2020

Search Results: Predicting Ranking Algorithms With User Ratings and User-Driven Data

Gary Michael Taylor
Walden University

Walden University

College of Management and Technology

This is to certify that the doctoral study by

Gary Michael Taylor

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Bob Duhainy, Committee Chairperson, Information Technology Faculty
Dr. Charlie Shao, Committee Member, Information Technology Faculty
Dr. Steven Case, University Reviewer, Information Technology Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2020

Abstract

Search Results: Predicting Ranking Algorithms With User Ratings and User-Driven Data

by

Gary Michael Taylor

MS, West Virginia University, 2010

BS, Marshall University, 2006

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Information Technology

Walden University

February 2020

Abstract

The purpose of this correlational quantitative study was to examine the possible relationship between user-driven parameters, user ratings, and ranking algorithms. The study's population consisted of students and faculty in the information technology (IT) field at a university in Huntington, WV. Arrow's impossibility theorem was used as the theoretical framework for this study. Complete survey data were collected from 47 students and faculty members in the IT field, and a multiple regression analysis was used to measure the correlations between the variables. The model was able to explain 85% of the total variability in the ranking algorithm. The overall model was able to significantly predict the algorithm ranking discounted cumulative gain, $R^2 = .852$, $F(3,115) = 220.13$, $p < .01$. The Respondent DCG and Search term variables were the most significant predictor with $p = .0001$. The overall findings can potentially be useful to content providers who focus their content on a specific niche. The content created by these providers would most likely be focused entirely on that subgroup of interested users. While it is necessary to focus content to the interested users, it may be beneficial to expand the content to more generic terms to help reach potential new users outside of the subgroups of interest. User's searching for more generic terms could potentially be exposed to more content that would generally require more specific search terms. This exposure with more generic terms could help users expand their knowledge of new content more quickly.

Search Results: Predicting Ranking Algorithms With User Ratings and User-Driven Data

by

Gary Michael Taylor

MS, West Virginia University, 2010

BS, Marshall University, 2006

Doctoral Study Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Information Technology

Walden University

February 2020

Table of Contents

List of Tables

List of Figures

Section 1: Foundation of the Study

Background

The web has quickly become a part of people's everyday life and the utilization of search engines is a key part of this (Schroeder, 2014). Search engines remain somewhat generic in how sites are ranked, meaning each user searching for a term will get the same result set (Luh & Huang, 2016). While this can provide the users with the information they need, it is often manipulated with techniques of search engine optimization (Kritzinger & Weideman, 2017). This study looked at the correlation between user-driven parameters, user ratings, and ranking algorithms to help content providers better understand this relationship.

Problem Statement

The manipulation of website content is a common practice, which allows content providers to trick search engines into thinking their site is more important than others (Luh & Huang, 2016). Content providers manipulate content to get their website higher in a result set, so their site will be more likely to be visited because 62% of users do not go past the first page of results (Giomelakis & Veglis, 2016). The general IT problem is that website content can be manipulated in order to make a page appear more important to ranking algorithms. The specific IT problem is that the content providers lack information on the relationship between user-driven parameters, user ratings, and ranking algorithms.

Purpose Statement

The purpose of this correlational quantitative study was to examine the possible relationship between user-driven parameters, user ratings, and ranking algorithms. This information could potentially help content providers to provide users with content that is better suited to their needs as opposed to content tailored towards search engine optimization. This study utilized a correlational design in which participants from a local university in Huntington, West Virginia ranked a listing of web pages for a specified search term on a 1 to 5 scale as well as provided their internet usage information. The independent variables were the user-driven parameters and the user ratings. The dependent variable was the ranking algorithm. The implications for positive social change include producing more relevant results for users which will help them find more relevant information more efficiently.

Nature of the Study

Quantitative methodology was utilized for this study. Quantitative research relies on statistical information regarding the connections between independent and dependent variables (Maher, Markey, & Ebery-May, 2013). The justification of the quantitative method for this study resulted from the need to determine what the correlation was between variables such as user-driven parameters, user ratings, and ranking algorithms. Qualitative and mixed methods rely, to some extent, on the collection of open-ended information through methods such as interviews or observations (Frels & Onwuegbuzie, 2013). This study focused on the analysis of data collected using different search ranking algorithms and the use of observational data would not have added to the results of this study, therefore qualitative and mixed methods was not appropriate.

In this study, I utilized a correlational design. Correlational design is utilized to find possible relationships between variables (Venkatesh, Brown, & Bala, 2013). In this study, I looked at the relationship between user-driven parameters, user ratings, and ranking algorithms. Utilizing a correlational design, I was able to examine the possible relationship between variables but was not required to manipulate the independent variables in order to determine if any relationship existed. Feedback from participants was from the same question set (see Appendix A) and this feedback was utilized to determine if a relationship existed between user-driven parameters, user ratings, and ranking algorithms.

An experimental design utilizes the manipulation of independent variables in a study to better understand how these affect a dependent variable (Frels & Onwuegbuzie, 2013). While an experimental design could be utilized to look at the effects between independent and dependent variables, the manipulation and randomization of the independent variable required for an experimental design was not needed for this study. The need to control the independent variables in a study made an experimental design inappropriate for this study.

A quasiexperimental design is similar to an experimental design, except it lacks the aspects of randomization utilized by an experimental design (Campbell, Parks, & Wells, 2015). Quasiexperimental designs are often utilized to evaluate the impact of a variable on a process (Campbell et al., 2015). Quasiexperimental design was not appropriate for this study because I was not trying to determine the specific impact variables had on one another, rather I was simply looking for a possible relationship between variables. While quasiexperimental design would have worked for this study at this phase it would have required more in-depth research that may not be necessary if no true relationship existed between variables.

Research Question

What is the relationship between user-driven parameters, user ratings, and ranking algorithms?

Hypotheses

Null Hypothesis (H_0): There is no statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms.

Alternative Hypothesis (H_1): There is a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms.

Theoretical Framework

Kenneth J. Arrow developed the impossibility theorem in 1951 (Arrow K. J., 2012). Arrow's impossibility theorem identifies the following five conditions that a social choice rule should satisfy (Arrow K. J., 2012): First is the complete condition. The complete condition states the social choice rule should provide a complete ranking of all alternatives. Second is the paretian condition. The paretian condition states if every individual prefers A to B, then the social choice rule should rank A above B. Third is the transitive condition. The transitive condition states if the social choice rule ranks A above B and B above C, then A should be ranked higher than C. Fourth is the independence of irrelevant alternatives condition. The independence of irrelevant alternatives condition states the ranking of A compared to B should not depend on preferences for other alternatives. Fifth is the nondictatorial condition. The nondictatorial condition states the social choice rule should not depend on the preferences of only one individual (a dictator). According to Arrow, it is impossible to find a situation that will satisfy all five of these conditions when three or more options exist.

Preference aggregation is an important concept in Arrow's impossibility theorem and social choice theory (Arrow K. J., 2012). The key variables considered in this model are the following: (a) the individual, (b) individual preferences, and (c) the individual rating for each option. Utilizing this information, the results are combined to determine how the group feels about each option presented. This concept aligns with the concepts of this study because I was looking at what the correlation is between user-driven parameters, user ratings, and ranking algorithms, which are similar to the key variables considered by Arrow.

For the purposes of recommendation engines for web searching the following key items need to be considered: (a) the number of users using the system, (b) the importance assigned to pages by the user, (c) the importance of pages assigned by ranking algorithms, and (d) the user's personal preferences. This information allows web pages, as well as users, to be placed into specific groups and can be utilized to help build relationships among both sites and users. As the amount of data available grows the relationships should begin to be more closely knit and the results offered based off of those recommendations should continuously improve.

There are more variables than those presented in the social choice theory, but they do tie together nicely. Sites visited and rated by users all tie to the individual's ratings for a given page and will be used in combination with a ranking algorithm that utilizes backlinks on pages to determine a page's importance to create a new ranking for the page based on user input. In addition, sites visited can help to determine users who have simply rated a site without visiting it first and can be utilized to help weight the importance of the ratings provided by the user. In addition, the ratings provided by the users can be combined with the ratings generated from a ranking algorithm that only uses backlinks to determine a page's importance to potentially create a more reliable listing of search results that would be much more difficult to manipulate utilizing search engine optimization techniques.

Definition of Terms

Collaborative filtering: Describes techniques used to recommend items to users based on similar preferences of other users (da Silva, Camilo-Junior, Pascoal, & Rosa, 2016).

Personalization: Describes the customization of web content to help provide users with a unique experience (Bodoff & Ho, 2016).

Assumptions, Limitations, and Delimitations

Assumptions

When conducting research, the reliability of data being utilized is important, however, it is common to encounter things that the researcher simply assumes to be true, even if there is no scientific evidence to support it, which needs to be identified to the reader (Grant, 2014). The first assumption of this study was that participants honestly rated the sites of the search results presented to them. The second assumption is that participants would not attempt to provide ratings that could sway the results for their own personal benefits.

Limitations

Research studies are not perfect. They have certain aspects that may limit the data that can be collected. The reader needs to be made aware of limitations in the data in order to better understand areas of the study that may limit the results of the current study (Brutus, Aguinis, & Wassmer, 2012). The first limitation of this study was that search engines require a large amount of data in order to allow the user to search the web. Datasets as large as popular search engines were not available to me for this study. The second limitation is that search topics were provided to the participants and the data sets were limited to sites within the category of the selected topic. The third limitation of this study is that the population was limited to students and faculty members in the IT field in the area of Huntington, WV.

Delimitations

Like many large-scale projects, research studies must have a well-defined scope, or delimitations, that the reader needs to be aware of so to better understand the boundaries the researcher has defined (Spitzmuller & Warnke, 2011). The primary delimitation of this study was the ranking algorithm examined. Many different algorithms are utilized by search engines to determine the order of results presented to the user, for this study I was only looking at the relationship between user ratings, user-driven parameters, and ranking algorithms. Another delimitation of this study was the location of the participants. Participants were all either students or faculty members from a university in Huntington, WV.

Significance of the Study

Contribution to Information Technology Practice

The internet is filled with an almost endless supply of information and the amount of information is growing almost every day. As more information gets added searching the web becomes a much more time-consuming task. Users are provided with page after page of results that they must sort through in order to find the information they are looking for. Algorithms try to rank web pages in order to limit the results a user must go through before finding what they are searching for. Collaborative filtering techniques are commonly used to help users find information on popular sites such as Netflix and Amazon but are not as common with larger datasets, such as those found in search engines. This study is significant to IT practice in that it may provide content providers with a better understanding of how users view web content and how those views can potentially affect the results provided by search engines. Improvements to the content providers understanding of this topic can potentially result in higher quality content being provided by search engines and can lead to decreasing the amount of time required to find information on the web.

Implications for Social Change

The implications for positive social change include producing more relevant results for users and help them find more relevant information more efficiently. Utilizing collaborative filtering allows for search results to be grouped and ranked based off what similar users found the most helpful. Users searching for similar information would first be presented with results that others seeking the same information found to be the most successful and can potentially decrease the amount of time spent searching the web. Helping users find higher quality information can help them learn about topics much faster which can help to increase the level of innovation in any number of fields. An increase in innovation may lead to improvements that could potentially improve the quality of life for people around the world.

A Review of the Professional and Academic Literature

The literature review examines topics pertinent to this study, including Arrow's impossibility theorem, search engine ranking algorithms such as Google's PageRank algorithm, collaborative filtering techniques, discounted cumulative gain, and search engine personalization. Conducting searches on these topics yielded 157 articles of which 141(90%) are peer-reviewed and 137(87%) were published within the last five years. Seventy-nine articles reviewed have been utilized within the literature review section 74(93%) are peer-reviewed, and 73(92%) have been published within the last five years. The table below shows a breakdown of all the sources in this study.

Table 1.

