure 4, we can see that the selected experiments cover level 1 (L1) system performance components. Hence, this set of benchmark workloads can be used to evaluate the system performance from the perspective of the backend, frontend, and bad speculation TMA metrics. However, this set of

workloads does not provide sufficient coverage for all level 2 (L2) and level 3 (L3) system components.

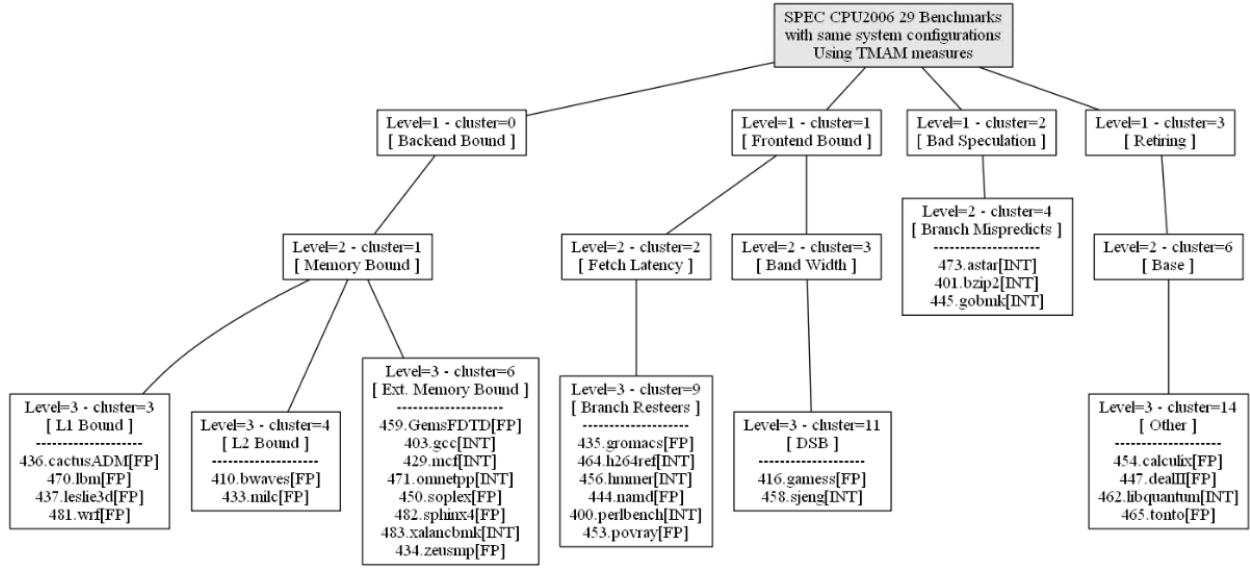


Figure 3. Workload characterization result for SPEC CPU 2006 with TMA metrics

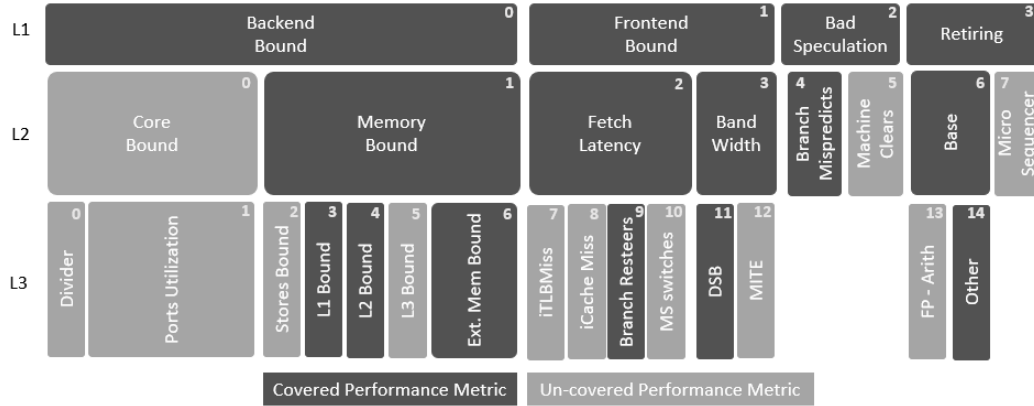


Figure 4. Coverage analysis result for SPEC CPU 2006 with TMA metrics

4.2. Workload universe mapping tool (WUMT)

While the methodology of TDWDC is insightful and novel, there are other ways in which the data can be analyzed, which will present additional, more granular interpretations of the workload data. WUMT is a software tool (written in Python) that seeks to discover these insights by providing a workload time-phased data analysis tool, a workload clustering engine, and workload data visualizations.

These analysis tools are implemented as stages in a workflow and can be run separately or in conjunction depending on the desired analysis scope. Fig. 5 shows the workflow of the full analysis. The four-step arrows illustrate the rationale for implementing the WUMT. In Step 1, WUMT determines the workload phases for each raw workload data file. In step 2, the WUMT computes the statistics based on the workload phases. In step 3, WUMT performs clustering of a given set of workloads using the workload phase statistic values. Finally, the WUMT displays the data visualization results in Step 4.

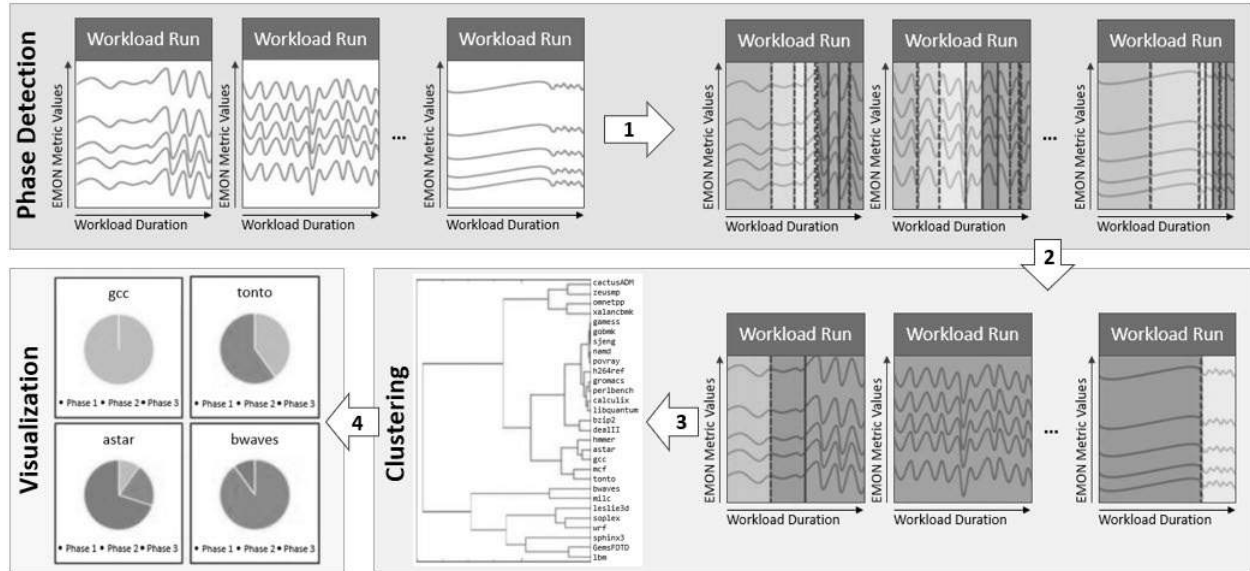


Figure 5. WUMT analysis components and workflow

We take some special focused on the clustering stage of WUMT. As articulated earlier in this document, it is not the norm for Intel to examine workloads in a general manner, making comparisons across different workload types and system configurations. Therefore, our methodology makes no assumptions regarding the nature of the workloads used and can cluster any arbitrary set of workload experiments.

Once classified, inferences can be made regarding the structure of the space occupied by the workloads. At this point, a good data visualization engine is required to allow us to retrieve results from the clustering algorithm and to represent workloads in a familiar and easily interpretable manner. Finally, the time-phased data analysis tool allows us to obtain a more granular interpretation of workloads, whose behavior may be highly variable and time-dependent. This last addition to the workflow is a valuable improvement over former methods which use the “median” TMA metric values across the whole execution to classify workloads. Using time-phased analysis, we can better capture the dynamic nature of the key workloads.

In addition to the benefits of implementing these decision tools as a workflow, the loose coupling of tools and their design as separate packages provides increased flexibility for creating valuable analyses. For example, we can choose to run a phase analysis to separate various workloads into smaller segments and then cluster them, as shown in Figure 5. Alternatively, we can choose not to run the phase-analysis stage and consider each workload as a single unit. For the purposes of this study, we skip the phase-detection stage entirely when we present the results below.

To illustrate the results of WUMT, we begin by representing the workload universe for the same set of 29 microbenchmarks used in the previous sections. These workloads were clustered using data from their top-level (L1) and second-level (L2) TMA metrics, and the results were visualized using the same set of metrics.

The results of the analysis are shown in Figure 6, where we show an overall representation of the workloads using the Principal Component Analysis (PCA) transformation of their TMA metric data, as

well as an example of two of the categories represented via radar diagrams. The five categories can be broadly described as follows: 1) memory-bound, 2) heavy retiring and bad speculation, 3) heavy retiring and memory-bound, 4) well-rounded, and 5) retiring.

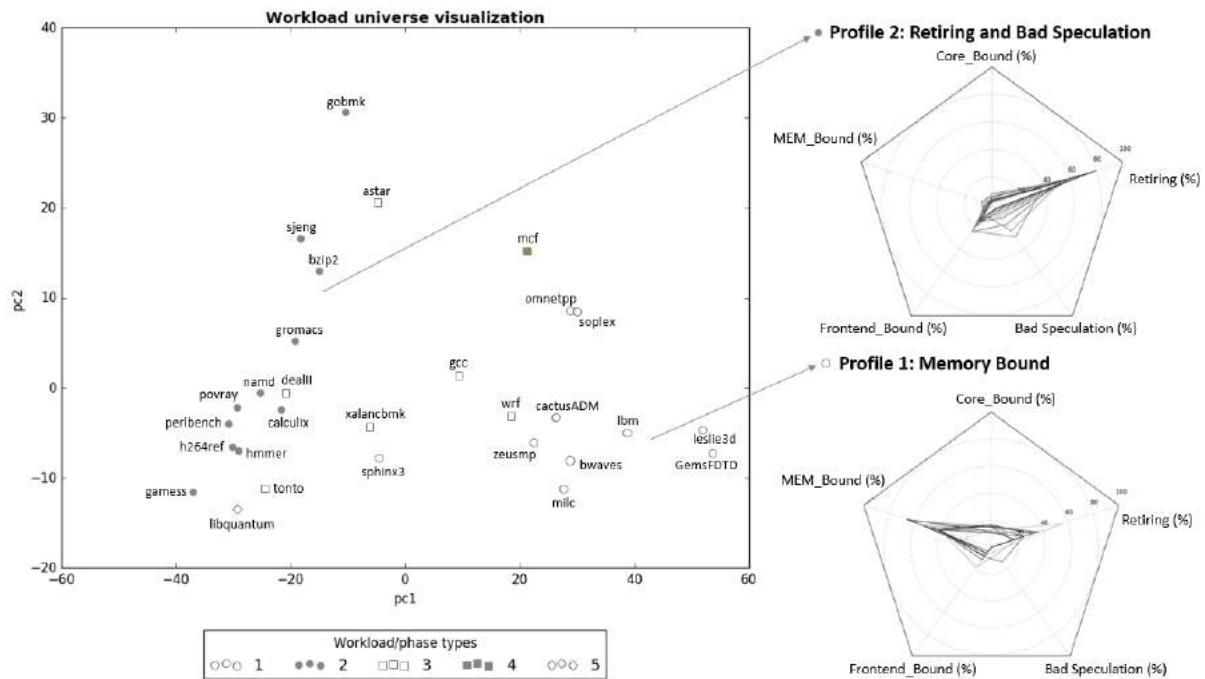


Figure 6. Workload category profiling from the automated clustering approach

We briefly examined the results by focusing on the memory-bound TMA metric. While our results are consistent with previous observations that group the milc, lbm, and omnetpp workloads together as memory bound, we note that the mcf workload is separated from them into its own cluster, suggesting that, even though these four are comparable if we look only at the “memory bound” aspect of the workload, they are sufficiently different to warrant separation into their own categories.

Applying this analysis to the SPEC CPU2006 Benchmark suite would be interesting and informative from a research perspective. However, we argue that the true potential of this analysis can be realized when scaled up to hundreds or thousands of experiments, across several workload types and system configurations. The combination of a larger scale and a tool for high-level visualization will allow us to realize the full benefits of this promising research area.

5. Conclusion

In this paper, we present a new way of thinking about workload analytics, looking at the universe of existing workloads from a high level and seek to derive high-level insights from the top-down approach. To achieve this, we relied on existing methods and results that provide a foundation for our analysis, such as the TMAM representation of the workload (Yasin, 2014). Our analysis is not possible without the existence of a Workload Analytics Platform (WAP), which is the main avenue for exploring new methodologies and scaling them to a level where all existing data on Intel can be used.

We present two novel methods for examining workload data at the segment level. These new methods allow us to derive insights that are critical to Intel’s long-term strategy but are still uncommon

in the workload research space. We hope that these new data analysis methods will help Intel to 1) develop better benchmarks with better coverage of all relevant system bottlenecks, 2) understand the size and scope of different workload types across the workload landscape, and 3) generate insights from the dynamic time-phased nature of workloads. These outcomes go hand in hand with scaling up our study; therefore, we look forward to expanding the workload library contained in WAP. We believe that these outcomes will increase Intel's competitive standing and its long-term strategic advantages.

6. Acknowledgment

The authors would like to thank Dr. Karl Kempf for his insightful comments and unwavering support of the foundational research required for this study. We thank the Intel Decision Engineering team for their contribution to this project. We also thank Andy Anderson and Sundar Iyengar for providing valuable data and expertise throughout the project lifetime.

7. Reference

- Brooks, D. (2015, October). Keynote I: Workload Characterization in the Era of Specialization. *2015 IEEE International Symposium on Workload Characterization (IISWC)*, xii.
- Calzarossa, M. C., Massari, L., and Tessera, D. (2016). Workload characterization: a survey revisited. *ACM Computing Surveys (CSUR)*, 48(3), 1-43.
- El-Refaey, M. A., & Rizkaa, M. A. (2009, June). Virtual system workload characterization: An overview. In *2009 18th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises* (pp. 72-77). IEEE.
- George, V. M. (2015, October). 3D workload subsetting for GPU architecture pathfinding. In *2015 IEEE International Symposium on Workload Characterization (IISWC)*, (pp. 130-139). IEEE.
- Hassan, M., Park, C. H., Black-Schaffer, D. (2021). A Reusable Characterization of the Memory System Behavior of SPEC2017 and SPEC2006. *ACM transactions on architecture and code optimization*, 18(2), 1-20.
- Henning, J. L. (2006). SPEC CPU2006 Benchmark Descriptions. *ACM SIGARCH Computer Architecture News*, 34(4), 1-17.
- Kanev, S., Darago, J. P., Hazelwood, K., Ranganathan, P., Moseley, T., Wei, G., & Brooks D. (2015, June). Profiling a warehouse-scale computer. *2015 ACM/IEEE 42nd Annual International Symposium on Computer Architecture (ISCA)*, (pp. 158-169).
- Lozano, E. S. A., Gerstlauer, A. (2023). Learning-based Phase-aware Multi-core CPU Workload Forecasting. *ACM transactions on design automation of electronic systems*, 28 (2), 1-27.
- Ould-Ahmed-Vall, E., Doshi, K. A., Yount, C., & Woodlee, J. (2008, April). Characterization of SPEC CPU2006 and SPEC OMP2001: Regression models and their Transferability. *ISPASS 2008*, (pp. 179-190). IEEE.
- Roodi, M., Moshovos, A. (2018, September). Gene sequencing: where time goes. *2018 IEEE International Symposium on Workload Characterization (IISWC)*, (pp. 84-85). IEEE.
- Saletore, V. A., Krishnam, K., Viswanathan, V., & Tolentino, M. E. (2013, September) HcBench: Methodology, development, and characterization of a customer usage representative big data/Hadoop benchmark. *2013 IEEE International Symposium on Workload Characterization (IISWC)* (pp. 77-86). IEEE.
- Yasin, A. (2014). A top-down method for performance analysis and counters architecture. *2014 IEEE International Symposium on performance Analysis of Systems and Software (ISPASS)* (pp. 35-44). IEEE.
- Yasin, A., Ben-Asher, Y., & Mendelson, A. (2014, October). Deep-dive analysis of the data analytics workload in cloud Suite. In *2014 IEEE International Symposium on Workload Characterization (IISWC)* (pp. 202-211). IEEE.
- Yasin, A., Haj-Yahya, J., Ben-Asher, Y., Mendelson, A. (2019). A Metric-Guided Method for Discovering Impactful Features and Architectural Insights for Skylake-Based Processors. *ACM transactions on architecture and code optimization*, 16(4), 1-25.