Breakdown of Sources

	> 5 years	< 5 years	Peer Reviewed	Not Peer Reviewed
Literature Review	6	73	74	5
Total Sources	20	137	147	10

The articles within the literature review provide an analysis of the relevance and importance of the independent and dependent variables. In addition, the articles utilized outline the key components behind the central idea of the study and how these components can be utilized together to potentially provide users with more relevant search results. The articles will be examined based on themes surrounding the impossibility theorem, search ranking algorithms, and collaborative filtering. These themes will show the individual components of the study. The articles utilized in the literature review are accessible through various databases including Google Scholar, ProQuest, IEEE, and ACM. To find articles for this study I utilized many different keywords, such as collaborative filtering, recommendation engines, social choice theory, impossibility theorem, search algorithms, search engine optimization, ranking algorithms, personalization, and discounted cumulative gain.

The purpose of this correlational quantitative study is to examine the correlation between user-driven parameters, user ratings, and ranking algorithms. This information could potentially help designers to consider new techniques for algorithms that could improve the ranking of pages. This study utilized a correlational design in which participants from a local university in Huntington, West Virginia ranked different websites for different search terms which were utilized to examine the correlation between a ranking algorithm that looks at a combination of backlinks and user provided information and searches conducted by a search tool using a ranking algorithm that only looks at backlinks on websites. The independent variables are the user-driven parameters and user ratings. The dependent variable is the algorithm rankings. The implications for positive social change include producing more relevant results for users which will help them find more relevant information more efficiently. The null hypotheses for this study is, there is no correlation between user-driven parameters, user ratings, and ranking algorithms. The alternative hypotheses for this study is, there is a correlation between user-driven parameters, user ratings, and ranking algorithms.

Theoretical Framework - Arrow's Impossibility Theorem

Kenneth J. Arrow developed the impossibility theorem in 1951(Arrow, 2012). Arrow's impossibility theorem identifies the following five conditions that a social choice rule should satisfy (Maddux, 2014):

1. Complete. The social choice rule should provide a complete ranking of all alternatives.
2. Paretian. If every individual prefers A to B, then the social choice rule should rank A above B.
3. Transitive. If the social choice rule ranks A above B and B above C, then A should be ranked higher than C.

4. Independent of irrelevant alternatives. The ranking of A compared to B should not depend on preferences for other alternatives.

5. Nondictatorial. The social choice rule should not depend on the preferences of only one individual (a dictator).

Arrow states that it is impossible to find a situation that will satisfy all five of these conditions when three or more options exist (Ben-Yashar & Nitzan, 2017). In terms of search engines, this would indicate that is not possible to provide a results list that would be an accurate representation of all users' opinions. Despite this, some researchers believe meeting all five exactly is not necessary (Coban & Sanver, 2014; Gibbard, 2014). Not meeting all five conditions does not mean that any condition is completely ignored, rather the rules for some conditions are a little more relaxed in order to accommodate scenarios where some conditions would be unrealistic. Exactly how Arrow's theorem is applied is also debated. Some groups see certain areas as an aggregation of criteria and not as individual preferences and with this difference Arrow's theorem would no longer apply (McComb, Goucher-Lambert, & Cagan, 2017). While it may not be possible for an algorithm to meet all five conditions when creating a result set for all users, an algorithm may come closer to satisfying all five conditions by creating a result set that is personalized to individual users. In addition to this as you become closer and closer to meeting some criteria, then the end results should also be closer to the group's overall opinion. This closeness to meeting the criteria is the fundamental concept for this study and how improvements could be made to search algorithms. To better understand this, we will look at the five conditions in more depth.

Complete. For this condition to be met every person voting would need to rank every item being voted on. An election with three candidates, republican, democrat, and libertarian can be utilized to better understand the complete condition. In this situation, if voters only choose their top pick there is no way of determining who their second choice would be. Limiting to only the top choice can lead to the overall winner not being an accurate representation of how the entire voting population feels (Kapeller, Schutz, & Steinerberger, 2013; Makovi, 2016). However, if all three candidates are ranked by every voter from top choice to bottom choice these rankings could be combined to arrive at a more accurate representation of the entire group. Consider the following scenario, there are 1000 voters and 450 are voting republican, 350 are voting democrat, and 200 are voting libertarian by simply tallying the top choices the republican is the favorite. When the second and third choices are considered the outcome can be different. Traditionally if the republican candidate drops out of the race, most of those votes would fall to the democrat, if the democrat drops out those votes would go to the republican, and if the libertarian drops out those votes would go to the democrat. Utilizing a borda count, which voters rank their options by preference and assign a value to each preference, a different result occurs (Morreau, 2015). The top choice gets three points, the second choice gets two points and the third choice gets one point. Utilizing the borda count method the tally is as follows:

Table 2.

Borda Count Example

	Republican	Democrat	Libertarian
Republican	450 * 3 = 1350	450 * 2 = 900	450 * 1 = 450
Democrat	350 * 2 = 700	350 * 3 = 1050	350 * 1 = 350
Libertarian	200 * 1 = 200	200 * 2 = 400	200 * 3 = 600
Total	**2250**	**2350**	**1400**

In this situation, by looking at the top choice only the republican is the winner. Utilizing the borda count method, where every candidate is ranked, the group as a whole prefer the democrat over the republican candidate. In terms of search engines, this condition is possibly the most difficult one to meet, in fact, it is arguably impossible. In order to meet this condition, every item in a result set needs to be ranked by every user that searches for a given term. While this condition may not be possible to meet entirely, as users rank more sites within a result set those that the user has ranked can be utilized in combination with other users' rankings to improve their results. The closer the system gets to meeting this condition the more accurate the results should become. This type of ranking is very similar to the concepts of approval voting. In approval voting, voters would be permitted to vote for as many candidates as they want but can only vote one time. Once voting is completed the votes are counted and the candidate with the most votes is the winner (Maniquet & Mongin, 2015). This type of voting would closely relate to how voting could be done with search engines. A user could place a vote on any website that is returned to them, but not have to vote for them all, essentially picking out the sites that are liked within the result set.

Paretian. This condition, by itself, is relatively simple. The paretian condition states that in situations where every voter prefers one option over another the system should as well (Ritesh, 2015). In terms of a search engine, when someone types in the search term “news” he/she is presented with a list of news related sites that users have rated. Among the list presented to the user are Fox News and CNN. In a situation where every user rates Fox News higher than CNN, Fox News should appear above CNN in the list of results. This condition states exactly what you would expect a search engine to do. The most popular item appears before all other items on the list. The difficulty with this condition is that rarely will all users prefer one site over another. User preferences will be a mixture of opinions (Cina & Endriss, 2016). Search engine optimization techniques are currently used by many websites in an effort to have their site appear higher on the results list (Luh & Huang, 2016). Search engine optimization techniques can lead to the possibility of skewed results that may not accurately reflect the opinions of the users of a given search engine. Personalization techniques could be utilized to help meet the paretian condition to some degree. The effects search engine optimization has on-site positions within a result set could also be limited with personalization techniques. By grouping users together based on their interests or past ratings of websites the likelihood of users being more similar in their preference of sites would increase (Orlikowski & Scott, 2015). By increasing the likelihood that users are grouped together based on similarities the chances the paretian condition could be met would increase for the specific groups. Grouping similar users would not guarantee that every user within that group prefers one option over another, but when users are grouped with other users with similar interests the likelihood of users having similar opinions would increase.

Transitive. This condition is related to the transitive property in mathematics. The transitive condition states that in situations where the system shows option one is better than option two and option two is better than option three then option one is better than option three (Jacobs, de Poel, & Osseweijer, 2014). Expanding on the example of a search for "news" from above, add BBC to Fox News and CNN. Above we said that Fox News is preferred to CNN and now we say that CNN is preferred to BBC, with this example we can also say that Fox News is preferred to BBC. The search results can be ordered Fox News, CNN, and then BBC (Cina & Endriss, 2016). This condition orders more than two options based on popularity (Coban & Sanver, 2014). Similar to the paretian condition the transitive condition does exactly what a search engine should do, order a result set based on popularity. Like the paretion condition, the transitive condition would benefit from grouping similar users because this condition would be extremely unlikely to ever be met for all users. Grouping similar users together can help to increase the likelihood of this condition being met within a group.

Independent of irrelevant alternatives. This condition is often seen as one of the more difficult conditions to meet (McComb, Goucher-Lambert, et al., 2017). When looking at two options the preferences for those two options should not change given any changes to any other options within a set (Stegenga, 2015). Using the example of news sites from above, if Fox News is preferred to CNN, then the addition of BBC should not affect that order. No matter what a person's opinion of BBC is that opinion should not cause Fox News to go below CNN. The addition of BBC could cause these options to move down on the list, but this addition should not change the preference of Fox News over CNN. While this condition seems fairly simple to satisfy, it is not easily met by many voting systems. Take the example given above regarding voting for republican, democrat, or libertarian. By doing a straight popular vote the republican candidate would come out the winner. If the libertarian drops out of the race the democrat would take the lead. In order to truly satisfy this condition, the votes for each candidate would have to be independent of one another and the libertarian dropping out should not affect the preference of republican over democrat (Morreau, 2015). This would be very difficult to accomplish in a typical voting system because voters typically know what they are going to do if their candidate drops out, and therefore the alternatives are not independent of one another.

Search ranking algorithms can potentially adhere to this a little more easily. Ranking in terms of a search engine would not be based on one site over another, just the user's overall thoughts of a given site (Shafiq, Alhajj, & Rokne, 2015). These overall thoughts can then be combined with the opinions of others and help to provide a listing to users (Saxena, Agarwal, & Katiyar, 2016). Since the ranking is done independently for each site changes to one ranking should not impact the general preference of other sites on the list (Kritzinger & Weideman, 2017). Looking at the example of news sites above, if users utilize a five-star rating system and give Fox News five stars, CNN three stars, and BBC two stars changes to the star rating of BBC will not affect the general preference of Fox News over CNN, because they will remain at five and three stars. This condition is a perfect fit for search results ranking because sites should be rated based on their own content and not the content of others.

Nondictatorial. This condition states that the results given are not based on the opinions of just one individual, rather they are based on the combination of the opinions of others (Saari, 2016). In terms of ranking algorithms, both collaborative filtering and the Page Rank algorithm satisfy this condition (Abdel-Hafez, Xu, & Jøsang, 2015; Jilani, Fatima, Baig, & Mahmood, 2015). The one exception to this is the beginning phases of collaborative filtering techniques. In the beginning, before any user input is received, collaborative filtering techniques do not have a way to prioritize results and therefore the prioritization is left up to the system. Depending upon how the system handles this prioritization the system could be considered a dictator and the sole means of ranking results (Igersheim, 2017; Wei, He, Chen, Zhou, & Tang, 2017). The Page Rank algorithm provides a ranking that is determined based on the number of links going to a page. The system tallies those links but does not determine how many links exist, which helps to prevent the system from being a dictator (Gleich, 2015). By combining the Page Rank algorithm with collaborative filtering methods a more meaningful result set can be provided in the beginning stages of a collaborative filtering system, without having the system act as a dictator.

Preference aggregation. Arrow's impossibility theorem and social choice theory do closely examine the ideas of preference aggregation. The concepts of preference aggregation are extremely beneficial to search engine ranking algorithms (Gibbard, 2014). Preference aggregation is utilized to take the preferences of several individuals and calculate a single collective set of the alternatives (Fotso & Fono, 2015; McComb, Cagan, & Kotovsky, 2017). The key variables for this study are the items being ranked, the individual preference for each item, collective preferences for each item, and general user information. While preference aggregation may include additional variables, the variables chosen for this study are the fundamental variables required for preference aggregation. A combination of individual preferences for the items being ranked makes up a profile. A profile is simply a set of all user preferences (Stegenga, 2015). At this point, some type of function is applied to these different preferences to come up with one single preference that represents the group's overall opinion, similar to how some collaborative filtering techniques work (Wei et al., 2017; Yu, 2015). According to Arrow's theorem, it is not possible to combine a set of individual preferences into a single group preference that is representative of every individuals' opinion when there are more than two items being ranked (Ninjbat, 2015; Okasha, 2015). While it may not be possible to meet all five of Arrow's criteria with more than two options, many have argued that loosening some of the restrictions can provide a good representation of group opinions (Gibbard, 2014). McComb, Goucher-Lambert, and Cagan defined what they call *conditional arrow fairness* (2017). In order to meet conditional arrow fairness, an aggregation function must meet the paretian, independent and irrelevant alternatives, and non-dictatorial criteria (McComb, Goucher-Lambert, et al., 2017). Conditional arrow fairness leaves two conditions that do not necessarily have to be met, the transitive and complete conditions. The complete condition can be extremely difficult to meet, depending on the population size and number of items being rated because the complete condition requires data regarding every participants' opinions of every possible option (McComb, Cagan, & Kotovsky, 2017). Relaxing this criterion can open the door for more aggregation functions. For this study, it is not possible to collect every single user's opinion about every single website returned within a search result, and therefore the complete condition could not actually ever be met on a large scale. Even though the complete condition cannot be met 100%, how close to 100% complete you are can be monitored. Monitoring how close aggregated results are to individual responses as you approach 100% complete it is possible to determine how the complete criteria affect the aggregated results. The concepts of collaborative filtering provide a good example of this concept. Collaborative filtering techniques cannot provide helpful recommendations to users without any type of feedback from the user (Wei et al., 2017). As the user begins supplying feedback, the recommendations that are given to the user become more useful (Moradi & Ahmadian, 2015). Netflix's recommendations work on this principle (Chen, Chen, & Wang, 2015). As users watch and rate programs within a system the recommendations provided by the system become more accurate. Each rating provided by users within their system brings them closer to meeting the complete criteria of Arrow's framework. While they will most likely never completely meet this criterion, each step closer provides an improvement to their recommendations.

The general idea behind Arrow's theorem is that in order to create a perfect single group preference set the aggregation function must meet all five criteria. One must keep in mind that we do not live in a perfect world and very few if any, scenarios will completely meet all five criteria. Even though the criteria may not all be met in their entirety, they are important to consider and to note situations that may arise that will prevent them from being met. The importance of each of the five criteria is not universal and depends upon what you are ranking (Gibbard, 2014; McComb, Goucher-Lambert, et al., 2017). Noting which criteria are met, and how close you are to meeting others can provide insight into the accuracy of the results set.

Variables

User-driven parameters. User-driven parameters are one of two independent variables that were discussed in this study. User-driven parameters are key elements that relate specifically to the user, such as user preferences, user interests, and profile information (Bostan & Ghasemzadeh, 2014; Chen et al., 2015; Yeung, 2016). Many popular websites, such as Netflix and Amazon, rely very heavily on this type of information in order to help provide their users with recommendations of things they may be interested in (Chen et al., 2015; Moradi & Ahmadian, 2015). The utilization of user-driven parameters has truly changed the internet. User-driven parameters can be used to help provide every user with a unique experience on any number of sites. Currently, user-driven parameters are primarily utilized by specialized sites where metadata being utilized for items can be controlled to some extent (Ramesh & Andrews, 2015). Being able to expand this methodology to more extensive data sets, like those utilized by search engines, could potentially lead to improvements in results provided to users when searching the internet. Content providers gaining a better understanding of how user-driven parameters effect result sets can lead to better content targeted to the users' needs rather than to how the search engines look at it. Following are examples of types of user-driven parameters that are commonly utilized within collaborative filtering algorithms and how these parameters can be used to help users find information they may be interested in.

User history. In some situations, content providers with websites that provide recommendations do not want to have to ask their users to provide them with information about their interests, but still want to be able to provide some kind of feedback regarding other things the user might be interested in (Guo & Chen, 2016). When this is the case content providers can utilize implicit information based off of the user's activity on the site (Jerath, Ma, & Park, 2014). For example, content providers can track the links a user clicks as well as the amount of time a user spends looking at specific pages. Using this information, algorithms can be implemented that can calculate a user's interest in specific items. These algorithms can categorize those interests to help group users and provide them with recommendations of other items they may be interested in. Utilizing history alone can have issues. Algorithms that only utilize a user's history must make a lot of assumptions when determining a user's interest in a given item (Singh & Sethi, 2016). Just because a user clicks on an item and spends some time looking over the page does not mean the user is actually interested in that item. As a result, the algorithm may place a higher interest value on that item then is actually necessary.

User history was not utilized within this study. User history data could have been utilized with the user-driven parameters variable. While user history data can provide valuable data, the ability to collect user history data was outside the scope of this study.

User ratings. A user's opinion is an extremely powerful resource that can be utilized in ways to benefit all users of a system. In collaborative filtering, items get assigned ratings by users. User ratings can be utilized to estimate the user's interest in the given item (Guo, Zhang, & Yorke-Smith, 2015). With a single user, this information is not extremely useful, however with multiple users, this information can be combined to help group users together who share similar interests. Once users are grouped together collaborative filtering algorithms can be utilized to find items that similar users were highly interested in. Items that users within the group showed interest in that others within that group have not given an opinion on can be suggested to those users (Yeung, 2016). As users continue to use these systems and rate more items the accuracy of the recommendations that are provided to the user should increase (Zhang & Min, 2016). Zhang & Min showed that user ratings combined with user data can help to provide users with more personalized data that can be utilized to improve the overall user experience. The true benefit to this is that the user responses help to build a large amount of data that has the potential to help many people, not just themselves. While using user ratings with collaborative filtering can provide benefits to users, it also has some issues that can make it a difficult solution to implement.

User ratings directly relate to Arrow's impossibility theorem. Arrow looked at traditional voting systems which is essentially the concept behind ranking algorithms, to determine what items a user prefers over other items. The ratings provided by the user can simply be looked at as a vote for that item. User ratings are an important part of evaluating ranking algorithms because how the user feels about sites within a result set is the best way to determine if the ranking algorithm performs to the user's satisfaction. User ratings were utilized within this study to determine the discounted cumulative gain for search terms used in this study.

User interests and demographics. Utilizing collaborative filtering requires some kind of data from the users and the items in a system. The need for this data presents a problem with new systems. One of the biggest issues with collaborative filtering is a lack of data regarding new users and items (Wei et al., 2017). In order for these systems to provide recommendations to users the system contain information about the user as well as information regarding the items that will be recommended to the user. When a user first joins his/her interests are not known to the system and therefore the users can't be grouped with similar users (Sehgal, Chaudry, Biswas, & Jain, 2016). Some sites can get around this by adding questions to the registration process about the users' specific interests (Singh & Sethi, 2016). The responses to questions about the user's specific interests can be utilized to group the user with those who have similar interests without the need to look at other information such as user history or user ratings. Being able to group users together without the need for user history or ratings can help to provide recommendations to the user when he/she first comes to a webpage. A user's interests can be utilized in place of detailed search history to group users into groups of similar users (Chen, Ji, & Wang, 2017). The first use of a collaborative tool can be limited in terms of recommendations because accurate recommendations are dependent upon the data provided by the user. The use of user interests and demographic questions can be utilized to alleviate this issue (Yeung, 2016). Questions pertaining to user interests and demographics can later be combined with additional information to help improve recommendations. Zhang and Min utilize a combination of information pertaining to the user, items, and item ratings in order to determine whether or not to recommend items (2016). Initially user interest and demographic questions can be helpful in grouping the user with similar users. As the user provides more information user interest and demographic data can be combined with additional information such as item ratings. Combining user provided interest and demographic questions with additional information such as the how the user rates items within the system can help to ensure that the users recommendations are being based on the appropriate group of users.

User interests and demographics can mean a variety of things. When discussed in terms of collaborative filtering the primary purpose for collecting user interests and demographic data is to help group similar users together. Al-Shamri utilized users' age, gender, and occupation in different combinations to explore different approaches to utilizing demographic data (2016). Son utilized age, education, number of children, and living standards to help group similar users together (2014). Using demographic based data can be helpful in developing an initial user profile. User demographics alone do not provide the best possible solution for grouping users, but grouping based on demographics is better than randomly assigning a user to a group (Pazzani, 1999). The initial grouping of users can be modified later on using other information provided by the user such as item history, ratings, or user interests. As users interact with a system more data gets collected. As the amount of data about the user increases the accuracy of recommendations should increase as well.

User interests and demographics were utilized in this study for the user-driven parameters variable. While this does not directly relate to Arrow's impossibility theorem it can be utilized to help address the viability of the concept of conditional Arrow's fairness. The idea behind conditional Arrow's fairness is that the complete rule is to strict but can be looked at as how close to complete a voting system is (McComb, Goucher-Lambert, et al., 2017). By breaking down ratings into groups you can compare smaller sets of meaningful data to the overall data set to determine the viability of conditional Arrow's fairness.

This study utilized user interests to help group participants together. Groups were based on how the participants primarily utilize the internet. Feng, Wong, Wong, and Hossain utilized the groups of Facebook usage, study usage, and entertainment usage (2019). For this study I utilized similar groupings, social media, news/research, and games/media. This type of information is easily collected and does not require the collection of large amounts of participant usage data. The additional benefit to using user interest questions is the ability to test how the user groups data can help with the cold start problem with collaborative filtering.

Item information. Whether you are looking at user ratings, user history, or some combination of the two, information about the items being looked at is extremely important. In order to provide a user with valid recommendations, you must have some kind of data that allows you to group these items into groups (Abdel-Hafez et al., 2015). This study looked at how these tools can potentially be utilized to recommend websites to users, using this as an example, keywords on a web page can be utilized to help group sites together and help to narrow down the pool of sites that are recommended to a user based on his/her own interests (Singh & Sethi, 2016). The links that are presented on a web page can also be utilized to help group pages together. In many cases pages are only going to link to other pages they feel their users might be interested in and therefore these pages could be considered similar (Brin & Page, 2012). Using this information, keywords from these pages can be utilized as well to help expand the potential pool of sites that may be of interest to a given user based off of his/her own interests. These techniques are commonly utilized by search engines when ordering links in result pages.

Ranking algorithms. Ranking algorithms are utilized by search engines to help determine the importance of a website given a specific search term and is one of the independent variables that was used in this study. Many different ranking algorithms are utilized by different search engines, but they all have the same end goal in mind, which is to provide users with the most relevant websites first in order to help shorten the amount of time a user spends going through the list of sites provided to them (Barboucha & Nasri, 2015). This study focused primarily on an algorithm designed by the founders of Google called the Page Rank algorithm (Brin & Page, 2012). While improvements have been made to the Page Rank algorithm since its creation, the original algorithm is still capable of producing quality results for the user when no feedback has been provided (Yan et al., 2014). In this study, I focused on the collaborative ranking of search results. With collaborative ranking, an initial ranking still needs to be present to be able to provide results to users that have not provided appropriate feedback (Wei et al., 2017).

By implementing collaborative filtering techniques around existing ranking algorithms these types of issues can be limited.

Page Rank algorithm. Search engines utilize different methods to prioritize the results given to the users. Page Rank is one algorithm utilized by Google to help determine the importance of a website (Jilani et al., 2015). In the Page Rank algorithm, all web pages in the database are assigned a value which is determined by the number of other pages that link to them. The more websites that link to a site the higher the importance of that site and therefore the higher the site is on the list of results provided to the user (Kumar & Prakash, 2015). The idea is relatively simple, the more sites that link to a page the more likely that page has useful content and is most likely a quality site. The algorithm gets a little bit more complicated when determining the weight of the importance of each of these backlinks (Yan et al., 2014). A website that has a higher importance value, or page rank, carries more weight than a site with a lower value. If this site links to a page this would increase that page's importance by more than a page with a lower value linking to it (Pirouz & Zhan, 2016).

In order to better understand the concept behind the Page Rank algorithm, we can look at it in respect to a random web surfer (Gleich, 2015). In this concept a user is looking at a web page, the user clicks a button that takes them to another page, the new page will either be a page that was linked from the page the user is currently looking at or a completely random page. The likelihood of the random page being useful to the user is relatively low, but the page linked off of the site they are looking at is much more likely to be useful to the user, and therefore is more important. This idea is fundamental to the

Page Rank algorithm. When a web page is created, there is typically a theme behind it and therefore the links present on the site are highly likely to follow that theme. By following the link structure of the web page a general model can then be derived counting the number of links going to a page. Links going to a page are looked at as a vote for that page, the more votes for the page the more important the page is considered.