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Service Quality Between Tourism and Pilgrimage: A Literature Review

Majid Alshaibi¹

Haitham Bahaitham²

Ahmad Elshennawy¹

¹*UCF Department of Industrial Engineering and Management Systems*

²*Department of Industrial and Systems Engineering, College of Engineering, University of Jeddah, Saudi Arabia*

ma550940@ucf.edu; hbahaitham@uj.edu.sa; ahmad.elshennawy@ucf.edu

Abstract

Service quality is a key success factor in rapidly developing markets that crucially acquire customer satisfaction and retention. In this study, a systematic literature review following the PRISMA review protocol addresses service quality approaches and models. Following the service quality model's conceptualization and dimensionality, the models' applications in the tourism industry, in general, have been covered while shedding light on their applications in the Hajj event, one of the world's largest annual massive gatherings in Saudi Arabia. The outcomes of this effort are aimed at developing a novel framework with standardized, relatively comprehensive dimensions that suit Hajj service clusters and assimilated stakeholders.

Keywords: Service Quality, Tourism, Religious Tourism, pilgrimage, Hajj

1. Introduction

The Hajj ritual is the fifth pillar of the Islamic faith, and it is considered one of the world's largest annual massive gatherings, with the participation of pilgrims from more than 185 countries and figures of participants reaching more than three million. It is mandatory once in a lifetime for financially and physically capable Muslims and must be performed in specific geographical areas in the Province of Makkah, Saudi Arabia, and over specific days of the last month of the Islamic lunar calendar. The Hajj pilgrimage stretch out between four zones: The Grand Mosque (Al-Masjid Alharam), Mina, Arafat, and Muzdalifah respectively with 13.9 miles (22.4 Km) length and area of 10 square miles (26 square km) (Abdo, 2022). Demand for Hajj has increased significantly. According to Weeks (2020), Muslim's shape 25% of the world, forming about 1.8 billion. Pew Research Center (2019) projected the Muslims' population growth to reach 2.98 billion by 2060, shaping 30 % of the global population.

The government of Saudi Arabia pays a great deal of attention to delivering a safe and pleasant experience to Hajj pilgrims. In 2019, The Custodian of the Two Holy Mosques, King Salman bin Abdul-Aziz Al Saud, inaugurated the pilgrims Service DARP (Dyof Al-Rahman Program) one of the Kingdom's Vision 2030 programs. The program has three main strategic objectives: facilitating more pilgrims to attend Hajj with ease of access to information about the experience and pleasant arrival to the holy cities of Makkah or Medina; delivering high-quality services to the participating pilgrims, including accommodation, subsistence, transportation, etc.; and enriching the pilgrims' spiritual and cultural

experiences through visits and events related to the heritage monuments of Islamic history (Doyof Al Rahman Program, 2019). Thus, improving Hajj service quality, one of the program objectives, represents a cornerstone for fulfilling other program objectives. This research activity aims to systematically explore the literature in similar massive gathering settings in the tourism industry while narrowing down to Hajj service quality to outline a set of comprehensive service quality dimensions that suit the complexity of the event and the diversity of the stakeholders included. Mapping Hajj service quality dimensions will aid in the development of a robust framework to measure service quality and generate action plans for improvement.

2. Methods

This literature review followed the systematic literature review protocol “Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).” The search scope was to investigate service quality approaches and model conceptualization and how they are applied in the Hajj pilgrimage. The search approach used related keywords interchangeably in the five databases. In addition, a narrative literature review approach was used to complement the areas that require further coverage. Exclusion criteria were research scope articles, observational design studies, opinion articles, publications older than 20 years (except theoretical foundations), and non-English language publications. On the other hand, the inclusion criteria were peer-reviewed quantitative and qualitative articles, books, and English-authored studies. To this point, the literature lacks a Systematic Literature Review (SLR) for service quality in Hajj, and the goal of this study is to investigate service quality models and dimensionality, and what has been done specifically in Hajj. This will exclude reviews on tourism and religious tourism for the sake of document size criteria.

3. Results

3.1. Service Quality

Attraction toward service quality has been fueled by the massive growth of service sectors, which has affected the world economy (Prakash & Mohanty, 2013). A recent study in the United States showed that from 2009 to 2019, more than 78.65% of the nation’s workforce was represented by service industries (Plecher, 2020). Service quality affects company profitability, improves operational performance, increases customer satisfaction, and secures customer loyalty (Chang, Wang, & Yang, 2009; Hallowell, 1996; Kunst & Lemmink, 2000; Sivadas & Baker-Prewitt, 2000). In addition, organizations delivering high-quality services have higher market segments and investment profits than organizations delivering lower-quality services (Ghobadian, Speller, & Jones, 1994).

Services have various characteristics that differentiate them from physical products. Such characteristics can be mapped in literature as intangibility, inseparability, perishability, and heterogeneity (Zeithaml, Parasuraman, & Berry, 1985). These characteristics are commonly relied upon as a source for developing commercialization strategies and realizing service consumers behavior (Wolak, Kalafatis, & Harris, 1998). Despite the popularity gained by these characteristics, in literature, to conceptualize service, some researchers view them as inadequate images of service and limit service concept generalizability (Edgett & Parkinson, 1993). As an effort to overcome this limitation, service has been conceptualized in literature based on service complexity and consumer level of involvement

(Prakash & Mohanty, 2013). Nonetheless, the literature shows that service and service quality are viewed differently, and they are challenging to define and measure (Some of the definitions found in literature are presented in Table1). As a result, scholars have created various service quality models based on different service quality dimensions, and the debate continues to focus on the most suitable definitions, models, and dimensions. Based on the variation of the definitions, and despite the thorough study of service quality observed in literature, it is considered as an obscure construct that is challenging to define, model, and scale (Crosby, 1979; Reeves & Bednar, 1994; Parasuraman, Zeithaml, & Berry, 1985; Prakash & Mohanty, 2013). With all the attention given to service quality from service providers and researchers, research on the topic has not reached maturity. Service organizations are lagging behind manufacturing organizations in implementing quality approaches, such as total quality management and quality assurance (Ghobadian et al., 1994). Stated differently, the manufacturing sector leads in quality management attributes, and the service sector leads in customer voice understanding and satisfaction (Sun, 2001).

Table 1. Service Quality Definitions

No.	Reference	Service Quality Definition
1	Gronroos (1984)	"The outcome of an evaluation process where the customers compare their expectations with service they have received".
2	Parasuraman et al. (1985)	"The discrepancy between consumers' perceptions of services offered by a particular firm and their expectations about firms offering such services".
3	Bitner and Hubbert (1994)	"The consumer's overall impression of the relative inferiority/superiority of the organization and its services".
4	Ghobadian et al. (1994)	"The level of service offered to meet customer expectations".

In this study, summaries of twenty-nine service quality models are chronologically listed in Table 2, with the understanding that no generic model fits all types of services to date. Several of the reviewed models were retrieved or modified from the gap model and SERVQUAL instrument, while others were IT-based for electronic-based services. Seth et al. (2005) assessed 19 models and reported that service quality measurement and results rely on service type, time of service, demand, and setting, which adds to the challenging nature of the context. In addition, consumer expectations for a certain service -fluctuate for the same factors, and there is a need to validate and modify the existing service quality models. Ghotbabadi et al. (2012) reviewed four service quality models: the Nordic model, SERVQUAL model, multilevel model, and hierarchical model, limiting the study to the most acceptable and used models by researchers. The study concluded that the SERVQUAL model is widely accepted among researchers, despite the lack of convincing support for all service applications. The hierarchical model is a representation of overcoming SERVQUAL shortcomings or other models through modifications to assess service quality. Another effort by Jain and Aggarwal (2015) mapped 16 models and concluded that none of the service quality models is perfectly applicable to all service sectors and cultural settings. A recent study agreed on the absence of consensus among scholars' definitions, factors, and indicators of service quality in the health care industry. This study focused on two practical models that have been commonly used and modified in the literature, SERVQUAL and SERVPERF, and recommended the customization of

models to cover the aspects of required services (Endeshaw, 2019). There are distinctions between online electronic services or e-service models and in person services models in literature. Beyond 2010, electronic service quality models have shown growth among researchers in literature compared to conventional service quality models.

Table 2. Service Quality Models

	Reference	Model Description	Notes
1	Lehtinen & Lehtinen (1982,1991)	Initially stated three dimensions for service quality physical quality, interactive quality, and corporate quality followed later by a model.	Limited application to restaurants service.
2	Gronroos (1984)	The Nordic model is the first SQ model with three dimensions technical quality, functional quality, and image quality.	Lack of a research instrument and implication.
3	Parasuraman et al. (1985,1988, 1991,1994)	Gap model: measures service quality as the difference between consumer expectations and perception of service using SERVQUAL instrument which started with ten dimensions then reduced to five dimensions namely tangibles, reliability, responsiveness, assurance, and empathy.	Ignited research in service quality and led cumulative efforts in the development of various models and derivatives.
4	Haywood-Farmer (1988)	Attribute service quality with no directions to management to identify service shortcomings.	Lack of a research instrument (Conceptual model).
5	Brogowicz et al. (1990)	Synthesized model of service quality model considering traditional marketing, external influences, and corporate image that affect technical and functional consumer expectations.	Modified from SERVQUAL scale and lack empirical evidence (Conceptual model).
6	Cronin and Taylor (1992,1994)	Applied perception only as predictor for service quality and developed SERVPERF performance only scale, an empirically proven common tool in literature.	Modified from SERVQUAL.
7	Mattsson (1992)	Ideal value model of service quality arguing value as an outcome of satisfaction based on ideal perceived standard and experienced outcomes.	Lack the adoption in literature.
8	Teas (1993)	Normed quality and evaluated performance model.	Modified from SERVQUAL scale with limited sample and validity.
9	Berkley and Gupta (1994)	IT alignment model that connects the service with the organization IT strategies.	Lack of measurement tool or research instrument
10	Rust and Oliver (1994)	Three component model a service quality three dimensions conceptualization includes service product, service delivery, and service environment.	Limited to dimensionality conceptualization only.

11	Dabholkar (1996)	Attribute and overall affect model for self-service technology-based services consist of two alternatives where attribute model capture cognitive expectation of the service and overall affect model capture the costumer feeling about the use of the self-service technology.	Limited application and setting to fit all sorts of self-services.
12	Dabholkar et al. (1996)	The multilevel model of service quality consists of three stages of overall retail service quality, primary dimensions, and subdimensions for retail stores. The model brought the idea that retail service includes products quality and not a pure service.	Modified from SERVQUAL scale with five dimensions and twenty-eight items.
13	Spreng and Mackoy (1996)	Model of perceived service quality and satisfaction adapted from Oliver (1993) students' satisfaction conceptual model to cover expectations, performance, and satisfaction.	Applied to educational services and measured by ten dimensions.
14	Philip and Hazlett (1997)	PCP attribute model a hierarchical structure based on peripheral, core, and pivotal attributes.	lacks dimensions to the stated attributes and empirical validity (Conceptual model).
15	Sweeney et al. (1997)	Retail service quality and value model with two alternatives, one influences value perceptions and the other influences customer's willingness to buy.	Value is limited to value of money in the model.
16	Oh (1999)	Service quality, customer value, and customer satisfaction model developed to assess consumer decision process and organization performance.	Lack generalization and has few items to measure in the model dimensions.
17	Mei et al. (1999)	HOLSERV model a developed model to address hospitality industry with three dimensions employees, tangibles, and reliability.	Lack generalization and validity.
18	Dabholkar et al. (2000)	Antecedents and mediator model a comprehensive model link service quality antecedents defined as reliability, personal attention, comfort, and features with customer satisfaction and behavioral intentions.	Modified from SERVQUAL.
19	Bahia and Nantel (2000)	The Banking Service Quality (BSQ) model includes seven dimensions: reliability, service portfolios, tangibles, access, effectiveness, price, and assurance represented by 31 items.	Based on literature and experts' opinion and lack validation.
20	Frost and Kumar (2000)	Internal service quality model focuses on internal customer or contractors to realize internal service quality.	Based on SERVQUAL and external effect is not considered in the model.
21	Soteriou and Stavrinides (2000)	Data Envelop Analysis (DEA) Internal service quality model designed for banking service to ideally utilize resources and point to the factors not utilized.	SERVPREF based with perception only and does not measure bank service performance.
22	Brady and Cronin	Hierarchical model with three dimensions interaction quality, physical environment quality, and outcome	A explanatory model with growing adaption

	(2001)	quality where each has three sub dimensions, the model combines four models.	in literature.
23	Broderick et al. (2002)	Online banking model an electronic service model	A conceptual model lacks validation.
24	Zhu et al. (2002)	IT-based model an electronic service quality model that focus on experience control and ease of use.	Lacks measure (conceptual model).
25	Santos (2003)	Model e-service quality an electronic service model.	Lacks scale and empirical validity.
26	Long and McMellon (2004)	Online shopping hierarchical electronic model with an addition of geographic distance and impersonal customer experience dimensions.	Focus more on the technological attributes of the online service.
27	Svenson (2004)	The Chronological service quality model is based on performance, context, and time with six dimensions.	Lacks application in literature.
28	Parasuraman et al. (2005)	The electronic service quality multi-items E-S-QUAL & E-RecS-Qual assessment model introduced by the scholars who developed the Gap model and SERVQUAL scale.	consists of total of seven dimensions.
29	Shahin and Samea (2010)	Comprehensive expanded Gap service quality model includes five new components and eight additional gaps.	Lacks scale and validity.

3.2. Service Quality in Hajj

Service quality in Hajj research is relatively limited in literature and is considered to be in the exploration phase. More research has addressed Umrah (a non-mandatory short form of pilgrimage that can be performed around the year) than Hajj, and in this review, only the research on Hajj has been included. Hajj includes physically exhausting activities, and the Hajj management's main concern is the safety of pilgrims. Effectively managing Hajj activities consider the clusters of services with complex processes and intersecting stakeholders (Al Jahdali, 2021; Hassan, Abdou, Abdelmoaty, Nor-El-Deen, & Salem, 2022; Kadi & Selim, 2022). To accommodate the assessment of the services offered in such gatherings, the dimensionality of the model that addresses Hajj service quality should be selected from the literature before being validated using feasible validation techniques.