One issue with the Page Rank algorithm is that a page's importance is determined by a combination of backlinks and keywords, and can, to some degree, be manipulated by the creator of a web page by setting up relationships with other websites to link to their site as well as wording their website in such a way to emphasize the importance of certain words to help bring the importance of the page up for specific keywords (Mavridis & Symeonidis, 2015).

There are two basic principles regarding how to conduct searches, keyword and semantic-based searches (Singh & Sethi, 2016). While this study is primarily focused on the ranking algorithms utilized by search engines, understanding how searches are conducted is also important because the two go hand in hand. Searching provides a listing of sites that can then be passed to the ranking algorithm to determine the importance of the site (Fang, 2016; Saxena et al., 2016). Understanding the two basic methods for search will help to provide some insight into how the searching and ranking algorithms work together.

Keyword-based searches look almost entirely at the words on the page to determine how relevant a page is to given search term. These search methods do not only look at the words that are visible on the page, but they also look at words utilized in tags, such as the <meta> tag within the HTML code (Mahendru, Singh, & Sharma, 2014). This method of searching, by itself, is highly susceptible to manipulation, because the more prevalent a term is on the page, the higher it will be ranked within the search results. Keyword-based searches can produce results relatively quickly. Keyword searches run a simple query for a specific set of words (Singh & Sethi, 2016). The downside to this method is the creators of websites can easily manipulate results by targeting specific keywords in their text. Targeting specific keywords increases the keyword count and causes these algorithms to place a site higher on the list of results (Singh & Varshney, 2013). Semantic-based searches have methods in place to try to prevent this from occurring.

Semantic-based searches are a little more complex. Semantic searches utilize a knowledgebase of search terms and refine and extend a search term to allow for more keywords to be included within a search (Dascalu, Trausan-Matu, McNamara, & Dessus, 2015). The concept behind semantic search is fundamentally similar to keyword-based searching. The user enters a search term and pages that contain that search term are returned. The difference is before the pages are searched the search term entered is compared to terms in the knowledge base to find other search terms that are connected to it (Singh, Sharma, & Dey, 2015). Search results are presented based on not only the original search term, but the connected terms as well. By expanding these terms new potentially beneficial sites will be presented to the users. Semantic methods are often able to present the user with higher quality results, because it does not look for a specific keyword set, rather a set of related keywords. Utilizing a keyword set can help to weed out the less desirable sites that are trying to manipulate the system (Meusel, Ritze, & Paulheim, 2016). The downside to semantic-based searching is that it is not as fast as keyword-based searching (Dong et al., 2017).

The types of search methods available are important to consider along with the ranking algorithm. The Page Rank algorithm is separate from the search method being utilized. The ranking of pages takes place after the search has been conducted and determines how those results will get ordered, but the 2 processes are closely connected. In situations where pages have the same Page Rank the prevalence of keywords among the pages could help to decide which site the user would rather see over the other. Utilizing either keyword or semantic searching methods you can track the count of keywords, or similar keywords, on a webpage to determine how closely related a site is to the original search term and then utilize this information in combination with a Page Rank value as a tiebreaker when necessary.

Combining Page Rank with User Opinions

The concepts outlined above played a key role in this study. The Page Rank algorithm was the first algorithm utilized by Google and helped make them a household name synonymous with search (Chen et al., 2017). Google has made many improvements to their search algorithms since its release, but the simplicity of this algorithm provides a perfect base for this study. The Page Rank algorithm takes into account the number of sites that link to an existing page in order to determine a site's importance (Jilani et al., 2015). User opinions are not taken into account in the original Page Rank algorithm (Brin & Page, 2012). Taking a calculated Page Rank and combining it with the opinions of users through a simple star-based rating system of web pages, can potentially provide a more accurate representation of how users feel about certain websites and their content (Mala & Hema, 2014). This information can be taken one step further because the information provided is enough to actually start personalizing the results provided to a specific user (Gupta & Chavhan, 2014; Zhang & Min, 2016). With enough user feedback, given a search term, we can utilize the concepts of collaborative filtering to take the results of the search and find the pages that similar users were more interested in and prioritize the list of results provided.

Correlation between user provided rankings and algorithm rankings. Balakrishnan, Ahmadi, and Ravana (2015) examined the differences between five different feedback models for ranking search results. The five models were the following:

1. Baseline. A model with no user feedback.
2. Rating. A model which included users' ratings.
3. Comment. A model which included users' comments.
4. Referral. A model which included users' referrals to websites.
5. CoRRe. An integrated model which included facets of all models.

Balakrishnan et al. (2015) utilized discounted cumulative gain to evaluate the precision of the result sets for each model using the top-five, top-ten, and top-15 results. The findings showed that the CoRRe and Rating models outperformed all of the other models. Balakrishnan et al. conducted an independent sample *t*-test and found the differences between the baseline and CoRRe model to be significant at all 3 document levels: top-five ($t_{(18)} = -3.049$, $p = 0.016$), top-ten ($t_{(18)} = -5.158$, $p = 0.001$), and top-15 ($t_{(18)} = -$

5.344, $p = 0.01$). Zhitomirsky-Geffet, Bar-Ilan, and Levene (2016) conducted a similar study looking at individual user ratings versus search engine rankings. Zhitomirsky-Geffet et al. (2016) found that aggregated rankings were much more reliable than individual user ratings. When looking at the difference between aggregated users' rankings and search engines' rankings Zhitomirsky-Geffet et al. found a difference of 20-60% for the top-ten ranked sites. The findings of Balakrishnan et al. and Zhitomirsky-Geffet et al. show that the addition of user ratings can provide improved results to search engine algorithm rankings.

Suruliandi, Rajkumar, and Selvaperumal (2015) conducted a study looking at the performance of personalization techniques in search engines by comparing web results without personalization, web results with personalization using content analysis, and web results with personalization using user groups. Suruliandi et al. (2015) collected precision, recall, and f-measure data for each method, see table 3.

Table 3

Comparison of Precision, Recall, and f-Measure data.

Method	Precision	Recall	f-measure
Web results without personalization	0.715	.8231	.7652
Web results with personalization using content analysis	0.7206	0.7513	0.7356
Web results with personalization using user groups	0.8437	0.8658	0.8546

Note. Adapted from "Validating the performance of personalization techniques in search engine," by Suruliandi et al., 2015, *ICTACT Journal On Soft Computing Note*, *5*(3), 965-970.

The findings of Suruliandi et al. (2015) show the addition of personalization using user groups provides higher quality results then utilizing the standard search engine algorithm and personalization using content analysis. The data collection within this study is based on the study conducted by Suruliandi et al. (2015). Sureuliandi et al. collected data by having participants rate sites on a search results page, for my study participants were asked to rate websites that were presented by a search engine for a given search term. The primary difference between the data collection methods is how the sites were presented, rather than presenting a results page in my study participants were presented with questions and asked to rate those pages in question format.

This study looked at the correlation between user-driven parameters, user rankings of websites, and algorithm rankings. Specifically, I looked at the difference between algorithm ranking and simple combined user rankings. The idea was that while the algorithm ranking and combined user rankings both can provide good results to a user, there may be a correlation between these methods. The understanding that a correlation may exist can help content providers realize that by focusing efforts solely on achieving a higher ranking on a search engine is not enough. By understanding this, content providers may focus more effort on the users and can provide higher quality content.

Collaborative filtering techniques utilize the opinions of similar users to help recommend options a user might like (Gautam & Bedi, 2017). In terms of a search engine, when a user provides a search term the system can order results based on the ratings from the user and similar users. This concept is used heavily to help recommend items to users in many popular websites today, like Netflix (Li, Zhang, Wang, Chen, & Pan, 2017). As more and more users utilize the system and provide more ratings for web pages the accuracy of the results provided should begin to increase. In addition to increased accuracy of search results, the system could eventually begin to recommend new topics of interest to specific users based on this information (Abdel-Hafez et al., 2015). While collaborative filtering is capable of providing good recommendations to the users, it can only do so if the user provides with ratings for items within the system (Najafabadi, Mahrin, Chuprat, & Sarkan, 2017). Other methods of collaborative filtering have been devised in order to get around this limitation.

Memory-based recommendation is one method commonly utilized to get around the issue with sparse user ratings (McKechnie & Prithwiraj, 2016; Wei et al., 2017). Memory-based systems monitor user activity to determine areas of interest and can suggest other areas of possible interest based off of the activity of other users (Abdel-Hafez et al., 2015; Bodoff & Ho, 2016; Yang, Guo, Liu, & Steck, 2014). This type of collaborative filtering is most commonly found in commercial systems, such as Amazon, to help users find other products they may be interested in (Abdel-Hafez et al., 2015; Bodoff & Ho, 2016). This method of collaborative filtering has its issues, in order to make suggestions the system must be able to determine the similarity between items, but if this data is sparse the ability to find similar items becomes difficult (Moradi & Ahmadian, 2015). One solution to this issue is to use a model-based approach which simply combines a rating based system with a memory based system (da Silva et al., 2016). With the model-based system ratings are not simply used to find similar users, rather the ratings are utilized in combination with user activity to help develop a model of similarities between items. These similarities can then be utilized to make connections between items within the system rather than connections between users.

Regardless of the collaborative filtering method utilized a significant amount of data is required. New users will not have provided enough feedback to truly benefit from collaborative filtering methods (Chen et al., 2015). Combining collaborative filtering techniques with an existing ranking algorithm could help to alleviate this issue. The Page Rank algorithm does not take into account any user feedback and only produces results based off of links going to a specific site (Brin & Page, 2012). The general idea is that the more links going to the site, the more popular the site must be. This concept is extremely powerful and can produce fairly good results. The information provided by this algorithm can be utilized as a baseline for this study. One of the biggest problems with personalization techniques is acquiring the data required in order to successfully implement the solution (Orso, Ruotsalo, Leino, Gamberini, & Jacucci, 2017). Personalization techniques require a large amount of data that must be provided by the user in order to work, before this data is acquired there is no way to effectively rank items within the system (Bostan & Ghasemzadeh, 2014; Ghose, Ipeirotis, & Li, 2014). The Page Rank algorithm can potentially provide a mechanism to rank items and present them to the user in a somewhat meaningful fashion. Utilizing Page Rank as a base algorithm for collaborative filtering could potentially provide new users with quality results that can improve over time as they provide feedback to the system.

Collaborative filtering and algorithm rankings. This study was interested in the possible relationship between the techniques of collaborative filtering and algorithm rankings. Both concepts are extremely powerful on their own, but they also have limitations. Collaborative filtering requires huge amounts of data that must be provided by the user (Aghaiipour-Chafuchahi & Ahmadi-Abkenari, 2015). Algorithm ranking does not require any user data to be provided to rank a web page. This lack of user input results in generic result lists that may or may not be what the user is actually interested in

(Gleich, 2015). The combination of the two techniques can potentially alleviate the weaknesses of one another. The algorithm ranking can provide a good starting point for new users who utilize the system. As these users provide ratings for sites this information can be utilized to combine collaborative filtering values to the ranking provided by the algorithm to present the user with a personalized listing of results.

Algorithm ranking utilizes the links going to a page as a type of vote for the site. Using this concept, the votes can be combined with user votes to establish a new importance value for a web page (Liu, Chen, et al., 2016). This combination of values can help to produce a ranking value for search results that can be driven more by the users rather than just the structure of the web. Early implementations of such a ranking system would rely heavily on the values produced by the ranking algorithm because user ratings would be very limited. As more users utilize the system more user ratings will be put into the system, as this occurs the rankings of the web pages will become more based on the users as opposed to the algorithm ranking value alone.

Differences among Algorithms

Each of the algorithms used in this study can provide an ordered listing of websites given a search term. These listings should all be ordered in a meaningful way, but these orderings can potentially be very different and may not be representative of how the user may order these sites. In this study, I looked at the correlation between different types of personalization methods and algorithm rankings. By looking at each algorithm with an identical set of search results we should be able to determine whether or not the addition of user opinions to existing algorithms can help to provide results that would be an improvement to the group as a whole or if the existing algorithms more closely represent the group's views. In order to compare the different algorithms, a measurement tool is needed.