One of the early studies on Hajj service quality was introduced by Jabnoun (2003), who developed a HAJQUAL instrument based on the SERVQUAL model. The HAJQUAL resulting dimensions after factor analysis are human service, Makkah accommodation, Mina-Arafa, accessibility, bathroom accessibility, bathroom cleanliness, and accommodation outside Makkah with questions related to the original dimensions of SERVQUAL and a new dimension named accessibility. Haq and Jackson (2009) qualitatively researched service perception and pilgrims' behavioral patterns by comparing two groups of pilgrims in Australia and Pakistan. The results showed more critical responses in interviews with Australian pilgrims due to their culture of individualism, criticism attitude, and authority openness. A study by Al-Hoqail et al. (2010) addressed ambulatory health services in Hajj, part of a major stakeholder in pilgrimage healthcare, and considerable related dimensions of Hajj services. One influential study in the context of Hajj service quality was conducted by Alsharief in 2005, where the empirical results were published by Eid (2012), followed by a book by Alsharief and El-Gohary (2017). This research covered pilgrims from countries with the highest percentage of participants from Asian, African, European, and North American cotenants. Specifically, the five countries selected were Indonesia, Egypt, the United

Kingdom, the United States, and Saudi Arabia, with a total sample of 934 pilgrims. The research applied the original SERVQUAL original instrument complemented by semi-structured interviews conducted at the port of departure. Darfoon (2013) added to the context with mixed method research, where he examined US Hajj agencies' service quality, value, and pilgrims' satisfaction. Similarly, Peck (2013) developed a framework to clarify the service delivery elements, service processes, and challenges under the South African Hajj and Umrah Council (SAHUC). In 2017, Arasli et al. realized the gap in Hajj quality research and used semi-structured interviews among pilgrims, Hajj agents, and coordinators to explore Hajj service quality dimensions. Content analysis revealed that the concluded six dimensions were procedural, accommodation, guidance, external and internal transportation, assistantship, and health services. Post the pandemic of Covid-19, Hassan et al. (2022) investigated pilgrims' experiences with Hajj services, addressing a domestic sample of 216 pilgrims in the Hajj season of 2021. The researchers developed an instrument based on the literature with five service subdimensions food, accommodation, religious guidance, transportation, and medical services to explore pilgrims' experiences. Table 3 shows the studies reviewed on Hajj service quality and the most recent publications in the context.

Table 3. Hajj Service Quality Studies and Dimensionality

No.	Reference	Model Dimensions	Notes
1	Jabnoun (2003)	Human services, Makkah accommodation, Makkah Accommodation, Mina-Arafa, Accessibility, Bathroom accessibility, Bathroom cleanness, and Accommodation Outside Makkah.	HAJQUAL model with 44 items and sample of 110 pilgrims in season 2002.
2	Haq and Jackson (2009)	Semi structured qualitative interviews to compare the cultural effect on Hajj perceived service by comparing Australian pilgrims' responses with Pakistani pilgrims.	Studying pilgrims' cultural behavior and variation associated in season 2005.
3	Al-Hoqail et al. (2010)	The study addressed ambulatory health services in Hajj with three dimensions health facility, physicians, allied medical services.	An instrument consist of 15 items and 478 sample in season 2008.
4	Eid (2012)	SERVQUAL scale conventional dimensions tangibles, reliability, assurance, empathy, and responsiveness to address Hajj service provider (HSP) generically.	The research data and results are from Dr.Alsharief book published 2017.
5	Darfoon (2013)	SERVPREF based scale developed from semi-structured interviews and literature review to include transportation services and accessibility, accommodation services, Hajj tour guide services, food services, money value, and overall satisfaction dimensions.	A scale of 29 items and sample size of 183 US pilgrims in season 2012.
6	Peck (2013)	The study aims to reveal service quality components, processes, and challenges with five dimensions namely Hajj agents' services, Airlines, SAHUC regulations, Ministry of Hajj regulations, and pilgrims' preparation.	Research instrument includes 22 items, sample of 326 SAHUC pilgrims in seasons 2006 & 2007.
7	Alsharief and El-	The book reviewed the context extensively using	Total of 44 items and a

	Gohary (2017)	Gap model's SERVQUAL conventional scale and open-ended interviews. Proposed a model of integrated Hajj service excellence. The research data and results published by Eid (2012) paper reviewed previously.	sample of 934 pilgrims from five different countries each represent a continent and the host country in Hajj season 2004.
8	Robbany (2018)	Applied SERVQUAL five dimensions with modified questions.	A scale of 29 items and a sample of 95 Indonesian pilgrims.
9	Hassan et al. (2022)	The research addressed accommodation, food, medical, religious guidance, transportation, and spiritual experience.	Instrument of 39 items, multinational domestic sample of 216 in Hajj season 2021.

4. Conclusion and Discussion

Among the reviewed service quality models, the gap model instrument SERVQUAL showed domination and support in the academic field and practical implications. One of the main reasons behind this is the model's capacity for modification to fit the application, which resulted in having ten of the addressed models or conceptual models as SERVQUAL derivatives or somehow based on it. For instance, SERVPRE is a highly anticipated instruments based on measuring service quality performance instead of applying the two measures of service expectation before the service and service perception after consuming the service and measuring service quality from the disconfirmation model. Both scales were empirically tested, and the performance can only avoid the researchers the challenging nature of addressing the same respondent twice. The SERVQUAL scholars in 1994 publication confirmed the validity of counting on perception of service only. Flexibility in implementation to fit the context of the research and intuitiveness of the instrument are among the reasons for adoption among researchers.

The dimensionality of service quality has been thoroughly discussed in the literature, and there is agreement that there are no universal dimensions for service quality, and each study should be adjusted to fit a service context. The gap model originally introduced ten service quality dimensions (assurance, communication, credibility, competence, courtesy, empathy reliability, responsiveness, security, and tangibles) trying to hold arms in service applications. However, the dimensions were reduced to five to fit service providers, such as bank branches, automotive service units, or a telecommunication service provider, with fewer items to be brief and manageable by respondents. Adding variables to the service quality dimensions is very common in literature depending on the service application.

In the Hajj service quality context, relatively limited contributions found in the literature and the research could add to quality improvement efforts with a novel framework. The dimensionality in Hajj and similar settings should include all stakeholders as dimensions, and under each dimension, sub-dimensions can be placed to diagnose the significant areas in the pilgrims' experience. From the conducted review, the dimensions considered in Hajj service quality should include accommodation, food and beverages, guidance and informative support, safety and security, transportation, spiritual experience, accessibility and hygiene, extracurricular activities or heritage visits, and healthcare. In addition, a contemporary element that should be included is digitalization, which refers to electronic

services associated with and supporting gadgets, such as phone applications and Radio Frequency Identification (RFID) bracelets. The anticipated Hajj service quality dimensionality needs to be further validated by an expert study to finalize a comprehensive instrument construction. Figure 1 represents the Hajj service quality anticipated dimensionality. Further research to develop a model for the Hajj service quality (HSQ) assessment is highly needed. It was noticed that the sample size in HSQ is relatively low compared to Hajj population due to the challenging nature of judgmental or convenience sampling. This matter can be mitigated by proper coordination with Hajj authorities to support data collection and overcome the challenging nature of service quality assessment.

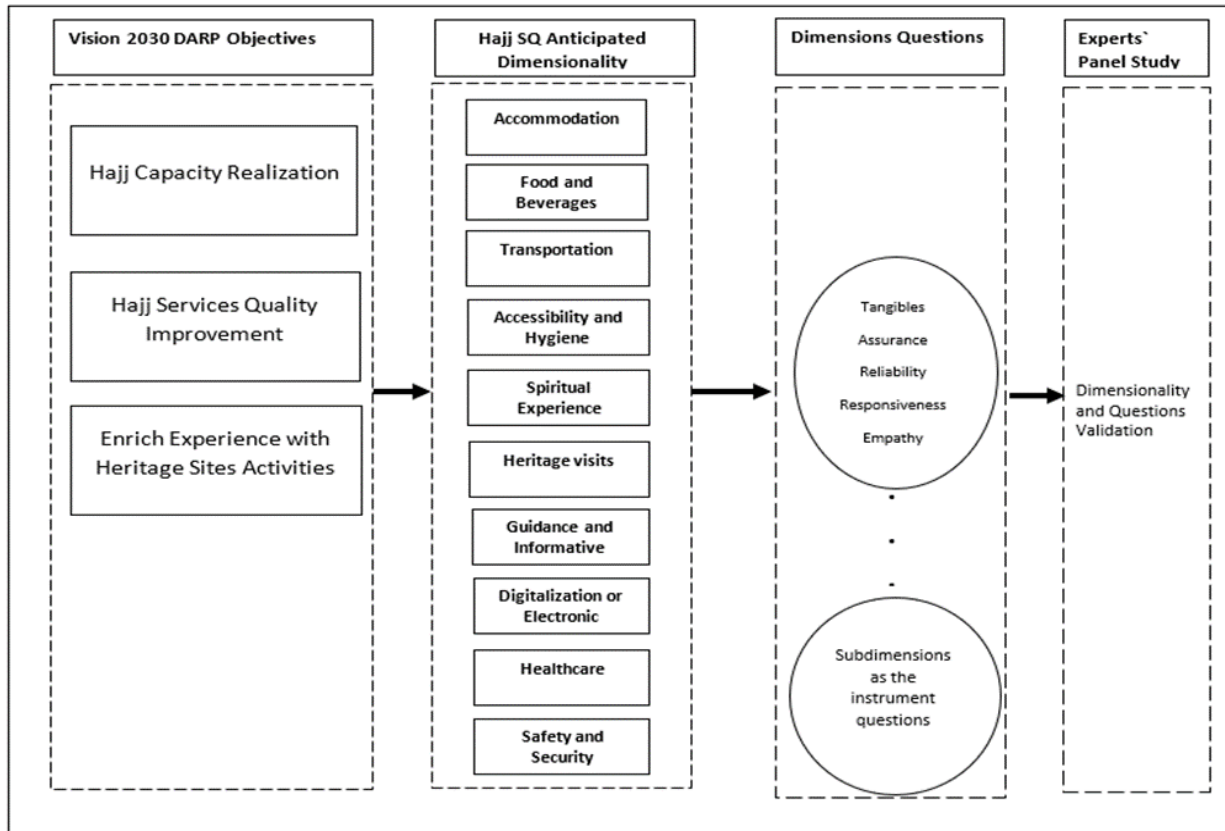


Figure 1. HSQ Anticipated Dimensionality and conceptual framework.

5. References

- Abdo, S. a. G., John Bagot. (2022). Mecca. In Encyclopedia Britannica.
- Al-Hoqail, I. A., Abdalla, A. M., Saeed, A. A., Al-Hamdan, N. A., & Bahnassy, A. A. (2010). Pilgrims satisfaction with ambulatory health services in Makkah, 2008. *Journal of Family and Community Medicine*, 17(3), 135.
- Al Jahdali, K. D. (2021). *Towards an intelligent Holy City: assessing the current training needs for the Ministry of Hajj and Umrah*. Nottingham Trent University (United Kingdom).
- Alsharief, R., & El-Gohary, H. (2017). *Service Quality and Religious Tourism: The Context of Hajj (Islamic Pilgrimage)*. Noor Publishing.
- ARASLI, H., ÇAKMAKOĞLU, N., ARICI, H. E., & ARASLI, F. (2017). DETERMINANTS OF SERVICE QUALITY IN HAJJ TOURISM. *PROCEEDINGS BOOKS*, 190.
- Bahia, K., & Nantel, J. (2000). A reliable and valid measurement scale for the perceived service quality of banks. *International journal of bank marketing*, 18(2), 84-91.

- Berkley, B. J., & Gupta, A. (1994). Improving service quality with information technology. *International journal of information management*, 14(2), 109-121.
- Bitner, M. J., & Hubbert, A. R. (1994). Encounter satisfaction versus overall satisfaction versus quality. *Service quality: New directions in theory and practice*, 34(2), 72-94.
- Brady, M. K., & Cronin Jr, J. J. (2001). Some new thoughts on conceptualizing perceived service quality: a hierarchical approach. *Journal of marketing*, 65(3), 34-49.
- Broderick, A. J., & Vachirapornpuk, S. (2002). Service quality in internet banking: the importance of customer role. *Marketing Intelligence & Planning*, 20(6), 327-335.
- Brogowicz, A. A., Delene, L. M., & Lyth, D. M. (1990). A synthesised service quality model with managerial implications. *International Journal of Service industry Management*, 1(1), 0-0.
- Chang, H. H., Wang, Y.-H., & Yang, W.-Y. (2009). The impact of e-service quality, customer satisfaction and loyalty on e-marketing: Moderating effect of perceived value. *Total quality management*, 20(4), 423-443.
- Corrêa, H. L., Ellram, L. M., Scavarda, A. J., & Cooper, M. C. (2007). An operations management view of the services and goods offering mix. *International Journal of Operations & Production Management*, 27(5), 444-463.
- Cronin Jr, J. J., & Taylor, S. A. (1992). Measuring service quality: a reexamination and extension. *Journal of marketing*, 56(3), 55-68.
- Cronin Jr, J. J., & Taylor, S. A. (1994). SERVPERF versus SERVQUAL: reconciling performance-based and perceptions-minus-expectations measurement of service quality. *Journal of marketing*, 58(1), 125-131.
- Crosby, P. B. (1979). *Quality is free: The art of making quality certain* (Vol. 94): McGraw-hill New York.
- Dabholkar, P. A., Shepherd, C. D., & Thorpe, D. I. (2000). A comprehensive framework for service quality: an investigation of critical conceptual and measurement issues through a longitudinal study. *Journal of retailing*, 76(2), 139-173.
- Dabholkar, P. A., Thorpe, D. I., & Rentz, J. O. (1996). A measure of service quality for retail stores: scale development and validation. *Journal of the Academy of marketing Science*, 24(1), 3-16.
- Darfoon, M. (2013). *An examination of service quality and satisfaction in a religious tourism setting* (Doctoral dissertation, Clemson University).
- Diamant, J. (2019). The countries with the 10 largest Christian populations and the 10 largest Muslim populations. Retrieved from <https://www.pewresearch.org/fact-tank/2019/04/01/the-countries-with-the-10-largest-christian-populations-and-the-10-largest-muslim-populations/>
- Donnelly Jr, J. H. (1976). Marketing Intermediaries in Channels of Distribution for Services: Service marketers should take a fresh look at the channels of distribution for services as distinct from the channels concept followed for goods. *Journal of Marketing*, 40(1), 55-57.
- Doyof Al Rahman Program, V. (2019). Retrieved from <https://darf.gov.sa/en.html>
- Edgett, S., & Parkinson, S. (1993). Marketing for service industries-A review. *Service Industries Journal*, 13(3), 19-39.
- Eid, R. (2012). Towards a high-quality religious tourism marketing: The case of Hajj service in Saudi Arabia. *Tourism Analysis*, 17(4), 509-522.
- Endeshaw, B. (2019). A review on the existing service quality measurement models. *Science Journal of Business and Management*, 7(4), 87-201.
- Frost, F. A., & Kumar, M. (2000). INTSERVQUAL—an internal adaptation of the GAP model in a large service organisation. *Journal of services marketing*.
- Garrigos-Simon, F. J., Narangajavana-Kaosiri, Y., & Narangajavana, Y. (2019). Quality in tourism literature: A bibliometric review. *Sustainability*, 11(14), 3859.
- Ghobadian, A., Speller, S., & Jones, M. (1994). Service quality. *International journal of quality & reliability management*.
- Ghotbabadi, A. R., Baharun, R., & Feiz, S. (2012). A review of service quality models. Paper presented at the 2nd International Conference on Management.
- Gronroos, C. (1984). A service quality model and its marketing implications.
- Hallowell, R. (1996). The relationships of customer satisfaction, customer loyalty, and profitability: an empirical study. *International journal of service industry management*.