Search engines utilize different methods to rank the websites in a set of search results when making modifications to algorithms search engines need some way to determine if changes are improvements or not. One of the most common methods of doing this is calculating a discounted cumulative gain value (Jayashree & Christy, 2015). Discounted cumulative gain takes user opinions into consideration to determine an ideal ranking of web pages and then compares values from algorithms to determine which method more closely resembles the opinions of the users (Almulla, Yahyaoui, & Al-Matori, 2015; Benabbou, Perny, & Viappiani, 2017). The calculations for the discounted cumulative gain are fairly simple but can produce very good results for determining the quality of a ranking algorithm.

The discounted cumulative gain not only looks at the quality of sites within a result set but also takes the order of those sites into account. In order to calculate the discounted cumulative gain, you must have a rating for the sites in the list as well as the position of those sites in the results set (Demeester, Aly, Hiemstra, Nguyen, & Develder, 2016; Liu, Song, et al., 2016). The calculation simply takes the log of the position plus one and divides it by the ranking of that site, this is done for all pages in a result set and then the summation of those values is calculated. This calculation must be done multiple times, first, the ideal value is calculated, this is the value when the result set is ordered exactly the way the user would order the sites. Then the calculation is done again for how the ranking algorithm orders the sites, the algorithm which has a discounted cumulative gain closer to the ideal discounted cumulative gain would be the algorithm that more closely represents the ideal ranking set (Ignatov, Nikolenko, Abaev, & Poelmans, 2016; Palotti, Hanbury, Muller, & Kahn, 2016; Sugumaran, Ravi, & Shanmugam, 2017). Rather than comparing entire lists of results discounted cumulative gain provides a single value for each result set that can easily be compared to determine how close an algorithm's results reflect the opinions of the user. This study was interested in the differences between search algorithms and the discounted cumulative gain provided a perfect method for comparing the different algorithms in the study.

Transition and Summary

In Section 1 the objective of this study and a review of the literature about the topic was discussed. The objective of this study is to determine if a significant correlation exists between user-driven parameters, user rankings, and rankings derived from other algorithms. To help determine this, a correlational quantitative study was conducted to examine the correlation between user-driven ranking methods and traditional ranking algorithms and whether or not the combination of the two can help create a more personalized search experience for the users. The literature review examined topics pertinent to this study, including Arrow's impossibility theorem, search engine ranking algorithms such as Google's Page Rank algorithm, collaborative filtering techniques, discounted cumulative gain, and search engine personalization. The articles utilized helped to provide an analysis of the relevance and importance of the dependent and independent variables in the study and how these components can potentially be utilized together to provide users with more relevant search results.

In Section 2, I provide a description of the role of the researcher, the research method and design, population and sampling, data analysis, ethical considerations of the participants, and issues concerning validity and instrumentation. The role of the researcher section will detail how I was involved with the study and its participants. Key ethical points from the Belmont Report will be looked at and how this study was conducted to ensure fair treatment of all participants. The research method and design section will provide a more detailed look at quantitative methodologies and experimental designs and why they were chosen for this study. Population and sampling will look at the specific populations being looked at for this study and the samplings methods that were applied. The data analysis section will look at how data collected will be looked at and the tools that were used for these comparisons. Ethical considerations will look at how the participants were protected throughout this study, specifically it will show how privacy was maintained and how participants could opt out of the study. The validity section will examine issues that can occur with study validity and how I worked to alleviate these concerns as needed.

Section 3 will provide detailed information regarding the findings of this research study. This will include a presentation of findings, possible applications to professional practice, implications for social change, recommendations for action, and reflections of my experience while working on the study.

Section 2: The Project

This section will outline more detailed information regarding how this research study will be conducted. More detailed information regarding plans for this study will be provided to give the reader a better understanding of how I moved forward with this study. Specifically, it will discuss the purpose of the study, the role of the researcher, the specifics of participants, the research method and design utilized, ethical considerations, the data instruments, data collection, data analysis techniques, and study validity concerns.

Purpose Statement

The purpose of this correlational quantitative study was to examine the possible relationship between user-driven parameters, user ratings, and ranking algorithms. This information could potentially help content providers to provide users with content that is better suited to their needs as opposed to content tailored towards search engine optimization. This study utilized a correlational design in which participants from a local university in Huntington, West Virginia ranked a listing of web pages for a specified search term on a one to five scale as well as provided their internet usage information. The independent variables were the user-driven parameters and user ratings. The dependent variable was the ranking algorithm. The implications for positive social change include producing more relevant results for users which will help them find more relevant information more efficiently.

Role of the Researcher

The role of the researcher is to develop the research for the problem he/she has identified. This role includes formulating a hypothesis, collecting and analyzing data, and formulating conclusions from the data analysis (Kyvik, 2013). In the data collection and analysis process, I was able to separate my own opinions on the topic from the data in order to provide accurate information without any personal bias (Thamhain, 2014). In this correlational study, I ignored my own personal experiences and thoughts regarding search engine ranking algorithms. The interpretation of data can be prone to personal bias and it was imperative that I did not allow my own opinions to influence the interpretation of data (Sohn, Thomas, Greenberg, & Pollio, 2017). All the data collected from the participant responses were looked at in an unbiased fashion and the interpretation of the results was not influenced by my own personal experiences and thought processes.

Participants for this study were selected from higher education institutions in the area of Huntington West Virginia. I have worked for these institutions and have a continuing professional relationship with many employees at these institutions. As the researcher, it was my responsibility to ensure that these relationships did not affect the results of this study. The existing relationships I have with these institutions were utilized to help find participants who were willing to be included in the study, but it was my responsibility to ensure those participants selected had no previous relationship with myself in order to prevent any potential bias in their responses.

This study was conducted to determine the possible relationship between user ratings, user-driven parameters, and ranking algorithms. With a better understanding of this relationship, content providers may potentially see that tailoring content to the user instead of to search engines can be an effective technique that could lead to higher quality content. My interest in this topic has grown from an interest in better understanding search engine technologies and a passion for recommendation engines. Over many years I have formed my own thoughts on how these technologies could be related, but have never tested these thoughts. This study was my opportunity to test these thoughts on a small scale to determine whether or not further research on the topic would be viable.

Working with human participants presents challenges. It is important to ensure that participants are treated fairly. As the researcher, I must ensure that I act ethically when working with the participants and their data. The Belmont Report outlines three basic ethical principles that should be considered when conducting research with human participants respect for persons, beneficence, and justice (U.S. Department of Health and Human Services, 1979).

Respect for persons is to ensure the autonomy of all participants and if autonomy is diminished that those participants are protected (Faden et al., 2013). In addition to autonomy, it is important to ensure that all participants are aware of what the research entails and can make an informed decision about whether or not to participate (Judkins-Cohn & Kielwasser-Withrow, 2014). As the researcher, it was my responsibility to ensure that all participants were provided with sufficient information in order to make an informed decision regarding whether or not they wanted to be included in the study. This information included how the information they provided was utilized, the potential benefits of the study, and how the information they included was protected. The anonymity of the participants is extremely important, and it was my responsibility to ensure that participants names were in no way associated with any responses that were provided.

Beneficence is to ensure that participants are treated ethically, no harm comes to them, and possible benefits of the study (Laage et al., 2017). Experiments in this study was based on data collected via survey monkey, so participants were not in danger of any physical harm. The identity of participants was completely confidential, and no names were associated with any responses. The results of this study can potentially help to improve the quality of search result ranking algorithms, which can potentially benefit all participants in the study.

Justice looks at the fair distribution of any potential benefits from the study and is to ensure that all participants have an equal opportunity to share in the potential benefits (Lantos & Spertus, 2014). In this study all participants were given the same set of questions to answer, so all participants had an equal say in all data that was utilized when formulating the results. The weight of responses was equal for all participants regardless of the person's background or responses.

Participants

I select participants from a local community college and university in Huntington West Virginia. Participants included both faculty members and students from programs within the IT field who are 18 years of age or older. Faculty and students from higher education institutions are among the group whose use of the internet is greater than other groups (Salam et al., 2015; Verma, Jha, & Mitra, 2016). Limiting to a group whose use of the internet is higher than others helped to ensure that participants were familiar with browsing the web. By narrowing it down to a specific field, the search terms utilized in the study could be controlled to areas that all participants were at least somewhat familiar with. The research question was looking for any correlation that may exist between user-driven parameters, user ratings, and ranking algorithms. Anyone who utilizes search engines would truly align with this research question, I chose to limit the participants to a specific field because user rankings of web pages could then be narrowed down to only sites within that field and all participants should have knowledge of the subject matter.

In order to gain access to participants within these organizations, I contacted the deans of the appropriate programs with an invitation and consent form via e-mail to forward on to current students and faculty. Utilizing a sponsor that the participants are familiar with can help to improve the number of potential participants who take part in the study (Rindfuss, Choe, Tsuya, Bumpass, & Tamaki, 2015). Participants must be made aware of what they are participating in and their anonymity must be protected (Faden et al., 2013). The Belmont report outlines the need to protect all participants' privacy (Judkins-Cohn & Kielwasser-Withrow, 2014). Privacy of participants can be protected by providing informed consent and ensuring the participant's anonymity (Roberts & Allen, 2015). The usage of a sponsor can potentially lead to issues where potential participants feel obligated to participate in the study. Ensuring the participants of their anonymity will help to alleviate this concern (Lantos & Spertus, 2014). Emails sent to potential participants contained details about the study as well explained the potential benefits participating in the study could bring to the field. In addition to this, the e-mail included details of how participant’s information was secured to ensure they understand their information would remain confidential if they chose to participate.

In order to establish a working relationship with the participants, I began with an introduction email where I explained my background and my intentions for the study. This email can be found in Appendix B. By doing this the student participants should have felt a connection to me because they were able to see that I was once in their position. Similarly, the faculty members were able to relate because of their past experiences conducting research projects. Establishing common ground between the researcher and the participants can help to build rapport between the two (Bowden & Galindo-Gonzalez, 2015). Limiting participants to the IT field helped to establish this rapport with the participants because I am a member of the IT community, this provided some common ground between myself and the participants. Allowing participants the opportunity to receive personal attention from me can help to establish rapport as well (Goode, Lin, Tsai, & Jiang, 2015). In order to help give the participants the opportunity to receive this personal attention, I provided personal contact information for myself to allow the participants to easily follow up with me with any questions or concerns they may have had.

Research Method and Design

Method

Quantitative methodology was utilized for this study. Quantitative research relies on statistical information regarding the connections between independent and dependent variables (Thamhain, 2014). Data collection within quantitative studies is done by different methods including questionnaires, correlational analysis, and systematic observations (Green et al., 2015). In quantitative research, a hypothesis is formulated and then tested using the data that the is collected during the research (Bryman, 2001; Westerman, 2014). In this this study I needed to utilize statistical analysis to determine if a correlation existed between user-driven parameters, user ratings, and ranking algorithms. In this study, I collected data from participants regarding their opinions on various websites given within a set of search results and their internet usage. This data was then compared to results of the ranking algorithms to determine if there was a correlation between the user ratings, user-driven information, and the ranking algorithm.

Qualitative methodology would not be appropriate for this study because qualitative methodologies utilize open-ended information regarding a phenomenon in order to examine the relationship between variables (McCusker & Gunaydin, 2015; Yin, 2014;). Typically, qualitative study results in a more open-ended analysis of participants' thoughts on the subject (Green et al., 2015). The results of a qualitative study are broken down based on themes that emerge during the study to form a narrative around the phenomenon being studied (Bryman, 2001; Yin, 2014). Qualitative methodology could potentially have been utilized for this study to determine the individual participants own feelings about the differences between the ranking algorithms, but this would have given the participants opinions and not a true statistical difference. The results of this study needed to be more statistical than open-ended and therefore a quantitative method was a better choice than a qualitative method.