- Haq, F., & Jackson, J. (2009). Spiritual journey to Hajj: Australian and Pakistani experience and expectations. *Journal of management, spirituality & religion*, 6(2), 141-156.
- Hassan, T. H., Abdou, A. H., Abdelmoaty, M. A., Nor-El-Deen, M., & Salem, A. E. (2022). The impact of religious tourists' satisfaction with Hajj services on their experience at the sacred places in Saudi Arabia. *Geo J. Tour. Geosites*, 43, 1013-1101.
- Haywood-Farmer, J. (1988). A conceptual model of service quality. *International journal of operations & production management*, 8(6), 19-29.
- Jabnoun, N. (2003). Development of Hajqual: A marketing research tool to measure Hajj service quality. *Journal for International Business and Entrepreneurship Development*, 1(1), 22-28.
- Jain, P., & Aggarwal, V. (2015). Service quality models: A review. *BVIMSR's Journal of Management Research*, 7(2), 125-136.
- Jaw, C., Lo, J.-Y., & Lin, Y.-H. (2010). The determinants of new service development: Service characteristics, market orientation, and actualizing innovation effort. *Technovation*, 30(4), 265-277.
- Kadi, A., & Selim, G. (2022). The Impact of Urban Management on Crowd Movement and User Experience through the Pilgrimage of Hajj (Frequent Temporary Mega-Events). Paper presented at the Proceedings of the International Conference of Contemporary Affairs in Architecture and Urbanism-ICCAUA.
- Kunst, P., & Lemmink, J. (2000). Quality management and business performance in hospitals: A search for success parameters. *Total Quality Management*, 11(8), 1123-1133.
- Lehtinen, U., & Lehtinen, J. R. (1982). Service quality: a study of quality dimensions: Service Management Institute.
- Lehtinen, U., & Lehtinen, J. R. (1991). Two approaches to service quality dimensions. *Service Industries Journal*, 11(3), 287-303.
- Long, M., & McMellon, C. (2004). Exploring the determinants of retail service quality on the Internet. *Journal of services marketing*.
- Lovelock, C. H. (1983). Classifying services to gain strategic marketing insights. *Journal of marketing*, 47(3), 9-20.
- Mattsson, J. (1992). A service quality model based on an ideal value standard. *International Journal of Service industry Management*, 3(3), 0-0.
- Mei, A. W. O., Dean, A. M., & White, C. J. (1999). Analysing service quality in the hospitality industry. *Managing Service Quality: An International Journal*, 9(2), 136-143.
- Foris Oh, H. (1999). Service quality, customer satisfaction, and customer value: A holistic perspective. *International journal of hospitality management*, 18(1), 67-82.
- Oliver, R. L. (1993). A conceptual model of service quality and service satisfaction: Comparative goals, different concepts. *Advances in service marketing and management*, 2, 65-85.
- Parasuraman, A., Berry, L. L., & Zeithaml, V. A. (1991). Refinement and reassessment of the SERVQUAL scale. *Journal of retailing*, 67(4), 420.
- 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.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). Servqual: A multiple-item scale for measuring consumer perc. *Journal of retailing*, 64(1), 12.
- Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1994). Reassessment of expectations as a comparison standard in measuring service quality: implications for further research. *Journal of marketing*, 58(1), 111-124.
- Parasuraman, A., Zeithaml, V. A., & Malhotra, A. (2005). ES-QUAL: A multiple-item scale for assessing electronic service quality. *Journal of service research*, 7(3), 213-233.
- Peck, N. (2013). Development of a service delivery framework for South African Pilgrims travelling to Saudi Arabia. Cape Peninsula University of Technology,
- Philip, G., & Hazlett, S. A. (1997). The measurement of service quality: a new P-C-P attributes model. *International journal of quality & reliability management*.
- Plecher, H. (2020). Distribution of the Workforce Across Economic Sectors in the United States from 2009 to 2019. *Statistica*, July, 15.
- Prakash, A., & Mohanty, R. (2013). Understanding service quality. *Production Planning & Control*, 24(12), 1050-1065.
- Reeves, C. A., & Bednar, D. A. (1994). Defining quality: alternatives and implications. *Academy of management Review*, 19(3), 419-445.

- Robbany, R. B. (2018). The Relationship Between Service Quality And Hajj Applicant Satisfaction: Study About Administration Process At Banyumas. Universitas Brawijaya,
- Rust, R. T., & Oliver, R. L. (1994). Service quality: insights and managerial implications from the frontier. *Service quality: New directions in theory and practice*, 7(12), 1-19.
- Santos, J. (2003). E-service quality: a model of virtual service quality dimensions. *Managing Service Quality: An International Journal*, 13(3), 233-246.
- Seth, N., Deshmukh, S., & Vrat, P. (2005). Service quality models: a review. *International journal of quality & reliability management*, 22(9), 913-949.
- Shahin, A., & Samea, M. (2010). Developing the models of service quality gaps: a critical discussion. *Business Management and Strategy*, 1(1), 1.
- Sivadas, E., & Baker-Prewitt, J. L. (2000). An examination of the relationship between service quality, customer satisfaction, and store loyalty. *International Journal of Retail & Distribution Management*.
- Soteriou, A. C., & Stavrinides, Y. (2000). An internal customer service quality data envelopment analysis model for bank branches. *The International Journal of Bank Marketing*, 18(5), 246-252.
- Spreng, R. A., & Mackoy, R. D. (1996). An empirical examination of a model of perceived service quality and satisfaction. *Journal of retailing*, 72(2), 201-214.
- Sun, H. (2001). Comparing quality management practices in the manufacturing and service industries: learning opportunities. *Quality Management Journal*, 8(2), 53-71.
- Svensson, G. (2004). A customized construct of sequential service quality in service encounter chains: time, context, and performance threshold. *Managing Service Quality: An International Journal*, 14(6), 468-475.
- Sweeney, J. C., Soutar, G. N., & Johnson, L. W. (1997). Retail service quality and perceived value: A comparison of two models. *Journal of retailing and Consumer Services*, 4(1), 39-48.
- Teas, R. K. (1993). Expectations, performance evaluation, and consumers' perceptions of quality. *Journal of marketing*, 57(4), 18-34.
- Weeks, J. R. (2020). Population: An introduction to concepts and issues: Cengage Learning.
- Wolak, R., Kalafatis, S., & Harris, P. (1998). An investigation into four characteristics of services. *Journal of Empirical Generalisations in Marketing Science*, 3(2).
- Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1985). Problems and strategies in services marketing. *Journal of marketing*, 49(2), 33-46.
- Zhu, F. X., Wymer, W., & Chen, I. (2002). IT-based services and service quality in consumer banking. *International journal of Service industry management*, 13(1), 69-90.

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

U.S. Wildlife Strikes by Phase of Flight

Saleem Hassan S. Alhumaidi¹

Ibrahim Yousef N. Alruwaili¹

Brooke E. Wheeler^{*,1}

¹*Florida Institute of Technology*

bwheeler@fit.edu

Abstract

Aircraft collisions with birds and other wildlife are of increasing concern to the aviation industry. The U.S. Federal Aviation Administration (FAA) developed the National Wildlife Strike Database to better define the wildlife strike problem. The purpose of this study was to determine the difference in the annual wildlife strike reports by phase of flight in the United States of America from 2010 to 2019. The phases of flight were categorized as found in the FAA Wildlife Strike Database: take-off run, landing roll, departure, arrival, climb, descent, en route, local, taxi, parked, and approach. We excluded reports with an unknown phase of flight from the study. This information provides airports with an objective baseline to aid in the evaluation of wildlife risk mitigation programs. The analysis showed a significant difference in annual wildlife strike reports by phase of flight. Approach had the highest mean number of wildlife strikes, whereas parked had the lowest mean number of strikes.

Keywords: Wildlife strikes, phase of flight, aviation safety.

1. Introduction

Blackwell et al. (2009) noted that aircraft collisions with wildlife are increasingly severe economic and safety concerns for the United States (US) aviation industry. In the years 1990-2004, wildlife strikes cost the US aviation industry an estimated \$500 million annually (Dolbeer, 2006). According to Dolbeer et al. (2018) and Allen (2020), wildlife collisions cost commercial airlines \$1.2 billion globally annually from 1999 to 2000. Collisions between wildlife and aircraft have impacted many people, both military and civilian (Hedayati & Sadighi, 2016). Between 1990 and 2022, wildlife strikes have continued to increase in the US, with 17,190 reports in 2022 alone. The average repair costs per strike were \$155,330 with an average aircraft downtime of 43.4 hours, for approximately \$385 million total impact of strike costs in 2022 (Dolbeer et al., 2023). While the financial and safety consequences of these wildlife strikes have been documented, their ecological consequences remain understudied (Fernandez-Juricic et al., 2018). In that regard, establishing the difference in wildlife strike reports by the phases of flight will aid in developing a sustainable solution to avoid wildlife strike damages.

The purpose of this study was to determine the differences in the annual wildlife strike reports by phase of flight in the US from 2010 to 2019. The phases of flight were categorized as found in the Federal Aviation Administration (FAA, 2022) Wildlife Strike Database into take-off run, landing roll, departure, arrival, climb, descent, en route, local, taxi, parked, and approach.

This study provides an analysis of the rate of wildlife strike reports that aircraft experience in

different phases of flights annually from 2010 to 2019. This understanding is essential to all aviation professionals because it will help improve aviation safety and minimize financial expenses. Therefore, this research could be used to improve anticipation of the risk of wildlife strikes in such scenarios by increasing the activation of NOTAM advisories. The outcomes of this study are generalizable to all civil aviation aircraft operations within the US because we used all strikes reported to the FAA (2022) Wildlife Strike Database for all flight operations inside the US for a period from January 1, 2010 to December 30, 2019.

2. Literature review

We reviewed previous research on causal factors for wildlife strikes, namely the geographical location of airports, aircraft characteristics, habitat management, altitude, species behavior, and aircraft lighting. Additionally, the current state of knowledge regarding wildlife strikes was examined.

Wildlife strikes are of great concern to the aviation community. Moreover, wildlife strikes pose a significant threat to flight safety. For example, U.S. Airways Flight 1549 had a series of strikes, resulting in the failure of both engines and causing a serious inflight emergency (Ferra et al., 2021). In today's ever-increasing flight operations, civil aviation has had to ensure sufficient revenues with acceptable safety levels to provide a sustainable aviation business. Wildlife strikes, particularly bird strikes, are a serious threat to aviation safety and cause a major dent in revenue (Dolbeer, 2022). It is estimated that the cost of damage caused by wildlife strikes is more than \$1.2 billion per annum (Allan, 2000). In addition to financial costs, these strikes have also claimed over 200 aircraft and 219 lives since 1988. In the US, a total of 82,057 wildlife strikes were reported to FAA during 1990-2007, which consisted of 97.5% of bird strikes and cost almost \$628 million per annum either in direct or indirect losses (Dolbeer et al., 2009). It is pertinent to note that statistics on wildlife strikes have shown an increased risk to flight safety (Blackwell et al., 2009). As such, the factors associated with wildlife strikes should be established to aid in solving this problem.

2.1. Wildlife strikes by regions

Ferra et al. (2021) noted that wildlife strikes were unavoidable, and that the aviation industry could mitigate them through several means. These include airport management plans, national wildlife strikes reporting systems, and deterring animal inhabitants around airports. Ferra et al. (2021) and Allan (2006) agreed that animal inhabitants near airports are a significant cause of wildlife strikes. Ferra et al. (2021) aimed to determine the difference in the frequency of wildlife strikes at non-military class B airports in the contiguous US by geographical region. The authors used data from the FAA (2022) Wildlife Strikes Database for the period from 1/1/2015 to 12/31/2019. They established that the western mountains and Texas had the highest frequency of wildlife strikes and that the Pacific Coast had a lower frequency of wildlife strikes. Furthermore, they noted that the number of wildlife strikes on the Pacific coast was significantly lower than in the other four regions. Ferra et al. (2021) speculated that the results could be attributed to the vast part of the area that is deserted and the fewer daily airport operations at most regional airports. The findings of Ferra et al. (2021) suggest that there is a need to mark specific regions in the contiguous US due to the frequency of wildlife strikes.

Many factors attract wildlife to airports including food, shelter, and water. Food sources include

grassy and weedy zones around airport operating surfaces, which are attractive to herbivores as well as carnivores hunting for food. Shrubs, trees, grass lawns, and airport infrastructure offer shelter to different species. Lastly, water sources include natural and artificial water bodies that provide habitat to species of animals. Moreover, many airports have a water drainage system where the excess water is drained from the operating surfaces of the airport and collected far away to ensure safe operations. However, this temporary pooling of drained water can attract wildlife, thus increasing the risk of wildlife strikes (Blackwell et al., 2009). To address the airport land use issues relevant to wildlife strikes, Blackwell and colleagues (2009) concluded that in order to mitigate the risk of wildlife strikes, airports needed to be constructed and managed as per specific guidelines addressing issues such as the management of storm water, quantification of species population on the airport, wildlife use of agricultural crops and enforcement of buffer areas in close vicinity of airports (Blackwell et al., 2009). Airport management must strive to limit the attractiveness of airport environments to these bird species.

2.2. Impact of aircraft characteristics

With advancements in the aviation industry, air traffic has drastically increased. Moreover, individual aircraft characteristics are also prominent decisive factors for evaluating the likelihood of bird strikes. Modern technologies have resulted in larger airframes and turbojet or turbofan engines replacing piston-powered engines; both design changes have resulted in a higher probability of wildlife strikes. Moreover, the larger airframe provided more surface area to hit and suction associated with intakes of turbojets/turbofans, making it very difficult for any wildlife to change its flight and avoid ingestion/strike (Blackwell et al., 2009). Moreover, turbofan engines are more prone to bird ingestion than any other engine type because of their large inlet size and greater suction effect (Australian Transport Safety Bureau, 2014). However, the intakes of commercial aircraft have seen a significant increase in size (diameter), thus further increasing the risk of ingestion (Hall, 2005). Similarly, there has been an increase in the number of aircraft with turbofan engines. Moreover, the percentage of commercial aircraft with turbofan engines in relation to all commercial aircraft increased from 79.6% in 2006 to 86.5% in 2015. Furthermore, according to Canadian data, turbofan aircraft faced 1.7 times more bird strikes than any other aircraft type from 2008 to 2018 (Metz et al., 2020). Hence, aircraft designs should limit collisions with birds and other wildlife.

Aircraft noise emissions also affected the rate of wildlife strikes. Quieter aircraft noise results in an inability to detect or late detection of the aircraft, thus affecting the ability of birds to initiate collision avoidance maneuvers (Sodhi, 2002). Owing to the adverse effects of noise emissions on neighboring communities, airlines have implemented newer and quieter aircraft, resulting in increased bird strike risk (Hall, 2005). As such, there is a need to devise ways to scare birds from flight paths without causing noise pollution.

2.3. Impact of habitat management

An in-depth analysis of wildlife strike statistics reveals that habitat management in the vicinity of an airport significantly affects the wildlife strike rate (Cleary & Dolbeer, 2005). More importantly, wildlife habitats (especially for birds) in the departure and approach zones of aircraft flight paths increased the probability of wildlife strikes exponentially because in those zones, the aircraft was operating at lower

altitudes, and the aircrew was more focused on doing checklist procedures (Blackwell et al., 2009). Hence, airport environments must be made unattractive to birds and wildlife.

Another significant aspect of climate variation is the migration of different wildlife species. Migratory birds living in northern regions move towards warmer regions during the early winter months and return to their original habitats in the early spring months. Consequently, more bird strikes were experienced during the migration months (Dolbeer et al., 2015). Additionally, areas surrounding the migration path were susceptible to increased bird strikes. Hence, flight operations in these areas during specific time frames required extra vigilance to avoid bird strikes and ensure safe flight operations (Ferra et al., 2021). In contrast, bird strikes were increasing during summer, especially in countries of the Northern Hemisphere, where many bird species have breeding seasons (Gasteren et al., 2008). Similarly, bird strikes increased during spring and autumn, owing to the increased flying activity of migratory birds (Metz et al., 2020). Airport management must improve vigilance during specific times of the year to reduce the incidence of wildlife strikes.

2.4. Impact of altitude

Wildlife strike statistics show that aircraft experience a higher likelihood of bird strikes while flying at low altitudes (McKee et al., 2015). According to a study conducted by Dolbeer et al. (2015), almost 88% of bird strikes in the US over a period of 27 years (1993-2018) happened below 2,500 ft, and 71% occurred below 500 ft. Likewise, a European study found that worldwide almost 95% of wildlife strikes were experienced below 2,500 ft, and nearly 70% of strikes occurred below 200 ft (Metz et al., 2020). The likelihood of wildlife strikes is greater at lower altitudes (ICAO, 2017). Fernández-Juricic et al. (2018) also found that low altitudes have the highest number of wildlife strikes. Conversely, the probability of bird strikes decreases with increasing altitude.

However, bird strikes at higher altitudes resulted in significant damage compared with bird strikes occurring at lower altitudes. Moreover, the greater damage can be attributed to increased aircraft velocity at higher altitudes and higher kinetic energy of the birds due to increased bird size (Metz et al., 2020). Furthermore, Fernández-Juricic et al. (2018) stated that at airports, collisions with low-wing loading birds have a low safety risk for humans because such collisions do not typically cause significant damage to aircraft. While the statistics suggest insignificant damage to aircraft, collisions at low altitudes greatly reduce the populations of low wing loading birds. As such, airport management authorities should reduce the attractiveness of the airports to low wing loading birds.

2.5. Impact of species hazard level and behavior

Not all wildlife species present the same level of hazard to aircraft upon collision. The hazard level of any species can be estimated empirically by calculating the percentage of strikes that result in damage to aircraft. For example, 49% of Canadian goose strikes caused aircraft damage compared to 8% of ring-billed gulls (Dolbeer et al., 2018). Hence, based on the hazard level of each species, the airport management can prioritize wildlife strike mitigation strategies (Dolbeer et al., 2018).

Fernández-Juricic et al. (2018) studied wildlife strikes in relation to species' aerial maneuverability with a primary focus on body mass, wing loading, wing aspect ratio, eye size, and brain size. Wing loading accounted for the difference in bird strikes across the studied species and had a negative linear

association with the frequency of bird-aircraft collisions. Thus, aircraft often struck bird species with lower wing loading, such as the Iwa, due to lower maneuverability. According to Fernández-Juricic et al. (2018), high-wing loading bird species, such as Wandering Albatross, fly faster and can avoid approaching aircraft when in the flight path. Fernández-Juricic et al. (2018) noted that low-wing loading species are struck at a higher rate at airports. Therefore, airport environments should be made unattractive to these specific bird species.

2.6. Impact of aircraft lighting

Bird strikes were more common for communication towers fitted with warning lights of higher wavelengths (e.g., red) than those equipped with lights of shorter wavelengths (e.g., blue). In a recent study conducted by Dolbeer (2018), data from the National Wildlife Strike Database were used to examine the hypothesis that aircraft wings with a red navigation light would experience more bird strikes. Hence, it was concluded that the modification of red navigation light the light having a shorter wavelength could reduce bird strikes (Dolbeer et al., 2018). This among other technological advances may assist in minimizing wildlife strikes near airports.

This section reviewed the literature on several factors associated with wildlife strikes in the US, such as the geographical location of airports, aircraft characteristics, habitat management, altitude, season, species behavior, and aircraft lighting. More specifically, most wildlife strikes occur at 500 ft and below. However, strikes above 500 ft often result in significant damage to aircraft. In addition, wildlife body mass is related to damage during collisions. The literature suggests that wildlife species is the most studied factor contributing to reported strikes. Although there are many existing studies on factors that contribute to wildlife strikes, as reviewed above, in-depth research is needed on other factors, such as the phase of the flight during collisions.

3. Methods

This study adopted an ex post facto design to provide an analysis of the rate of wildlife strike reports that aircraft experience in different phases of flights annually from 2010 to 2019. The data used in this study were obtained from the FAA (2022). The Wildlife Strike Database consists of all types of wildlife strike reports occurring within the US and submitted by relevant agencies by filing FAA Form 5200-7 - Bird / Other Wildlife Strike Reports. This database is maintained and validated by the FAA. The main limitation of this approach is that reporting is voluntary, which means that the dataset contains reported strikes, not all strikes, and there may be missing data for given variables in the reports. The phases of flight that were investigated were categorized as in the FAA (2022) Wildlife Strike Database: take-off run, landing roll, departure, arrival, climb, descent, en route, local, taxi, parked, and approach.

All reports of strikes occurring in the US from January 1, 2010, to December 31, 2019 in the Wildlife Strike Database were downloaded into a Microsoft Excel spreadsheet. Any strike reports with an unknown phase of flight were excluded. Using a Pivot table, the annual rate of wildlife strikes was tallied. Descriptive statistics were computed using Excel, and inferential statistics were determined using R Studio version 4.1.3, including one-way Analysis of Variance (ANOVA), Tukey's pairwise comparison test, and eta squared.

4. Results

Beginning on January 1, 2010, and ending on December 31, 2019, a total of 131,456 wildlife-aircraft collision reports were filed with the FAA (2022). We excluded 55,817 reports from the dataset because their phases of flight were not listed. We omitted entries with an unknown phase of flight from the analysis because we were unable to categorize them properly. Therefore, the total number of wildlife strike reports used in the analysis was 75,639.

Descriptive statistics for wildlife strikes for each phase of flight are presented in Table 1. The phase of flight with the highest average annual wildlife strike reports was approach ($n=32,966$, $M = 3,296.60$), followed by landing roll ($n = 13,634$, $M = 1,363.40$), and then take-off run ($n= 11,535$, $M = 1,153.50$). The departure phase had 1,732 strikes ($M = 173.20$). The phase of flight with the least strikes reported was parked ($n = 64$, $M = 6.40$). Figure 1 illustrates that the approach phase had the highest mean number of strikes, whereas the parked phase had the lowest mean number of strike reports.

Table 1. Descriptive statistics for wildlife strikes by phase of flight (2010 -2019)

Phase of Flight	Number of Strikes	Mean	Median	Mode	Range	Standard Deviation
Parked	64	6.4	6.5	10	1-10	3.2
Taxi	251	25.1	25.5	15	15-39	8.66
Take-off Run	11535	1153.5	1165	NA	1013-1363	104.77
Departure	1732	173.2	200	17	4-338	145.62
Local	645	64.5	52.5	NA	31-119	31.84
Climb	10891	1089.1	1050.5	NA	965-1344	121.69
En Route	2412	241.2	238.5	220	173-344	47.96
Descent	991	99.1	96	67	67-148	27.48
Arrival	517	51.7	42	42	4-147	51.24
Approach	32966	3296.6	3265	NA	2657-4150	521.07
Landing Roll	13634	1363.4	1409	NA	1127-1803	218.76

The one-way ANOVA indicated a statistically significant difference in annual wildlife strike reports by phase of flight, $F(10, 99) = 299.14$, $p < .001$. Tukey's pairwise test results showed the phases of flights that were significantly different. The approach phase was different from all other phases ($p < .001$), while the take-off run, climb, and landing roll were statistically different from all the other phases ($p < .001$) except themselves. The eta-squared value was 0.96. Therefore, the effect size was large.

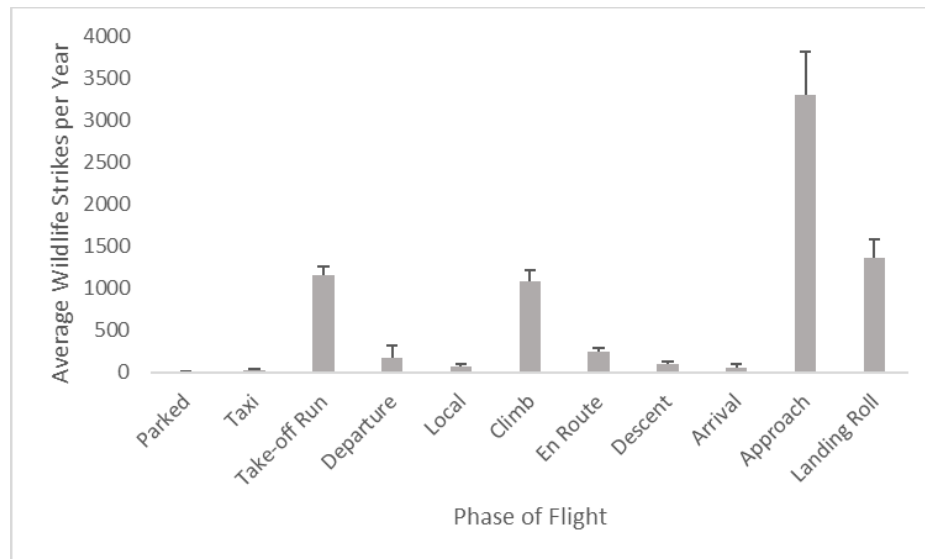


Figure 1. Mean Annual Wildlife Strike Reports by Phase of Flight (2010-2019)

5. Discussion

This study evaluated the differences in annual wildlife strike reports by phase of flight in the US from 2010 to 2019. We hypothesized that there would be a difference in the annual wildlife strike report by the phases of flight. Figure 1 illustrates that the approach phase had the highest annual strike, whereas the parked phase had the lowest mean number of reports. The findings supported our hypothesis; there was a significant difference in the annual wildlife strikes reported by the phases of flight, $F(10, 99) = 299.14$, $p \leq .001$. In addition, post-hoc pairwise comparisons showed that the annual wildlife strike reports for most of the flight phases were significantly different. The approach phase had a statistically higher rate of wildlife strikes than all other phases, while the take-off run, climb, and landing roll were different from all the phases except themselves. Eta squared (0.96) indicated a large effect size.

One potential explanation for the higher strikes during approach is that commercial aircraft now operate with more efficient and quieter engines, which leads to birds reacting late to avoid collisions (Metz et al., 2020). In addition, birds are more active at lower altitudes, which coincides with the approach phase. Dolbeer et al. (2015) established that almost 88% of bird strikes in the US over a period of 27 years (1993-2018) happened below 2,500 ft, and 71% occurred below 500 ft. These results align well with our findings, as the approach, landing roll, and take-off run are the highest reported phases, with lower averages in the En Route and Local phases. Findings from this study should be used in flight training to help pilots understand which phases of flight are the most hazardous for wildlife strikes: approach, landing roll, take-off run, and climb.

Further research should be conducted to establish which species might be at risk of wildlife strikes during different flight phases. It would also be interesting to investigate why there were more wildlife strikes in the approach phase of flight than other flight phases. This research was delimited to the US and the FAA (2022) database. Expanding research outside the US may confirm our findings in other countries.

The approach phase had the highest mean number of wildlife strike reports, whereas parked aircraft

had the lowest mean number of reports. In addition, the mean number of wildlife strikes reported during the approach phase was different from all other phases, while the take-off run, climb, and landing roll were different from all phases except themselves. More research is needed to understand wildlife strikes by the phases of flight to develop pilots' training to avoid wildlife strike damage and to advance the available measures to prevent wildlife strikes during flight. Given that the approach phase of flight is the most at risk for a wildlife strike, future research with airports should focus on mitigation strategies for this phase.

6. References

- Allan, J. (2000). The costs of bird strikes and bird strike prevention. *Human Conflicts with Wildlife: Economic Considerations*, Proceedings of the Third NWRC Special Symposium. pp. 147–153. Retrieved October 3, 2022
- Allan, J. (2006). A heuristic risk assessment technique for Birdstrike Management at airports. *Risk Analysis*, 26(3), 723–729. <https://doi.org/10.1111/j.1539-6924.2006.00776.x>
- Australian Transport Safety Bureau. (2014). Australian Aviation Wildlife Strike Statistics 2004 to 2013. Canberra, Australia: Australian Transport Safety Bureau. Retrieved October 3, 2022, from https://www.atsb.gov.au/media/5177450/ar2014075_final.pdf
- Blackwell, B. F., DeVault, T. L., Fernández-Juricic, E., & Dolbeer, R. A. (2009). Wildlife collisions with aircraft: A missing component of land-use planning for airports. *Landscape and Urban Planning*, 93(1), 1–9. <https://doi.org/10.1016/j.landurbplan.2009.07.005>
- Cleary, E. C., & Dolbeer, R. A. (2005). Wildlife Hazard Management at airports: A manual for airport personnel. *USDA National Wildlife Research Center-Staff Publications*, 133.
- Dolbeer, R. A. (2006). Height distribution of birds recorded by collisions with civil aircraft. *The Journal of Wildlife Management*, 70(5), 1345–1350. [https://doi.org/10.2193/0022-541x\(2006\)70\[1345:hdborb\]2.0.co;2](https://doi.org/10.2193/0022-541x(2006)70[1345:hdborb]2.0.co;2)
- Dolbeer, R. (2022). Population increases of large birds in North America pose challenges for aviation safety. *Human-Wildlife Interactions*, 14(3), 345–357. <https://doi.org/https://doi.org/10.26077/53f9-edc3>
- Dolbeer, R. A., Begier, M. J., & Weller, J. (2018). The National Wildlife Strike Database: A Scientific Foundation to enhance aviation safety. *Proceedings of the Vertebrate Pest Conference*, 28. <https://doi.org/10.5070/v42811027>
- Dolbeer, R. A., Wright, S. E., Weller, J., Begier, M. J., (2009). Wildlife strikes to civil aircraft in the United States, 1990-2008. United States. Department of Transportation. Federal Aviation Administration, & United States. Department of Agriculture. Animal and Plant Health Inspection Service. Wildlife Services. (2009, September 1). Retrieved October 3, 2022, from <https://rosap.ntl.bts.gov/view/dot/6461>
- Dolbeer, R., Wright, S., Weller, J., Anderson, A., & Begier, M. (2015). Wildlife strikes to civil aircraft in the United States, 1990–2014. FAA Wildlife Strikes Database Serial Report Number 21. Washington, DC, USA. Retrieved October 3, 2022, from <http://i2.cdn.turner.com/cnn/2016/images/09/07/faa.bird.strike.report.2014.pdf>
- Dolbeer, R.A., Begier, M. J., Miller, P.R., Weller, J., & Anderson, A.L. (2023). Wildlife strikes to civil aircraft in the United States, 1990-2022. FAA Wildlife Strikes Database Serial Report Number 29. Department of Transportation. Federal Aviation Administration, & United States. Department of Agriculture. Retrieved October 12, 2023, from <https://www.faa.gov/sites/faa.gov/files/Wildlife-Strike-Report-1990-2022.pdf>
- FAA. (2022). FAA wildlife strike database. FAA Wildlife Strike Database. Retrieved September 28, 2022, from <https://wildlife.faa.gov/home>.
- Fernández-Juricic, E., Brand, J., Blackwell, B. F., Seamans, T. W., & DeVault, T. L. (2018). Species with greater aerial maneuverability have higher frequency of collisions with aircraft: A comparative study. *Frontiers in Ecology and Evolution*, 6, 17. <https://doi.org/10.3389/fevo.2018.00017>
- Ferra, G., Alghamdi, H., Wheeler, B. E., & Li, T. (2021). Comparing the Frequency of Reported Wildlife Strikes by Region in the United States. *Journal of Management & Engineering Integration*, 14(1), 37-45.

- Gasteren, H. V., Holleman, I., Bouten, W., Loon, E. V., & Shamoun-Baranes, J. (2008). Extracting bird migration information from C-band Doppler Weather Radars. *Ibis*, 150(4), 674–686.
<https://doi.org/10.1111/j.1474-919x.2008.00832.x>
- Hall, C. (2004). Engine configurations for the silent aircraft. *The Journal of the Acoustical Society of America*, 116(4), 1164-2005. <https://doi.org/10.1121/1.4785127>
- Hedayati, R., & Sadighi, M. (2016). Bird strike: An experimental, theoretical and numerical investigation. Elsevier, WP Woodhead Publishing.
- ICAO. (2017). International Civil Aviation Organization - ICAO. Retrieved October 3, 2022, from <https://skybrary.aero/sites/default/files/bookshelf/4069.pdf>
- McKee, J., Shaw, P., Dekker, A., & Patrick, K. (2016). Approaches to wildlife management in aviation. In *Problematic wildlife: A cross-disciplinary approach*, 465-488. Springer, Cham.
https://doi.org/10.1007/978-3-319-22246-2_22
- Metz, I. C., Ellerbroek, J., Mühlhausen, T., Kügler, D., & Hoekstra, J. M. (2020). The bird strike challenge. *Aerospace*, 7(3), 26. <https://doi.org/10.3390/aerospace7030026>
- Sodhi, N. S. (2002). Competition in the air: Birds versus aircraft. *The Auk*, 119(3), 587–595.
<https://doi.org/10.1093/auk/119.3.587>

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Unhelpful and Unaware of It: A Dyadic Analysis of Online Product Reviews

Scott D. Swain¹

¹*Clemson University*

sdswain@clemson.edu

Abstract

Much research focuses on identifying characteristics that predict whether consumer reviews are perceived as helpful. In contrast, little is known about whether review writers themselves know if their reviews will be helpful or whether they can be provided with effective writing prompts to improve the helpfulness of their reviews. Across two studies, the evidence suggests that while review writers are overconfident, their reviews are most helpful when their attentional focus during writing is on others (versus themselves) and when reviewing products characterized predominately by search (vs. experience) qualities.