Mixed methods utilize a combination of both quantitative and qualitative methodologies (Bryman, 2001; Green et al., 2015; McCusker & Gunaydin, 2015). The combination of both quantitative and qualitative methods can help to alleviate certain restrictions of each individual method (Bryman, 2001). Mixed methods are often done by utilizing qualitative research to help formulate a research question and hypothesis that can then be tested utilizing quantitative methods (McCusker & Gunaydin, 2015). This study focused on the statistical analysis of data collected utilizing different search ranking algorithms. While open-ended information could potentially have been collected to get individual participants overall views, the addition of this data collection would not have added to the results of this study and therefore a mixed method study was not appropriate.

Research Design

In this study, I utilized a correlational design. This design was chosen because it is utilized to find possible relationships between variables (Venkatesh et al., 2013). A correlational design tries to determine how variables in a study are related and how strong this relationship is (Simon, 2013). Data in a correlational design is looked at without any control or manipulation (Becker et al., 2016). This study explored the possible relationship between user-driven parameters, user ratings, and ranking algorithms. A correlational design allowed me to collect data from participants without a need to control or randomize data collected. Specifically, this study presented the user with multiple lists of web pages and asked them to provide a ranking for each site on the list. The rankings the user provided were utilized with their internet usage data to set a baseline score for each page which was then utilized to see if a correlation existed between the user ratings, user-driven parameters, and the ranking algorithms.

Experimental design allows the researcher to manipulate the independent variables in a study to better understand how they affect the dependent variable (Spector & Meier, 2014). The manipulation of variables is often done to demonstrate a cause and effect (Lucero et al., 2016). Another characteristic of experimental design is the ability to randomize the data collection and utilization (Roosta, Ghaedi, Daneshfar, Sahraei, & Asghari, 2014). This study was trying to determine the correlation between the independent and dependent variables and not to determine the causation for that relationship. Since an experimental design is utilized to determine what the relationship between independent and dependent variable is it was not be appropriate for this study.

A quasiexperimental design is similar to an experimental design, except it lacks the aspects of randomization utilized by an experimental design (Campbell et al., 2015). Quasiexperimental designs are often utilized to evaluate the impact of a variable on a process (Campbell et al., 2015). Quasiexperimental design was not appropriate for this study because I was not trying to determine the specific impact variables have on one another, rather I was simply looking for a possible relationship between variables. While quasi-experimental design would have worked for this study at this phase it would have required more in-depth research that may not be necessary if no true relationship existed between variables.

Population and Sampling

The population for this study included students and faculty from higher education institutions in the computer science field located near Huntington West Virginia. Limiting the population to a specific field made it easier to narrow down search results to a specific knowledge area that all participants were familiar with. Including both faculty and students in the study helped to broaden the age and educational background of participants.

The sampling method for this study was nonprobability sampling. In nonprobability sampling each participant does not have a known probability to be selected to participate in the study (Raschke, Krishen, Kachroo, & Maheshwari, 2013). When utilizing a nonprobability sampling method researchers are not able to generalize the results of the study (Acharya, Prakash, Saxena, & Nigam, 2013). Nonprobability sampling methods are often utilized when the researcher does not have enough information about the population (Raschke et al., 2013). Convenience sampling was the subcategory of nonprobability sampling that was utilized in this study. With convenience sampling, everyone within the target population who was willing to take part in the study could do so (Acharya et al., 2013). This sampling method worked well with this study because it allowed everyone who wanted to participate the opportunity to do so regardless of the type of participant. The exact makeup of the population being sampled was not known, therefore trying to limit to specific groups would be difficult. Responses from all participants that completed the survey were loaded into an Access database which was queried to pull random rating and ranking data for each site within the result set for each group of participants. This data was then utilized to calculate the discounted cumulative gain for both the user data and the ranking algorithm data. The ranking algorithm's ranking data was derived from the location each site was found in when doing a simple search from a fresh browser installation.

For this study, I have utilized G*Power 3.1 to conduct an F-test for linear multiple regression to calculate a priori the required sample size given the effect size, the error probability, the power, and the number of predictors. G*Power is a statistical software designed to help determine appropriate sample size using power analysis (Faul, Erdfelder, Buchner, & Lang, 2009). Doing an a priori power analysis, assuming an effect size (f = .30) and an error probability (a = .05), indicated a minimum sample size of 36 participants is required to achieve a power .80. Increasing the sample size to 75 will increase the power to .99. The effect size (f = .30) and error probability have been derived from correlational studies that utilized multiple linear regression for data analysis (Hoxha, 2017; Johnson, 2017; Udom, 2017). These studies all utilized an effect size of .30 and an error probability of .05. In addition to other researchers utilizing a medium effect size in studies where they used the same methodologies studies have been conducted that validated the appropriateness of utilizing a medium effect size in scientific research (Bosco, Aguinis, Singh, Field, & Pierce, 2015; Eisend, 2015). Therefore, I utilized a medium effect size and seek between 36 and 75 participants for this study (Figure 1).

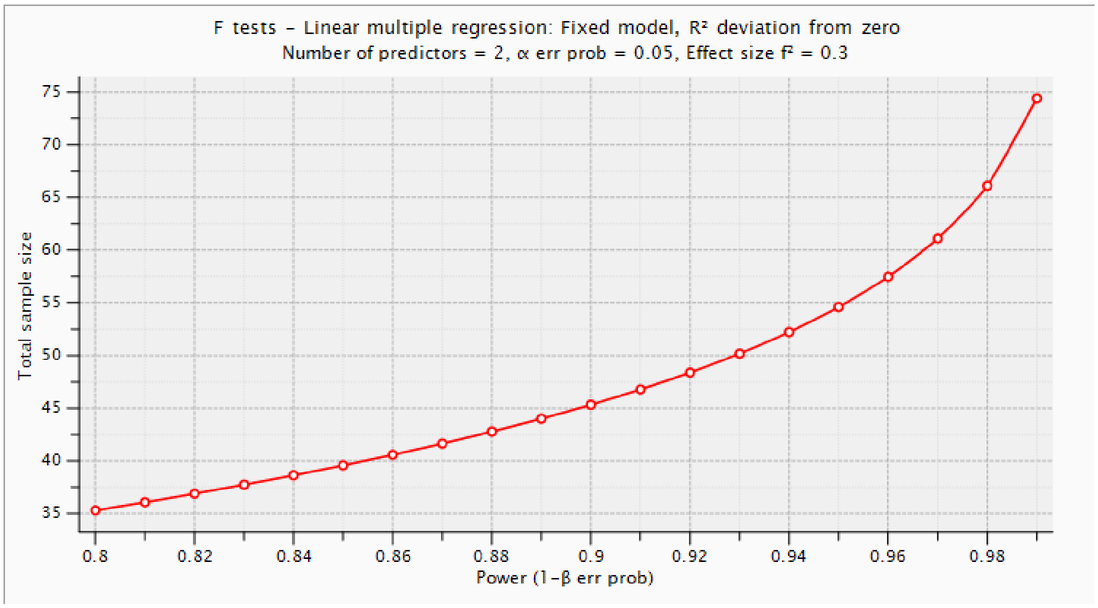

Figure 1. Power as a function of sample size.

Ethical Research

In this study many ethical considerations were taken into account before any communication with potential participants occurred I obtained approval from the Institutional Review Board at Walden University to ensure my study followed appropriate ethical considerations. Participants in research must sign up willingly and understand exactly what they are agreeing to participate in (Stang, 2015). Invitations to this study were sent out via email which contained information regarding the purpose of the study, how their responses would be protected, how they could withdraw from the study, and a link to the questionnaire itself, which can be found in Appendix A. Participation in this study was entirely voluntary and no incentives were given to participants. While offering incentives to participate in the study may help to attract participants, this can lead to biased responses from participants that are only interested in what they get out of participating (Ardern, Nie, Perez, Radhu, & Ritvo, 2013). When the user clicked the link provided in the email they were presented with a consent statement that they had to agree to before beginning the survey. By agreeing to this statement, participants acknowledged that they understand this information and were willing to participate in this study.

Throughout the study participants could decide that they no longer wanted to participate. Informed consent to participate in research is not a one-time thing at any point during the study if the participant changes their mind they should be able to withdraw from the study (Marshall et al., 2014). This study was conducted with an online questionnaire, see Appendix A, at any time during the questionnaire participants could leave. When data was compiled any incomplete questionnaires were considered as withdrawn and deleted from the results. In addition to this, at the end of the questionnaire participants were presented with a question asking if they would like their responses included in the study. If they answered no their responses were not included in the results of the study.

Part of my responsibility was to ensure that the participants privacy was protected, and if at any point an individual participant's responses are needed when presenting the results of the study I must be able to present this without any information that might identify the participant (Dempsey, Dowling, Larkin, & Murphy, 2016). In order to do this, all participants were given a unique identifier. This unique identifier will allow me to call out any specific participant information needed to explain the results without using any identifying information. Similarly, the names of the organizations where participants were found should be protected. For this study, the names of these organizations will be known only by me and will not be included in any of the study's findings. To ensure the data remains secure and meets all of the IRB requirements outlined by the University, all data collected for this study has been stored within a password protected cloud-based system and is accessible only to me. To help ensure participant's privacy no personally identifying information was collected. All data will be maintained for a period of five years. When this period is over, all data will be destroyed.

Data Collection

Instrumentation

This study utilized a close-ended questionnaire, comprised of Likert scaled questions for the data collection instrument. The questionnaire that will be utilized can be found in Appendix A. The Likert scale was developed by Rensis Likert in 1932 (Artino Jr, La Rochelle, Dezee, & Gehlbach, 2014). In this study, I measured participants' opinions of websites that would have been presented to them given a specific search term. A Likert scale is an interval scale of measurement commonly utilized to measure someone's opinions about a given idea, this is commonly done utilizing a one to five scale ranging from one being strongly dislike and five being the strongly like (Wu, Jia, & Enders, 2015). For this study one indicated the participant strongly dislikes a website, two was dislikes a website, three was neither likes nor dislikes a website, four was likes a website, and five indicated a strong like for a website. This method of rating is commonly seen on websites to get user feedback about items and can then be utilized to help personalize the user's experience on the site (de Sáa, Gil, González-Rodríguez, López, &

Lubiano, 2015). Given that this method of rating items is common in existing sites it was fitting to have participants use a similar scale to rate items within this study.

This study was administered using a web-based survey tool called Surveymonkey.com. The data will be maintained on Surveymokey.com for a period of five years, during this time I will be the only person with access to the data stored on Surveymokey.com. Data will be made available upon request. After five years I will delete all data from Surveymonkey.com. Upon deleting the data Surveymonkey.com will maintain a backup of the data for 90 days, after 90 days Surveymonkey.com purges all data associated with the deleted survey.

Participants in this study were presented with a series of questions pertaining to how they utilize the internet as well as a series of web pages and a search term that was utilized to find the webpage. Web pages were rated using a one to five scale, one being a strong dislike for a website and five being a strong like for a website. This rating provided me with an ordinal value that was utilized to determine how participants feel about specific sites returned for a given search term. In addition to the ratings, participants were asked to order the websites previously rated in the order they felt they should appear in the result set. This provided me with an ordinal value that was utilized to sort the participant’s ratings. Questions regarding the participant’s internet usage provided me with a nominal value that was utilized to group similar participants together. The values provided by user ratings were utilized to calculate the discounted cumulative gain for the search results provided. The discounted cumulative gain is a common measurement utilized to measure how effective a search engine result set is (Liu, Song, et al., 2016). The nominal data provided in regards to the participant's internet usage was applied to the user-driven parameters variable. The user-driven parameters variable was a nominal value, this variable was utilized to group participants together into three groups and is represented by Games/Media, News/Research, and Social Media in the model. The ordinal data provided in regards to the participant's ratings and ordering of web pages was utilized to calculate the discounted cumulative gain which was applied to the user rating variable. The user rating variable was an interval value. These ratings were also applied with the ordering of web pages by the ranking algorithm to calculate a discounted cumulative gain for the ranking algorithm which was applied to the algorithm ranking variable. The ranking algorithm was an interval variable.