Keywords: Product reviews, Online reviews, Helpfulness, Overconfidence, Search goods, Experience goods, Attentional focus, Others-focused, Self-focused

1. Introduction

Online consumer reviews have become a pervasive source of product information in the marketplace. For example, YouGov (2022) reports that, on average, one-third of global consumers regularly use consumer reviews before purchasing homeware products, while a similar percentage do so before buying clothes. When asked which source of product information was most important in the shopping process, 34% of consumers listed online reviews, second only to search engines at 54%, and just ahead of online stores at 33% (Statista, 2023). Similarly, 37% of consumers say customer reviews and ratings help “a lot” (as opposed to “some,” “a little,” or “not at all”) when it comes to making them feel confident about product decisions (Pew Research Center, 2018).

Reviews provide value not only for consumers but also manufacturers, service providers, and retailers. Consumers use reviews as a form of word-of-mouth to acquire information about options and attributes that might otherwise be difficult or impossible to obtain prior to making a purchase (Hanna, Swain, & Smith, 2016; Smith, Menon, & Sivakumar, 2005; Tractinsky & Rao, 2001; Voight, 2007). Similarly, manufacturers and retailers have learned that facilitating review technologies leads to greater customer awareness as well as lower costs due to less need for direct communications and lower probabilities of product returns (Sahoo, Dellarocas, & Srinivasan, 2018; Swain, 2022).

Recent research in the areas of information systems and marketing provide a number of insights on the impact of product reviews in the marketplace. Reviews have been linked to outcomes such as profitability and viewership (e.g., Cui, Lui, & Guo, 2012; Sun, 2012), as well as website and product evaluations (e.g., Hu, Zhang, & Pavlou, 2009). To understand these effects, researchers have focused on

identifying the review characteristics that influence shoppers' perceptions of review helpfulness (Hong et al., 2017; King, Racherla, & Bush, 2014). This work finds that helpfulness is enhanced when reviews are perceived as diagnostic (e.g., Mudambi & Schuff, 2010), credible (e.g., Wu, 2013), and appropriate for the type of product (e.g., Weathers, Swain, & Grover, 2015).

A second stream of research on review helpfulness examines the motivations and characteristics of review writers (e.g., Dellarocas, Gao, & Narayan, 2010). While this stream of research has proven useful for understanding or predicting whether a person will write a review, it thus far offers little insight as to when or how eventual writers create reviews that readers deem helpful. The present research seeks to address this limitation in the literature.

2. Writers' Motivations

Prior research has documented a variety of motivations that lead consumers to write reviews. One motivation is to reward (or punish) sellers. Positive reviews allow writers to express their gratitude (Lafky, 2014), while negative reviews allow writers to express dissatisfaction (Dellarocas et al., 2010). A second motive some writers have is their concern for others, which they can express by warning others about deceptive or deficient products or by drawing attention to attractive products or promotions (Bronner & de Hoog, 2011; Cheng et al., 2019; Swain, Berger, & Weinberg, 2014). A third motive present in some writers is the need for self-enhancement. Writers can help maintain a positive self-image by associating the self with positive product outcomes and by disassociating the self from negative product outcomes (Dixit, Badgaiyan, & Khare, 2019; Gonçalves et al., 2018). Lastly, some writers have a desire to be heard or to be viewed as an opinion leader (Chai et al., 2023). Similarly, some writers like to feel like part of group by sharing in the exchange of opinions (Munzel & Kunz, 2014).

It is important to bear in mind that most contributors to online review writers are not professional writers or communication specialists. As such, it should not be surprising that many reviews are not deemed helpful by readers. Recognizing this, retailers and review aggregators often provide "tips" for writing helpful reviews. For example, Google, Trustpilot, TripAdvisor, and Yelp have variously instructed writers that reviews should follow certain guidelines such as focusing on facts, providing details, staying balanced, using proper grammar, avoiding pickiness, and remaining respectful.

However, little evidence to date supports the idea that providing writing instructions as to the content or tone of a review is effective for eliciting reviews that readers find more helpful. Drawing on persuasion theory, the present research investigates a different approach to prompting writers, one in which writers are not advised as to what to write about but instead from which perspective to write (Campbell & Kirmani, 2000; Friestad & Wright, 1994; Weathers, Swain, & Carlson, 2012). This approach is described in the next section.

3. Attentional Focus

When constructing a review, writers can focus their attention on themselves, or they can focus attention on others (especially other consumers interested in the same products). Attentional focus is readily observable in everyday communications, as it manifests in common linguistic mechanisms. For example, individuals tend to use personal pronouns such as "I" and "me" when communicating with a self-focus, whereas they tend to use others-related terms such as "you" when communicating with a

focus on others (Ickes, Reidhead, & Patterson, 1986; Pennebaker, Mehl, & Niederhoffer, 2003).

Since attentional focus is prominent in human interactions (Galinsky and Moskowitz, 2000), writers who adopt different foci may create reviews that differ in ways that have an impact on consumers' perception of reviews. Indeed, the ability to be more helpful to others by virtue of shifting attentional focus is well-recognized in the literature on frontline employee training (e.g., Korschun, Bhattacharya, & Swain 2016). Consistent with this prior research, it is argued that prompting writers to focus on others (versus themselves) results in greater perceptions of helpfulness among readers (Lei, Yin, and Zhang, 2021).

This is likely to occur for at least two reasons. First, when individuals focus their attention on others (versus the self), they are more likely to display empathic concern, leading others to perceive them as more caring and trustworthy (Goldstein, Vezich, & Shapiro, 2014). In the context of product reviews, this suggests that others-focused reviews will lead to positive feelings or associations that spill over to readers' evaluations of the review. Second, writers who place their focus on others (versus only themselves) are more likely to consider and address a broader range of information needs in their product reviews. This results in a higher probability that the information needs of a given reader will be met. Thus,

H₁: Writers who are others-focused (versus self-focused) will write reviews that are perceived as more helpful by readers.

4. Product type

Prior research suggests that the type of product reviewed is an important factor for understanding the potential helpfulness of a review (Hong, Xu, Wang, & Fan, 2017). Though there are different paradigms for classifying products, Nelson's (1974) search versus experience paradigm has proven particularly useful (e.g., Baek, Ahn, & Choi, 2012; Brunel & Swain, 2007; Jiménez & Mendoza 2013). The search versus experience paradigm distinguishes between products based on the extent to which consumers must experience the products prior to deciding which option to purchase. However, the emergence of online (and virtual) retail channels means that the search versus experience distinction can vary for the same product depending on which retail channel it is encountered (Carlson, Weathers, & Swain, 2016; Mittal 2004; Sun, Han, & Feng, 2019).

Thus, in the present research, the search versus experience distinction is conceptualized as the extent to which consumers feel that they need to evaluate a product in person. The greater the perceived need to evaluate a product in person, the more experience versus (search) qualities the product possesses. Thus, readers should generally perceive reviews of search (versus experience) products as more helpful since any information conveyed is more directly evaluable for search (versus experience) products. Additionally, readers are more likely to discount writers' communications about personal experiences as idiosyncratic. Thus,

H₂: Writers' reviews will be perceived as more helpful by readers for search (versus experience) products.

A further question of interest in the present research is whether any effects of a writer's attentional

focus differ depending on the type of product reviewed. For products that must be experienced in-person to be evaluated, readers may have an enhanced preference for reviews written with a self-focus since the writer possesses the relevant experience. In contrast, for search products, readers may have an enhanced preference for reviews written with an others-focus since the product information can be evaluated independent of the writer's personal experience. Thus,

H₃: Writers who are others-focused (versus self-focused) will write reviews that are perceived as more helpful by readers, with the effect being stronger for search (versus experience) products.

Two studies were conducted to test the effects of attentional focus and product type on the helpfulness of writers' reviews.

5. Overview of methodology

This research generated data from both sides of the online consumer review dyad (writers and readers). One group of consumers served as review writers. These consumers were randomly assigned to one of several experimental conditions, asked to write reviews, and then asked to predict how helpful they thought their reviews would be to readers. A second, independent group of consumers served as the readers and were asked to evaluate the helpfulness of the writer's reviews.

6. Study 1: Exploratory analysis of review-writers' self-predicted helpfulness

6.1. Participants, Design, and Procedure

While the research hypotheses focus on the conditions under which writers' reviews are perceived as helpful by readers, Study 1 provides context by exploring writers' perceptions of their own reviews. A total of 24 participants (41.7% female, $M_{\text{age}} = 31.4$ years) were recruited to serve as review writers. The writing experiment employed a 2 (product type: search or experience) $\times$ 2 (attentional focus: self or others) between-subjects design. Writers were assigned to review either a search product (e.g., portable battery chargers, game consoles) or an experience product (e.g., shoes, pants) that they had recently purchased. They were then provided varying writing prompts depending on which attentional focus condition they were assigned to:

Self: *"The purpose of writing a product review is to express yourself. Experts have noted that the best product reviews are those that are written with your own product needs in mind. Please write a brief review that focuses on how the product did (or did not) satisfy your personal needs."*

Others: *"The purpose of writing a product review is to help other people. Experts have noted that the best product reviews are those that are written with other people's product needs in mind. Please write a brief review that focuses on the likely questions or concerns that other shoppers might have about the product."*

6.2. Measures

After completing their review, writers were asked to predict how helpful their review would be to readers: *"Imagine that 100 people read your product review. How many do you think would rate it as 'helpful' (versus 'unhelpful')? Please type a number between 0 and 100."*

To check the manipulation of product type, writers were provided with a 7-point Likert item (1 = Strongly disagree and 7 = Strongly agree), *“The product I reviewed is one that a person needs to evaluate in person.”* Greater agreement with this item indicated that the product was perceived to be an experience type rather than a search type. An independent samples t-test confirmed the success of the manipulation ($M_{\text{Experience}} = 4.83$ versus $M_{\text{Search}} = 3.58$, $t_{22} = 2.31$, $p = .03$). To check the manipulation of attentional focus, writers were asked whether they focused on their personal needs or the needs of other shoppers when writing. Their responses were coded as correct or incorrect based on their experimentally assigned condition. A binomial test confirmed the success of the manipulation, with the rate of correct responses significantly exceeding the rate one would expect if the writers were guessing (87.0% versus 50.0%; $z = 3.69$, $p < .001$).

6.3. Results

As shown in Table 1 and Figure 1, the writers predicted that their reviews would be helpful to a large majority of potential readers ($M_{\text{Overall}} = 68.8\%$). Additionally, the writers' predictions did not differ by product type ($F_{1,18} = 0.01$, $p = .91$), attentional focus ($F_{1,18} = 0.05$, $p = .82$), or the interaction of the two ($F_{1,18} = 0.01$, $p = .94$). These results indicate that writers are broadly confident in their ability to write helpful reviews, thus setting the stage for a comparison with readers' perceptions. Study 2 was designed to enable such a comparison and to test the research hypotheses.

Table 1. Model for review-writers' self-predicted helpfulness

	Sum of Squares	df	Mean Square	F	p
Product Type	1.50	1	1.50	0.003	0.95
Attentional Focus	37.50	1	37.50	0.080	0.78
Product Type * Attentional Focus	8.17	1	8.17	0.018	0.90
Residuals	9323.33	20	466.17		



Figure 1. Results for review-writers' self-predicted helpfulness

7. Study 2: Readers' predictions about the helpfulness of the review-writers

7.1. Participants, Design, and Procedure

The reviews written by the participants in Study 1 were shown to a second group of 172 participants (readers) in Study 2 (45.9% female, $M_{\text{age}} = 33.0$ years). These readers were randomly assigned to one of the products reviewed by the writers. For their assigned product, readers saw three different reviews and were instructed to, "Read the three reviews below and imagine you are actually shopping for each product." Thus, the effective sample size of reviews in Study 2 is $3 \times 172 = 516$. Since the readers were assigned to the reviews written by the writers, Study 2 inherits the experimental design of Study 1 with respect to the stimuli.

7.2. Measures

After seeing each of the three reviews for a second time, the readers in Study 2 were first asked, "If 100 other shoppers read this review, how many do you think would rate it as 'helpful' (versus 'unhelpful')? Please type a number between 0 and 100." They were then asked, "Was this review helpful to you?" (Yes or No). To check the manipulation of product type (from the perspective of the readers), a 7-point Likert item was provided (1 = Strongly disagree and 7 = Strongly agree), "*The product I reviewed is one that a person needs to evaluate in person.*" An independent samples t-test confirmed the success of the manipulation ($M_{\text{Experience}} = 4.71$ versus $M_{\text{Search}} = 3.39$, $t_{607} = 12.4$, $p < .001$).

7.3. Results

The analysis of the Study 2 data examines how helpful the writers' reviews were, as indicated by readers' personal assessments (i.e., the true helpfulness of the reviews). This allows for an assessment of the accuracy of the writers' predictions about the helpfulness of their reviews. Overall, the writers were overconfident. Whereas Study 1 revealed that writers predicted that their reviews would be helpful for 68.8% of readers, readers only rated their reviews as helpful in 60.7% of cases.

Since the dependent variable is binary ("*Was this review helpful to you?*" Yes or No), the data were submitted to a generalized linear mixed model with a logit link function. As in the first analysis, the fixed effects are product type, attentional focus, and the interaction of the two. Additionally, a random effect (on the model intercept) was included to account for the individual reviewers. The model provided a good fit to the data, with a marginal $R^2 = .158$ and conditional $R^2 = .175$. The Intra-Class Correlation was .020. Table 2 displays the results of the model estimation, while Figure 2 displays the estimated marginal means by experimental condition.

Table 2. Model for readers' personal assessments of review helpfulness

Names	Estimate	SE	Z	p
(Intercept)	-0.43	0.10	-4.25	< 0.001
Product Type (0 = Search, 1 = Experience)	-0.84	0.20	-4.11	< 0.001
Attentional Focus (0 = Self, 1 = Others)	1.40	0.21	6.66	< 0.001
Product Type * Attentional Focus	-0.08	0.04	-0.21	0.837

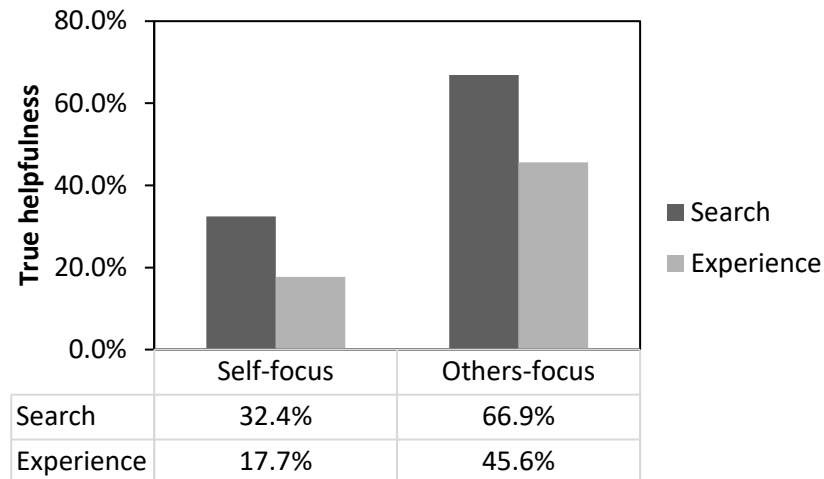


Figure 2. True helpfulness of writers' reviews

Consistent with H_1 , there was a significant main effect of attentional focus such that a higher percentage of readers found writers' reviews more personally helpful when writers' attentional focus was on others rather than the self (Others-focus: 56.6% vs. Self-focus: 24.3%; $B = 1.40$, $z = 6.66$, $p < 0.001$). Simple contrasts revealed that the advantage for others-focused reviews existed for both search products (others-focus: 66.9% vs. self-focus: 32.4%, $z = 5.16$, $p < 0.001$) and experience products (others-focus: 45.6% vs. self-focus: 17.7%, $z = 4.56$, $p < 0.001$).

Consistent with H_2 , there was a significant main effect of product type such that a higher percentage of readers found writers' reviews more personally helpful for search products than experience products (search: 49.6% vs. experience: 29.8%; $B = -0.84$, $z = -4.11$, $p < 0.001$). Simple contrasts revealed that the advantage for search product reviews occurred whether writers' attentional focus was on the self (search: 32.4% vs. experience: 17.7%, $z = 2.74$, $p = 0.006$) or on others (search: 66.9% vs. experience: 45.6%, $z = 3.18$, $p = 0.001$).

Contrary to H_3 , the interaction between attentional focus and product type was not significant ($B = -0.08$, $z = -0.21$, $p = 0.837$). That is, the effects of attentional focus and product type on perceptions of review helpfulness are additive.

8. Discussion and future research

Across two studies, the evidence suggests that while review writers are overconfident, their reviews are most helpful when their attentional focus during writing is on others (versus themselves) and when reviewing products characterized predominately by search (vs. experience) qualities. These findings hold managerial implications, as reviews that are more helpful are more valuable and influential for consumers. Further, retailers can benefit from being able to anticipate which reviews are likely to be helpful and promote them to the.

The present research also holds implications for review platforms since helpful reviews are "stickier" for consumers and enhance engagement with the platforms (Mudambi and Schuff 2010). Retailers and review platforms can also consider incorporating attentional focus as a writing prompt. Additionally, advances in text-based analyses (e.g., natural language processing) allow for large-scale, automated

detection and analysis of drivers of helpfulness such as attentional focus and product type.

This research also presents several opportunities for future research. First, the two studies reported here involved reviews for selected product categories. More product categories should be considered to further test the usefulness and scope of the search-experience paradigm. Second, attentional focus was varied by prompting writers to focus on others or on the self. Less obtrusive prompts may be of interest, such as encouraging writers to use first-person pronouns (to encourage self-focus) or encouraging writers to only use second-person pronouns (to encourage adoption of an others-focus). Third, future research is needed to explore additional outcome variables beyond perceptions of the helpfulness of a review. For example, brands and retailers may be interested in understanding the impact of writers' attentional focus on product-related beliefs, attitudes, and intentions to purchase.

9. References

- Baek, H., Ahn, J., & Choi, Y. (2012). Helpfulness of online consumer reviews: Readers' objectives and review cues. *International Journal of Electronic Commerce*, 17(2), 99-126.
- Bronner, F., & De Hoog, R. (2011). Vacationers and eWOM: who posts, and why, where, and what?. *Journal of travel research*, 50(1), 15-26.
- Brunel, F. and Swain, S. (2007) A moderated perceptual model of product aesthetic evaluations, *European Advances in Consumer Research*, Vol. 8, 444-445.
- Campbell, M. C., & Kirmani, A. (2000). Consumers' use of persuasion knowledge: The effects of accessibility and cognitive capacity on perceptions of an influence agent. *Journal of consumer research*, 27(1), 69-83.
- Carlson, J. P., Weathers, D. and Swain, S. D. (2016), Consumer responses to bonus pack and product enlargement claims, *Journal of Marketing Theory and Practice*, 24(1), 59-71.
- Chai, S., Choi, B., Kim, M., & Cheng, T. C. E. (2023). Why do people speak about products online? The role of opinion leadership. *Information Technology and Management*, 24(1), 1-17.
- Cheng, X., Fu, S., Sun, J., Bilgihan, A., & Okumus, F. (2019). An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust. *Tourism Management*, 71, 366-377.
- Cui, G., Lui, H. K., & Guo, X. (2012). The effect of online consumer reviews on new product sales. *International Journal of Electronic Commerce*, 17(1), 39-58.
- Dellarocas, C., Gao, G., & Narayan, R. (2010). Are consumers more likely to contribute online reviews for hit or niche products?. *Journal of Management Information Systems*, 27(2), 127-158.
- Dixit, S., Badgaiyan, A. J., & Khare, A. (2019). An integrated model for predicting consumer's intention to write online reviews. *Journal of Retailing and Consumer Services*, 46, 112-120.
- Friestad, M., & Wright, P. (1994). The persuasion knowledge model: How people cope with persuasion attempts. *Journal of consumer research*, 21(1), 1-31.
- Galinsky, A. D., and Moskowitz, G. B. 2000. Perspective-Taking: Decreasing Stereotype Expression, Stereotype Accessibility, and in-Group Favoritism, *Journal of Personality and Social Psychology* 78(4), pp. 708-724.
- Goldstein, N. J., Vezich, I. S., & Shapiro, J. R. (2014). Perceived perspective taking: When others walk in our shoes. *Journal of Personality and Social Psychology*, 106(6), 941-960.
- Gonçalves, H. M., Silva, G. M., & Martins, T. G. (2018). Motivations for posting online reviews in the hotel industry. *Psychology & Marketing*, 35(11), 807-817.
- Hanna, R. C., Swain, S. D., & Smith, S. (2016). Email marketing in a digital world: The basics and beyond. Business Expert Press.
- Hong, H., Xu, D., Wang, G. A., & Fan, W. (2017). Understanding the determinants of online review helpfulness: A meta-analytic investigation. *Decision Support Systems*, 102, 1-11.
- Hu, N., Zhang, J., & Pavlou, P. A. (2009). Overcoming the J-shaped distribution of product reviews. *Communications of the ACM*, 52(10), 144-147.
- Ickes, W., Reidhead, S., & Patterson, M. (1986). Machiavellianism and self-monitoring: As different as "me" and "you". *Social Cognition*, 4(1), 58-74.

- Jiménez, F. R., & Mendoza, N. A. (2013). Too popular to ignore: The influence of online reviews on purchase intentions of search and experience products. *Journal of Interactive Marketing*, 27(3), 226-235.
- Lei, Z., Yin, D., & Zhang, H. (2021). Focus within or on others: The impact of reviewers' attentional focus on review helpfulness. *Information Systems Research*, 32(3), 801-819.
- King, R. A., Racherla, P., & Bush, V. D. (2014). What we know and don't know about online word-of-mouth: A review and synthesis of the literature. *Journal of interactive marketing*, 28(3), 167-183.
- Lafky, J. (2014). Why do people rate? Theory and evidence on online ratings. *Games and Economic Behavior*, 87, 554-570.
- Korschun, D., Bhattacharya, C. B., & Swain, S. D. (2016). CSR and the frontline context: How social programs improve customer service. *NIM Marketing Intelligence Review*, 8(1), 24.
- Mudambi, S. M., & Schuff, D. (2010). Research note: What makes a helpful online review? A study of customer reviews on Amazon. com. *MIS quarterly*, 185-200.
- Munzel, A., & H. Kunz, W. (2014). Creators, multipliers, and lurkers: who contributes and who benefits at online review sites. *Journal of Service Management*, 25(1), 49-74.
- Mittal, B. (2004). Lack of attribute searchability: Some thoughts. *Psychology & Marketing*, 21(6), 443-462.
- Nelson, P. (1974). Advertising as information. *Journal of political economy*, 82(4), 729-754.
- Pennebaker, J. W., Mehl, M. R., & Niederhoffer, K. G. (2003). Psychological aspects of natural language use: Our words, our selves. *Annual review of psychology*, 54(1), 547-577.
- Sahoo, N., Dellarocas, C., & Srinivasan, S. (2018). The impact of online product reviews on product returns. *Information Systems Research*, 29(3), 723-738.
- Singh, J. (1988). Consumer complaint intentions and behavior: definitional and taxonomical issues. *Journal of Marketing*, 52(1), 93-107.
- Smith, D., Menon, S., & Sivakumar, K. (2005). Online Peer and Editorial Recommendations, Trust, and Choice in Virtual Markets. *Journal of Interactive Marketing*, 19(3), 15-37.
- Statista. (March 7, 2023). Attitudes towards online shopping in the U.S. in 2022
- Sun, M. (2012). How Does the Variance of Product Ratings Matter. *Management Science*, 58(4), 696-707.
- Sun, X., Han, M., & Feng, J. (2019). Helpfulness of online reviews: Examining review informativeness and classification thresholds by search products and experience products. *Decision Support Systems*, 124, 113099.
- Swain, S. D., Berger, P. D., & Weinberg, B. D. (2014). The customer equity implications of using incentives in acquisition channels: A nonprofit application. *Journal of Marketing Analytics*, 2(1), 1-17.
- Swain, S. D. (2022). An exploratory study of the effect of in-store recommendation technology on wine shoppers' search behaviors, *International Journal of Business Research*, 22(2), 124-137.
- Tractinsky, N., & Rao, V. S. (2001). Incorporating Social Dimensions in Web-Store Design. *Human Systems Management*, 20(2), 105-121.
- Voight, J. (2007). Getting a Handle on Customer Reviews. *Adweek*, 48(26), 16-17.
- Weathers, D., Swain, S. D., and Carlson, J. P. (2012) Why consumers respond differently to absolute versus percentage descriptions of quantities, *Marketing Letters*, 23(4), 943-957.
- Weathers, D., Swain, S.D., & Grover, V. (2015). Can online product reviews be more helpful? examining characteristics of information content by product type. *Decision Support Systems*, 79(November), 12-23.
- Wu, P. F. (2013). In search of negativity bias: An empirical study of perceived helpfulness of online reviews. *Psychology & Marketing*, 30(11), 971-984.
- YouGov, (February 2, 2022). Share of consumers who regularly use online reviews to help with homeware purchase decisions worldwide in 2021, by country.

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

Internal and External Precursors to Favorable Adoption of Industry 4.0 Technologies

Andrew Couch¹

Nicholas Loyd¹

*¹Department of Industrial & Systems Engineering and Engineering Management
The University of Alabama in Huntsville*

andrew.c.couch@gmail.com; loydn@uah.edu

Abstract

In the modern era of rapidly emerging technologies, organizations are faced with pressing concerns about how to strategically maneuver their trajectory and corresponding operations to accommodate the rapidly changing environment. Associated with this dilemma is a dynamic perception of importance and urgency. Just as the physical environment is capable of guiding human behavioral tendencies, the competitive climate of the marketplace as well as internal organizational tendencies hold influential power over how organizations innovate over time. In the context of Industry 4.0 adoption, the benefits for organizations are not universally identical. Rather, the benefits are situationally associated with operational trends of the surrounding industry space. Thus, this publication offers a discussion on a subset of the research findings pertaining to Industry 4.0 adoption ventures. For Industry 4.0 considerations, the first hurdle that confronts organizations is the decision about how extensively to adopt Industry 4.0 technologies, if at all. As such, an urgency-importance matrix is considered from a reliability context to help organizations identify favorable circumstances that are generally associated with higher levels of Industry 4.0 adoption necessity and success. Monitoring the competitive environment is essential to tailoring organizational strategies. The Industry 4.0 dimension plays an instrumental role in this area.

Keywords: Industry 4.0, Organizational Strategy, and Technology Innovation

1. Introduction and motivation

Observably, technology innovation throughout history has occurred in patterned waves that shape the technology landscape of the time. Moreover, the nature of technology brought about by any given innovation wave serves to establish the parameters that govern the following wave of technology innovation. For organizations, monitoring the latest innovative developments are a vital component of strategy and organizational excellence. Noticeably, this value becomes remarkably transparent when considering two broad archetypes of strategy known as blue ocean strategy and red ocean strategy (Kim, 2005). Blue ocean strategies are often characterized as finding success in competitive markets by leveraging a significant innovation that offers capabilities which are new and untested in the market (Kim et al, 2014). In this manner, organizations are able to acquire new market space instead of competing in established spaces. On the other hand, red ocean strategies focus on establishing organizational success through enacting marginal improvements to products and services in order to

gain competitive standing in established markets (Lainos, 2011). In both cases, monitoring and leveraging technology innovation are critical. In one case, innovation is the most pivotal ingredient of success in the marketplace. In the other case, innovation is a necessary condition in order to gain standing amongst an existing market.

As characterized by the current literature, several waves of technology innovation have occurred over time. These waves have been studied and are more commonly known as industrial revolutions. Modern economic and technological phenomena suggests that the fourth prevalent wave of technology innovation is currently unfolding. These developments and the technologies associated with them are generally known as Industry 4.0. Analogously to other technological developments, the benefits and drawbacks of new capabilities are not uniform across all markets and organizations. Thus, all organizations face an important dilemma when confronted with the decision about whether or not to pursue technology adoption and implementation in accordance with the latest pioneering developments. Observably, there are discrepancies throughout society on the basis of information access, technological acquisition access, and knowledgeable talent. Moreover, the decision framework that serves as the structure of this dilemma is not transparent to the everyday organization. Thus, this research seeks to offer insights and characterizations about how such a decision should be approached in accordance with modern technological developments and internal organizational conditions.

2. Literature assessment

Although literature definitions for Industry 4.0 differ in their philosophical approaches, the undertones of commonality are best embodied by describing Industry 4.0 as “beget with emergent and disruptive intelligence and information technologies” (Bai et al, 2020). Conversely, other publications characterize Industry 4.0 by a specific set of interconnected layers including business, functional, information, communication, integration, and asset (Zezulka et al, 2016). Regardless of whether Industry 4.0 is defined from the internal organizational view or the external market developments that unfold, the essence of a technological innovation wave and rapid transformation remains consistent.

Although the literature describes the benefits of Industry 4.0 as situationally dependent, there are certain elements of commonality that have been identified. Benefits such as advanced planning and control, higher quality, increased productivity, enhanced customization capabilities, reduced waste, and safer working conditions are all recognized benefits of implementing Industry 4.0 in organizations (Mohamed, 2018). Naturally, benefits of Industry 4.0 extend to all major market sectors (Thames et al, 2017) although the manner in which benefits are leveraged may vary.

Conversely, there are also challenges that are introduced by Industry 4.0. Most notably, these challenges often emerge from a deficiency on the basis of personal talent or resource limitations. Research (Schröder, 2016) often cites a lack of digital strategy, resource scarcity, poor data security, and lack of standards as the primary pressing concerns for small and medium-sized enterprises. Other research (Kiel et al, 2017) suggests that challenges centered upon technical integration, organizational transformation, and data security. Analogously to the advantages, Industry 4.0 poses challenges on a situational basis, although challenges are uniquely concentrated in small and medium-sized enterprises (Stentoft, 2021).

3. Industry 4.0 urgency and importance dynamic

At the crux of the Industry 4.0 dilemma stands an assessment about the importance of Industry 4.0 adoption for a given organization. While it is known that Industry 4.0 offers varying advantages and disadvantages that are dependent upon the situational circumstances of an organization, there is no formal methodology for assessing the importance of Industry 4.0 relative to a particular organization (Arnold et al, 2018). Likewise, when managers and corporate leaders pursue such an assessment, the interpretation of the internal organizational circumstances as well as the surrounding environmental market conditions are observably influenced by the urgency of signaling. The decision landscape becomes the leading challenge in these scenarios. In essence, certain circumstances have a tendency to signal a false sense of urgency while other types of conditions fail to signal enough urgency for change. As such, this complicates the decision concerning whether or not to pursue Industry 4.0 adoption. Moreover, this decision appropriately departs from a boolean choice.

As such, an analysis was performed using several research methods. First, a literature review was conducted. The review method is traditional narrative and draws upon 137 publications occurring with key search terms such as “Industry 4.0”, “technology precursor”, and “business model”. The two scholarly databases applied were Google Scholar and Scopus Search. Duplicating results were filtered and eliminated. Likewise, irrelevant results were excluded from the analysis. In surplus of this, expert opinions from industry were gathered on an informal basis to supplement the research findings.