A Likert type scale can be used in many different types of ways (de Sáa et al., 2015). Netflix has used a rating system very similar to allow its users to provide opinions about what they are watching, in fact, portions of this data was made public during the Netflix prize (Wei et al., 2017). Wei et al. utilized data from Netflix to help explore possible solutions to the cold start problem suffered by collaborative filtering algorithms (2017). In general, recommendation systems often utilize some type of Likert type scale (Parsons & Ralph, 2014; Yeung, 2016). Unless the site is only looking at what the user is doing on their site in order to recommend other items they must collect some type of feedback from the user, which you will see very often via a star-based rating system as well as some type of user preference data to help group similar users together (Guo & Chen, 2016). The questionnaire was utilized with this study, see Appendix A, contains a combination of questions that allowed for similarities among participants to emerge.

Some basic internet usage questions, as well as the rating of various websites, could be combined to allow for various types of collaborative filtering to be tested.

To address the validity of an online questionnaire using the Likert scale as a data collection instrument I have found studies measuring similar data. Bodoff and Ho utilized a Likert like scale to have participants gauge the personalization of given pages (2016). This study was not focusing on personalization of pages but having participants rank sites overall, but the underlying concept is similar. Similarly, McKechnie and Prithwiraj used a Likert type scale to measure different areas of satisfaction for specific product listings on a webpage (2016). While this study was not looking at specific items on a web page the idea is very similar, thinking of the web pages participants rated as different products that were listed. Zhao, Zhang, Cheng, and Chen utilized a Likert type scale from data collected from MovieLens to test a rating based collaborative filtering technique (2014). This study was extremely similar to this, instead of rating movies, this study was rating websites.

To address the reliability of this study I utilized the internal consistency reliability methods. Internal consistency looks for similarities between participant responses. This study looked at concepts of recommendation engines, therefore, the comparison of user responses had to occur, during this analysis, the reliability of data collected was examined. If major discrepancies were found this data was removed from the study. It is understood that the participants had their own opinions and some differences would occur, but major discrepancies among similar users could have indicated an unreliable participant.

While some of the studies that were mentioned above utilized a scale of one to seven, this study only be utilized a scale of one to five. This change was made to mimic the five-star rating system commonly seen on the web. A copy of the data collection instrument that was utilized for this study can be found in Appendix A**.**

Data Collection Technique

Data for this study was collected using an online survey. The questions on this survey gave the participant a search term and a website that would be presented with the given search term and then asked to rate that site on a one to five scale, one being the worst and five being the best. This series of questions allowed me to see how the participants felt about each site individually within the search result set and help to provide a baseline to compare the results from other algorithms. After rating the sites, the user was then asked to order the sites how he/she believes they should be ordered by a search engine. These questions allowed me to calculate the discounted cumulative gain for the users' ratings. In addition to rating and ordering sites, participants were asked a series of questions about how they utilize the internet. Internet usage questions were used to group participants together with similar participants. While other methods to group users together exist, such as tracking user history, these methods require a lot of historical data. Collecting historical data would not be feasible for this study, therefore, collecting internet usage data was sufficient to group similar participants together.

The results of the participant data were utilized to calculate the discounted cumulative gain. Discounted cumulative gain is an analysis tool commonly utilized to measure the quality of the ranking of a search engine result set (Benabbou et al., 2017;

Jayashree & Christy, 2015). In this study, I was examining the correlation between user-driven parameters, user ratings, and ranking algorithms and the discounted cumulative gain provided a perfect mechanism for assigning a single value to each result set that could be easily compared. The discounted cumulative gain is calculated by getting the summation of (*User rating*)/Log2(*Position in results* + 1) for all sites within the result set (Almulla et al., 2015).

For this study, I looked at the discounted cumulative gain for individual users as well as the average score for all users. The average score would be representative of all users opinions. In addition to this, I looked at how the addition of user preferences affect the ranking algorithms, to do this, participants were grouped with similar users. User similarity was determined based on responses to how they utilize the internet. The discounted cumulative gain was calculated for these groups like was done for the entire participant population. The grouping of participants allowed me to determine the difference, if any, that existed between standardized algorithms and algorithms that focus more on user preferences.

There are other methods utilized to determine the effectiveness of search engines. The first method is simply the cumulative gain. Cumulative gain simply looks at the result set returned to the user and does not take into account the position of the results (Benabbou et al., 2017; Jayashree & Christy, 2015). This method is actually a portion of the discounted cumulative in which the scores of each site in the result set are combined to give an overall score of the results returned. While this value can be useful when looking at the search aspect of a search engine it does not provide any feedback for the ranking aspect, which is what this study was interested in.

Online survey data collections have many advantages that make it ideal for this study. Online surveys are relatively easy to administer (Bateson, McPeake, & O'Neill, 2014). Many different tools are available to help administer online surveys for little or no cost which makes them a cost-effective option (Rice, Winter, Doherty, & Milner, 2017). In addition to being cost effective and easy to administer online surveys also make it easier to reach a larger number participants and do not have any major time constraints for the participants (Khazaal et al., 2014). This study was focusing on the input from students and faculty in the IT fields of higher education institutions, which can result in difficulty finding a time when all potential participants can be available to take part in the study. For this study, I was primarily interested in how the participants feel about sites within a set of search results which is information that can easily be collected at the participant's convenience. An online survey was easily set up and sent out to a pool of potential participants to fill out whenever they had the time to do so.

While there are advantages to online surveys, there are also disadvantages. One of the biggest disadvantages of online surveys is participants misrepresenting themselves (LaRose & Tsai, 2014). Since online surveys are conducted without the researcher present there is no way to ensure that the individual participating in the study is who they say they are. Many survey websites, like surveymonkey.com, provide features to help reduce these types of situations, for example, surveys can be made private and require a password to access. While this is not a foolproof method to completely eliminate this concern it can help to alleviate this issue.

Data Analysis

The research question for this study was, what is the relationship between user-driven parameters, user ratings, and ranking algorithms? The null hypothesis was, there is no statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms. The alternative hypotheses for this study was, there is a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms.

Data sampling was done using convenience sampling. Responses from all participants that completed the survey were loaded into an Access database which was queried to pull random rating and ranking data for each site within the result set for each group of participants. This data was then utilized to calculate the discounted cumulative gain for both the user data and the ranking algorithm data. The ranking algorithm's ranking data was derived from the location each site was found in when doing a simple search from a fresh browser installation.

The results of participant data were analyzed using multiple linear regression. Multiple linear regression is utilized when two or more independent variables are utilized to help the researcher predict the value of a dependent variable (Woodside, 2013). Simple linear regression only looks at the relationship between a dependent variable and a single independent variable (Enayatollahi, Bazzazi, & Asadi, 2014). Other types of regression analysis, such as hierarchical and stepwise regression, allow some form of control over the order in which independent variables are included in the regression equation (Hanisch & Rau, 2014). In this study, I was determining if a relationship exists between two independent variables and one dependent variable and did not require any control over the variables, therefore traditional multiple linear regression was appropriate for this study. The general linear regression model for this study was Ranking Algorithm = B_0 + B_1 (Respondent DCG) + B_2 (User Interests) + B_3 (Search Term) + B_4 (Search Term)(User Interests) + B_5 (Search Term)(Respondent DCG). The ranking algorithm variable is an interval variable that represents the discounted cumulative gain value of result set provided by the ranking algorithm. The user ranking variable was an interval variable that represented the discounted cumulative gain value of the result set provided by the user ranking. The discounted cumulative gain was a continuous number that represented the effectiveness of ranking methods. The user-driven parameter variable was a nominal variable that represented the group participant responses belonged to. For this study, participants were grouped using the questions about how they primarily use the internet. Asking questions like this is commonly utilized in collaborative filtering methods to help provide recommendations to users before they have provided any information about items within the system (Chen et al., 2017). Feng et al. (2019) utilized the groups of Facebook usage, study usage, and entertainment usage. For this study I utilized similar groupings for the user-driven parameters, which were broken down into four groups, Games/Media, News/Research, Social Media, and Overall which contains participants from each group, these groupings were then utilized to compute discounted cumulative gains for specific groups. In order to utilize a nominal variable as an independent variable in multiple linear regression it is necessary to modify this variable. Each group was assigned a numeric value, 1 = Social Media, 2 = News/Research, 3 = Games/Media, and 4 = Overall. In addition to grouping participants the type of search term was utilized. The type of search term was utilized to differentiate between the search term all participants should be familiar with within their field of study and a more generic term. For this study all participants were students or faculty in the IT field. The search term “Programming” was utilized as a search term that all participants should be familiar with. The search term “Recipes” was utilized as a generic search term that participants may or may not be familiar with. The search term was a nominal variable where a value of 1 = Programming and 0 = Recipes.

B_0 is the intercept value which indicates the value where the regression line meets the y-axis, this indicates the point on the regression line where the user ranking and user-driven parameter values are all 0. B_1 indicates the slope of the regression line, the slope of the line indicates the strength of the correlation between the variables. A negative slope would indicate a negative correlation, for this model it would show that as the ranking algorithm increases the user ranking decreases. B_4 and B_5 represent the possible interactions between the independent variables. These were included to determine if the interaction between the independent variables have a significant effect on determining the Algorithm DCG.

For this study, the participant was asked questions about how he/she primarily uses the internet that was utilized to group similar participants together. Participant groups were represented by a nominal value that represented the user-driven parameters.

Multiple discounted cumulative gains were then calculated. The discounted cumulative gain was calculated for the entire participant pool as they would order the results, the entire participant pool as the algorithm would order the results, the participants grouped with similar users as they would order the results, and the participants grouped with similar users as the algorithm would order the results. Calculating the discounted cumulative gain for each scenario assigns a score for the result sets for both the user ratings and the ranking algorithm. The discounted cumulative gains are an interval value that represented the algorithm and user ranking variables.

Analysis of the data was done using SPSS. Many different statistical values are available from SPSS. For this study, I was primarily interested in the R^2 and the sig (*p)* values. R^2 is a value ranging from zero to one which indicates how well the independent variables can be utilized to predict the dependent variables. The sig (*p*) value, is utilized to determine if the null hypothesis is true. The *p*-value can range from zero to one. A *p*-value less than .05 indicates strong evidence against the null hypothesis. A *p*-value > .05 indicates weak evidence against the hypothesis.

The normality assumption is the assumption that there is a normal distribution of variables (Korkmaz, Goksuluk, & Zararsiz, 2014). Normality can be tested for with a Normal Probability Plot (P-P) of Regression Standardized Residuals and a scatterplot of the standardized residuals. The points on the P-P Plots should form a straight line (Zhou & Shao, 2014). In situations where normality is violated transformation values can be applied to the variables to try to correct this issue (Zhou & Shao, 2014). This study did not require transformations to be applied.

Linearity is the assumption that a linear relationship exists between the independent and dependent variables (Hopkins & Ferguson, 2014). Linearity can be tested with a scatterplot. A linear relationship should be able to be seen. In situations where linearity is violated, a transformation can be applied to the independent or dependent variables to correct linearity issues. A common transformation is to apply the log of the independent or dependent variable (Jamshidian, Jalal, & Jansen, 2014). Transformations help when linearity is not met by helping to make the data better fit the chosen model. In this study, no transformation was needed.

Multicollinearity is a correlation between independent variables (Yu, Jiang, & Land, 2015). In situations where the independent variables are not truly independent of one another, the results of the study can be biased (Yoo et al., 2014). Winship and Western discuss that correlations between independent variables should not be higher than .80 (2016). In order to assess multicollinearity, a correlation matrix can be utilized (Winship & Western, 2016). A correlation matrix is simply a table of correlations between variables. For this study, correlations were calculated using SPSS and put into a correlation matrix to assess multicollinearity. In addition to a correlation matrix, multicollinearity can be tested by calculating the Variance Inflation Factor (VIF) where values above 10 indicate multicollinearity (Yoo et al., 2014).