In order to address these collective concerns, this research puts forth a version of the Eisenhower matrix. Although it is traditionally used for personal task management, the crossover application of this tool offers a unique approach to dissecting the turbulent technological environment of the modern era. To produce a higher resolution of analysis, the Eisenhower matrix is applied twice: once to diagnose internal circumstances and once to diagnose external conditions as they relate to Industry 4.0. From the basis of these collective observations, Figure 1 depicts the emergent Eisenhower matrix relative to the most current conditions of Industry 4.0 technology for organizations.

Self Diagnosis: How Important and Urgent is the Adoption of Industry 4.0?		
	Urgent	Not Urgent
Important	1. Technology capability deficiency to serve customers 2. Widespread and ineffective scheduling or timing of operations 3. Wasteful operational practices	1. Lack of long-term product/service feature innovation 2. Lack of opportunistic action to improve processes 3. Failure to sustain improvement solutions to operations
Not Important	1. Occasional re-centering of automated processes 2. Trivial engineering design solutions 3. Struggles with internal support value streams	1. Technology may not improve product/service deficiencies 2. Technology may not improve lapses in high-level strategy 3. If culture and not technology is the leading struggle

Figure 1. Urgency/Importance Matrix for Internal Operations

Contrary to the internal nature and circumstances of organizations, an assessment of external conditions must take into account a variety of dynamic factors that may emerge from a seemingly infinite number of entities that are capable of inducing some amount of market change. However, the most critical influential elements that carry implications for Industry 4.0 adoption most commonly exist within the supply chain or the demands of customers (Fatorachian et al, 2021). Appropriately, the primary external cues that provide a meaningful signal to suggest favorable Industry 4.0 adoption should emerge from these sources. Thus, Figure 2 depicts the Eisenhower matrix relative to the most critical cues that signal Industry 4.0 adoption in accordance with modern tendencies.

Self Diagnosis: How Important and Urgent is the Adoption of Industry 4.0?	Urgent	Not Urgent
Important	1. If organizations upstream and downstream achieved full adoption 2. If customers demand value only offered by Industry 4.0 technology 3. Automated functions are routinely tweaked due to the environment	1. Upstream/downstream proposals to make cohesive supply networks 2. Product operational context begins to deviate from past conditions 3. If population skillsets shift to mental tasks and not manual labor
Not Important	1. If situational differences exist in capabilities compared to suppliers 2. The prospect of improvement 3. Routine complex solutions that cannot be replicated by technology	1. Value is delivered to the customer that only human labor can provide 2. If human-machine interaction harms customer satisfaction 3. If added value cannot be identified by involving new technology

Figure 2. Urgency/Importance Matrix for External Conditions

As illustrated, it is essential for organizations to identify external conditions that are both important and urgent towards implementing Industry 4.0. Specifically, these findings suggest that external conditions that are both important and urgent manifest primarily in the supply chain and customer basis. If organizations upstream and downstream implement Industry 4.0, there is an elevated emphasis on Industry 4.0 adoption in order to sharpen the cohesive stream of real-time data that may be transited through every link. Likewise, an apparent condition that signals strong importance and urgency exists when customers demand capabilities that only Industry 4.0 can deliver. On the other hand, comparable cases where similar urgency is conveyed without the backing of importance would include trivial (although salient) differences in technological capabilities compared to suppliers, complex and creative solutions that may not be solved strictly through increased technical sophistication, and the prospect of operational improvement due to attractive features of Industry 4.0. When considering conditions that are both important and not urgent, it is essential to recognize that Industry 4.0 is both supported and constrained by the talents of the populace. Slow moving phenomena such as shifts in the talents of the populace, new operational contexts for products, and preliminary plans to create a cohesive and data-driven supply network are all circumstances that embody importance without urgency for a given organization to adopt Industry 4.0.

4. Derived Industry 4.0 action plans

In a comparable manner to personal task management, action plans for Industry 4.0 implementation in organizations may be derived from the insights that are offered by the Eisenhower matrix. Notably, the theoretical development of the Eisenhower matrix framework for personal task management includes a hierarchy of task priority that should be implemented (Jyothi et al, 2016). As shown below in Figure 3, tasks that carry both importance and urgency should be addressed first. In terms of prioritization, tasks carrying importance with the absence of urgency should possess higher priority over tasks that carry urgency without the backing of importance.

	NOT URGENT	URGENT
IMPORTANT	B TASKS Issues important, but not urgent, that can wait, but require constant attention. You should plan, implement and regularly check the degree of implementation.	A TASKS Issues important and urgent to be taken up immediately personally because the consequences can be serious.
DOES NOT MATTER	D TASKS Cases not urgent and not important. Such tasks can be avoided	C TASKS Urgent but not important, to be carried out by someone else, or to be carried out immediately to be done and over with. Such tasks should be delegated.

Figure 3. Eisenhower Matrix Action Prioritization (Gajewska et al, 2017)

Appropriately, this framework encourages stability in long-term performance. Above all else, this serves to help structure strategic plans using a layered approach. Thus, the core elements of a strategic plan may be constructed from such an approach in a manner that is consistent with the appropriate timeline of action. At the forefront of these plans exist addressing Industry 4.0 technology adoption conditions that are both important and urgent. Amid this centerpiece of Industry 4.0 strategy, our findings suggest that organizations should apply the Scrum approach to rapidly implement Industry 4.0 principles, identify internal resources to accommodate workers in a shift towards intellectually-oriented jobs, and enact short-term plans for implementation of specific Industry 4.0 technologies.

This research finds that the corresponding action plans for Small and Medium-Sized Enterprises (SMEs) are more complex than for more established organizations. Uniquely, SMEs must construct plans by applying a prioritization to each precursor criteria for Industry 4.0 adoption. The traditional narrative literature review finds that SMEs are rarely able to pursue each precursor in a simultaneous fashion. As such, the most common starting point for SMEs in the journey to adopting Industry 4.0 is a needs assessment. Industry 4.0 must become a necessary ingredient for achieving the mission, vision, values, and goals of the organization, but it is not a necessary condition. SMEs may use the diagnostic tool in Figure 1 and Figure 2 to dissect the likelihood of relevance of Industry 4.0 to organizational strategy.

Outside of this core component of Industry 4.0 strategy exists conditions that carry importance for adopting Industry 4.0 technologies while possessing no urgency. In these situations, it is often advantageous to use these observations to make amends to the long-term strategy. If adopting a particular aspect of Industry 4.0 is important but does not possess short-term urgency, then an appropriate action plan would be to amend the long-term strategic plan in accordance with directed

growth towards adopting the particular aspect of Industry 4.0. Additionally, situations such as this impose newfound criticality on training and workforce development activities. Enacting short-term pivots to enhance training serves to favorably support the amended long-term strategy for Industry 4.0. Finally, another derived action plan for conditions that are important but not urgent involves searching for suppliers who are willing to undergo an Industry 4.0 transition. Industry 4.0 thrives when the supply chain is linked together with real-time sharing of data and operational statuses. Thus, this action plan seeks to establish strategic posturing that accommodates the situational importance of Industry 4.0 adoption.

Conversely, organizations may experience the stated scenarios corresponding to conditions which are not important towards adopting Industry 4.0, but they possess an element of urgency. In these situations, it is best to construct organizational strategy to include and implement root cause analysis targeted towards assessing operational issues in order to determine the necessary technical sophistication of a solution. This helps to dissect the true magnitude of importance for adopting Industry 4.0 technologies without the influence that urgent or salient conditions may impose. Also, another core action plan exists in constructing strategies to handle urgent conditions with no backing of importance. Relatedly, this may tie back to root cause analysis or may be integrated into a broader problem-solving framework for addressing short-term scenarios that convey urgency.

On the other hand, if an organization experiences conditions that correspond to conditions that carries neither importance nor urgency, our findings suggest that continuous improvement is the best facilitator to establish the groundwork for best identifying important conditions for Industry 4.0 adoption when they arise in the future. Moreover, adapting organizational strategy to include market analysis metrics on technology adoption patterns helps to facilitate better internal detection of important and urgent conditions as it relates to Industry 4.0. Cohesively, considering the totality of these potential conditions relating to Industry 4.0 are essential when constructing organizational strategy in accordance with modern technological developments. This research establishes a direct connection between observed internal or external conditions and the appropriate structuring of organizational strategy for technology.

5. Implementation considerations

The previous section established strategic action plans that serve to incorporate Industry 4.0 into corporate strategy in strict accordance with observed internal or external phenomena. Now, it is important to consider several implementation considerations that should be enacted in order to favorably support this type of high-level strategy. Most prominently, Industry 4.0 is best supported in environments where lean principles have already been implemented under continuous improvement (Shahin et al, 2020). In essence, Industry 4.0 technology often plays an elevated role during implementation when these technologies are used to further drive continuous improvement on established operations. Thus, Industry 4.0 technologies may be used as just another avenue for sharpening individual processes and operations that have already been optimized for under the bounds of human labor speeds and productivity. This shows the dynamic ability of Industry 4.0 technologies to be used in many aspects of operations throughout the supply chain. When implemented, Industry 4.0 technologies would bypass the bounds on productivity that may be imposed by the physical constraints

of manual labor.

An additional implementation consideration exists in the knowledgeable talent that identifies areas for optimal application of Industry 4.0, implements the technology, and sustains implementation of the desired time horizon in the strategic plan. This research finds that the acquisition of sophisticated technical knowledge is a barrier that is most strongly prohibitive for SMEs. Prior to embarking towards technology implementation, organizations must support the directed actions of the organizational strategic plan with talent in the proper technological domains. Moreover, the dynamic of long-term change becomes critical in this context. If graduate internal or external change appropriately result in adaptations to the strategic plan, this must carry repercussions to the pursuit of talent to support future strategic actions.

Externally, market analytics assist in Industry 4.0 technology implementation by uncovering the unique connection that exists between the posturing of other competing organizations and the observed outcomes of their Industry 4.0 technology implementations for those competing entities. The complex nature of such relationships are often best estimated by experienced individuals. As a whole, implementation should be guided by phenomena as they unfold in the market. In this manner, the favorability of Industry 4.0 implementation changes in a dynamic fashion. Despite this change, this research offers a methodology for assessing the favorability of Industry 4.0 technology adoption from on the basis of internal and external factors.

6. Areas of future research

Given the results of the findings presented here, a number of intriguing areas of future research arise. First, future research is needed on how specific internal organizational conditions influence Industry 4.0 adoption levels. Although observations based on correlation may be drawn, the causality effects should be better characterized through proper research investigations. Specifically, root cause analysis may help to diagnose these types of underlying relationships. Survey biases may help to obtain broad generalizations, but regional biases often unfavorably skew the results and corresponding implications.

An additional point of future research should examine how human labor should be accommodated within organizations that choose to undergo Industry 4.0 transitions. Observably, technological change introduces widespread changes in requirements that organization must adapt to. This mandates an internal transformation in order to supply the adequate talent and managerial structures to address such capabilities. In times of rapid or significant transition, organizations may establish employee exchange agreements in order to minimize the impact of displaced human labor. In this manner, human labor skills that become obsolete to one organization due to technological innovation may be supported by another organization. The formal details of such a transition present a unique area of future research that should be examined under particular industry tendencies and best practices relative to Industry 4.0.

7. Concluding remarks

The findings presented here help to characterize the action plans that should be derived based on internal organizational conditions and external market circumstances. Observably, this analysis helps to discern and separate the common signal and noise that is present internally amongst organizations.

The intuitive route to addressing what may seem to be a pressing issue is not always the most beneficial within the context of Industry 4.0 adoption. Urgency and importance are assessed in the context of Industry 4.0 adoption favorability while also accounting for internal and external conditions for an organization within a competitive market.

From this basis, the Eisenhower matrix allows actions to be prioritized. The principles that the Eisenhower matrix offers in the world of personal task management also assist in guiding organizational strategy and technological implementation ventures. Importantly, this approach to analyzing Industry 4.0 offers an avenue for constructing organizational strategy. Moreover, considering that the current literature on Industry 4.0 recognizes the lack of a digital strategy as a primary drawback, this research offers support for everyday organizations to derive digital strategy from metrics that are most familiar to managers and corporate leaders. Motivated by the pursuit to clarify organizational strategy amid a turbulent technological landscape, this research delivers a concise methodology for obtaining actions plans based on the situational circumstances that challenge an organization from both an internal perspective and an external competitive market perspective.

8. References

- Kim, W. C. (2005). Blue ocean strategy: from theory to practice. *California management review*, 47(3), 105-121.
- Kim, W. C., & Mauborgne, R. (2014). Blue ocean strategy, expanded edition: How to create uncontested market space and make the competition irrelevant. Harvard business review Press.
- Lainos, I. (2011). Red ocean vs blue ocean strategies (Doctoral dissertation, University of Piraeus (Greece)).
- Bai, C., Dallasega, P., Orzes, G., & Sarkis, J. (2020). Industry 4.0 technologies assessment: A sustainability perspective. *International journal of production economics*, 229, 107776.
- Zezulka, F., Marcon, P., Vesely, I., & Sajdl, O. (2016). Industry 4.0—An Introduction in the phenomenon. *IFAC-PapersOnLine*, 49(25), 8-12.
- Mohamed, M. (2018). Challenges and benefits of industry 4.0: An overview. *International Journal of Supply and Operations Management*, 5(3), 256-265.
- Thames, L., & Schaefer, D. (2017). Industry 4.0: an overview of key benefits, technologies, and challenges. *Cybersecurity for Industry 4.0: Analysis for Design and Manufacturing*, 1-33.
- Schröder, C. (2016). The challenges of industry 4.0 for small and medium-sized enterprises. Friedrich-Ebert-Stiftung: Bonn, Germany.
- Kiel, D., Müller, J. M., Arnold, C., & Voigt, K. I. (2017). Sustainable industrial value creation: Benefits and challenges of industry 4.0. *International journal of innovation management*, 21(08), 1740015.
- Stentoft, J., Adsbøll Wickstrøm, K., Philipsen, K., & Haug, A. (2021). Drivers and barriers for Industry 4.0 readiness and practice: empirical evidence from small and medium-sized manufacturers. *Production Planning & Control*, 32(10), 811-828.
- Arnold, C., Veile, J., & Voigt, K. I. (2018, April). What drives industry 4.0 adoption? An examination of technological, organizational, and environmental determinants. In *Proceedings of 27th Annual Conference of the International Association for Management of Technology*, Birmingham, United Kingdom, April (pp. 22-26).
- Fatorachian, H., & Kazemi, H. (2021). Impact of Industry 4.0 on supply chain performance. *Production Planning & Control*, 32(1), 63-81.
- Jyothi, N. S., & Parkavi, A. (2016, May). A study on task management system. In *2016 International Conference on Research Advances in Integrated Navigation Systems (RAINS)* (pp. 1-6). IEEE.
- Gajewska, P., & Piskrzyńska, K. (2017, March). Leisure time management. In *Forum Scientiae Oeconomia*, 5(1), pp. 57-69).
- Shahin, M., Chen, F. F., Bouzary, H., & Krishnaiyer, K. (2020). Integration of Lean practices and Industry 4.0 technologies: smart manufacturing for next-generation enterprises. *The International Journal of Advanced Manufacturing Technology*, 107, 2927-2936.

Copyright of the Journal of Management and Engineering Integration is the property of the Association of Industry, Engineering and Management Systems Inc., and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

SUMMER
2023



AIEMS

AIEMS is committed to furthering the integration between business management, engineering, & industry. Our objective is to promote research, collaboration, & practice in these multidisciplinary areas. AIEMS seeks to encourage local, national, & international communication & networking via conferences & publications open to those in both academia & industry. We strive to advance professional interaction & lifelong learning via human & technological resources, and to influence and promote the recruitment and retention of young faculty and industrialists.

AIEMS
VOL 16
NO 1