Homoscedasticity assumes that the variance of errors is similar among the values of the independent variables (Hopkins & Ferguson, 2014). Homoscedasticity can be tested for using a scatterplot, if a clear pattern exists homoscedasticity may be a concern. In situations where homoscedasticity is violated a transformation of variables can help to reduce the issue. Taking the square root of the independent or dependent variables is a common transformation. (Hopkins & Ferguson, 2014). In this study, no transformation was needed.

Bootstrapping can be utilized in situations where the violation of an assumption occurs (Banjanovic & Osborne, 2016). Bootstrapping helps to determine standard errors of the coefficients for the predictor variables and helps address issues with a random sampling (Bro & Smilde, 2014). In some situations, data may also need to be adjusted in order to address violations of assumptions in multiple regression analysis by applying a logarithm or square root to the variables (Hopkins & Ferguson, 2014). For this study, no transformation was needed.

Data for this study did not require any additional clean up, however ordering of the data was required to calculate the discounted cumulative gains. Site rankings were ordered from highest to lowest in order to calculate an ideal discounted cumulative gain for each participant (Benabbou et al., 2017; Jayashree & Christy, 2015). For this study, I assumed that participants provided honest rankings for the sites provided, but exceptions to this might have occurred. To help alleviate this, responses were examined for any major outliers as well as responses of all the same rankings for every site. No major outliers were found in the data.

For this study, only complete response sets were utilized. Any incomplete questionnaires were viewed as a withdraw from the study. This decision was made based on the complete rule of Arrow's impossibility theorem which states the social choice rule should provide a complete ranking of all alternatives (Morreau, 2015). This means that in order to have a vote that is truly representative of the group as a whole every possible item must be ranked by every participant. In reality, this would be extremely difficult, if not impossible to achieve, but for the purposes of this study, it was followed, as it showed under perfect circumstances what could be possible. Recommendation engines are a good example that shows the more items that are ranked the more accurate the results can be (Aghaiipour-Chafuchahi & Ahmadi-Abkenari, 2015; Gleich, 2015; Wei et al., 2017). The idea for this study was if all items are ranked, the recommendations, or result rankings should be more accurate than if only some are ranked.

I utilized an F-test to validate the model derived from my data analysis. Utilizing an F-test allowed me to assess multiple coefficients at the same time. The F-test compares the statistical model to an intercept only model and determines if the new model is significantly different from the intercept only model. This test can be utilized to determine whether or not to reject the null hypothesis and to determine if the statistical model provides a better fit for the data then the intercept only model.

To analyze the data in this study I utilized SPSS. I explored other tools, such as Microsoft Excel 2016. While I was more experienced with Excel, SPSS provided additional features that helped validate the analysis of data.

Study Validity

External validity needed to be addressed for this study. The external validity of a study is how the results of the study can be applied to populations other than those that participated in the study (Rouf, Grech, & Allman-Farinelli, 2017). For this study the population was very specific, higher education faculty and students in the IT field in

Huntington WV, therefore care had to be taken to ensure the study’s results were applicable to other areas. Both faculty and students were included in order to get a wider range of experience and age. The decision to limit the field of study to IT was made simply to limit the topic of websites being reviewed to a specific topic that all participants will be familiar with. For this study, it was not feasible to open the study to any topic available, by limiting to a specific population the subject matter could be controlled to that group's knowledge area. The questions that were asked of participants were simple satisfaction ratings of websites and were not specific to a single area. The main assumption of this study was that all participants answered questions honestly and they did not try to sway the results in any way. In order to address the reliability of participants answers, I utilized the internal consistency reliability method.

Internal validity was not a major concern in this study. Internal validity deals with the connections between the dependent and independent variables (Abowitz & Toole, 2010; Halperin, Pyne, & Martin, 2015; St. Clair, Cook, & Hallberg, 2014). The collection of data from participants did not truly affect the internal validity at all, because this data was only collected and then utilized for the study, any manipulation that occurred was done to the algorithms being tested. The nature of this study helped to ensure that there was no issue with the internal validity. The data provided by participants was be fed into various ranking algorithms and the results were compared. By simply running these various algorithms the internal validity was tested against the various algorithms.

Statistical conclusion validity was not a major issue in this study. Statistical conclusion validity looks at whether or not the relationship found between variables is correct, specifically looking at if a correlation is found and none actually exists or if no correlation is found and one actually exists (Aravamudhan & Krishnaveni, 2015; Kratochwill & Levin, 2014). While this was something to be concerned about, the way this study was conducted these types of errors should be greatly reduced. Each set of responses were tested against various algorithms so any relationships between variables that may or may not be found were explored against different data sets, by doing this I was able to determine whether the observed relationships occurred reliably across multiple algorithms.

While this study was conducted in a limited area the results will be able to be generalized to other areas as well. None of the questions asked were specific to a single area and therefore the location of the study did not have any effect on the results. The study was limited to a specific field of study, which could have potentially caused problems with generalizing the results, in order to help with this two types of sites were rated within the study, sites within the IT field and then another generic topic that is not field specific. By including both types of sites the study was able to obtain ratings for two different types of searches, research and entertainment based search results. Multiple types of sites allowed me to determine whether or not the result sets were applicable to specific field type searches or if they only applied to more generalized searches. This differentiation was important to this study because users do search for things out of their areas of expertise and the ranking method utilized may or may not prove to be as successful when searching for an unfamiliar topic. Participants could potentially have utilized different browsers when completing this survey, while this could cause some sites to appear differently I chose not to try to control this. Requiring a specific browser to be utilized would not be reflective of every day usage by the participant, rather than requiring this to be done the participant was asked what browser he/she was utilizing and was then be applied to user preferences to determine if the browser utilized had any effect on his/her ratings. A power analysis was conducted to determine the appropriate sample size. To achieve a minimum power of .80 36 participants were needed, increasing the sample size to 75 increased the power to .99, therefore I sought between 36 and 75 participants for this study.

Transition and Summary

In Section 2, I provided a description of the role of the researcher, the research method and design, population and sampling, data analysis, ethical considerations of the participants, and issues concerning validity and instrumentation. The role of the researcher section detailed how I am responsible for formulating a hypothesis, collecting and analyzing data, formulating conclusions, and how I will separate myself from personal experiences in order to conduct the study. The research method and design section provided a more detailed look at quantitative methodologies and experimental designs and why they have been chosen for this study and why other designs and methodologies were not appropriate. Population and sampling identified college students and faculty in the Huntington, WV area as the target population for this study with a minimum of 36 participants being required based on a priori power analysis. The data analysis section looked at the discounted cumulative gain analysis tool commonly utilized to measure the ranking quality of search engine result sets and why this method is appropriate when looking at the results of this study. Ethical considerations looked at how the participants will be protected throughout this study, specifically explaining how no personal information will be linked to the participant's responses. The validity section examined issues that can occur with study validity and how I will work to alleviate these concerns as needed.

Section 3 will provide detailed information regarding the findings of this research study. This will include a presentation of findings, possible applications to professional practice, implications for social change, recommendations for action, and reflections of my experience while working on the study.

Section 3: Application to Professional Practice and Implications for Change

This study utilized a correlational quantitative research method that analyzed the relationship between user-driven parameters, user ratings, and ranking algorithms. In this section I will present the results of the analysis of the data gathered through the online surveys completed by the participants of the study.

Overview of Study

The purpose of this correlational quantitative study was to examine the possible relationship between user-driven parameters, user ratings, and ranking algorithms. Using G*Power 3.1, I calculated the required sample size given the effect size, the probability, the power, and the number of predictors. The analysis indicated that a minimum of 36 responses would achieve a statistical power of 0.80 while 75 responses would increase the statistical power to 0.99. I gathered data from 47 participants including both faculty and students in the IT field at a local university in Huntington, WV. Survey invitations were sent out to a total of 178 possible participants meaning that I received a response rate of approximately 26%. For this study I utilized a 95% confidence interval and anything above a 0.05 significance level indicated that a significant relationship did not exist.

The results of the data analysis showed that there was not a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms. The user ratings did show a significant relationship to the ranking algorithms, but the user-driven parameters did not show a statistically significant relationship, therefore I cannot say with any statistical significance that there is a relationship between all three variable.

I must reject the alternative hypothesis of, there is a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms**.**

Presentation of the Findings

In this portion of the study, I will analyze the methods used to test the assumptions involved with the methodology, present the statistical results of the data analysis, provide a detailed reporting of the findings, and summarize the findings.

Test of Assumptions

In Section 2 I presented multicollinearity, normality, linearity, homoscedasticity, and independence of residuals as assumptions evaluated in this study. I evaluated these assumptions and present the findings below, which did not indicate any major violations to these assumptions.

Multicollinearity. I utilized a correlation matrix to test for multicollinearity within my data. Table 4 depicts the bivariate correlations within my data. The bivariate correlation between variables the variables Search Term and Algorithm DCG indicated a fairly strong correlation (.709), as well as between the interaction variables, therefore additional testing was needed to ensure multicollinearity was not a concern.

Table 4

Bivariate Correlations

	Algorithm DCG	User DCG	User Interest	Search Term	Search Term * User Interests	Search Term * User DCG
Algorithm DCG	1	.428	.057	.709	.627	.828
User DCG	.428	1	.051	-.210	-.132	.008
User Interests	.057	.051	1	-.011	.372	.008
Search Term	.709	-.210	-.011	1	.944	.966
Search Term * User Interests	.627	-.132	.372	.844	1	.830
Search Term * User DCG	.828	.008	.008	.966	.830	1

For additional testing I utilized the Variance Inflation Factor (VIF) where values above 10 indicate multicollinearity (Yoo et al., 2014). The interaction variables presented problems for multicollinearity because the VIF values were above 10. Removing the interaction terms from the model corrected this problem. Table 5 shows the VIF values for my variables after the interaction terms were removed. The VIF values for Respondent DCG, User Interests, and Search Term were 1.049, 1.003, 1.046 respectively. All the VIF values were much lower than the threshold of multicollinearity, which is a good indicator that multicollinearity was not an issue.

Table 5

VIF Values

			Collinearity Statistics	
	t-Statistic	Sig.	Tolerance	VIF
Constant	.612	.004		
Respondent DCG	.586	.000	.954	1.049
User Interests	.026	.324	.997	1.003
Search Term	1.366	.000	.956	1.046

Normality, linearity, homoscedasticity, and independence of residuals. Normality, linearity, homoscedasticity, and independence of residuals were evaluated by looking at the Normal Probability Plot (P-P) of Regression Standardized Residual (Figure 2) and the scatterplot of the standardized residuals (Figure 3). Examining these plots did not indicate that there were any major violations of these assumptions. The points on the P-P Plots lie close to a straight line which is a good indicator that there was not a major violation of normality (Zhou & Shao, 2014). The scatterplots did not indicate a clear pattern of the standardized residuals which supported the assumptions being satisfactory.

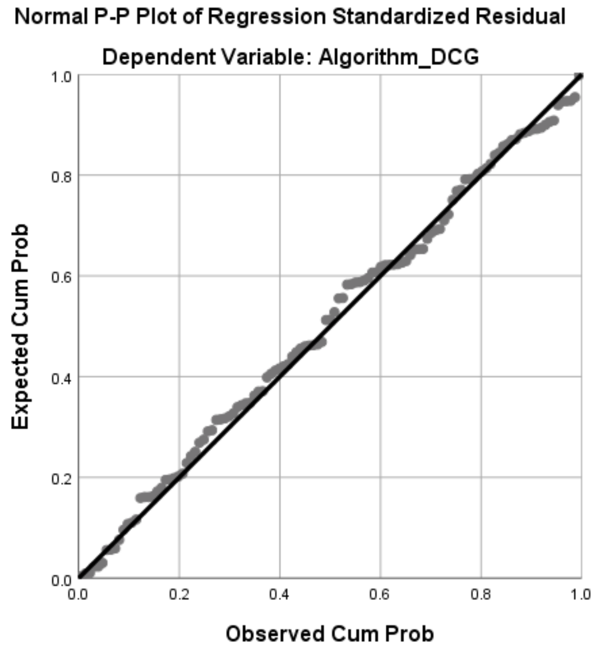

Figure 2. Normal P-P Plot of regression standardized residual.

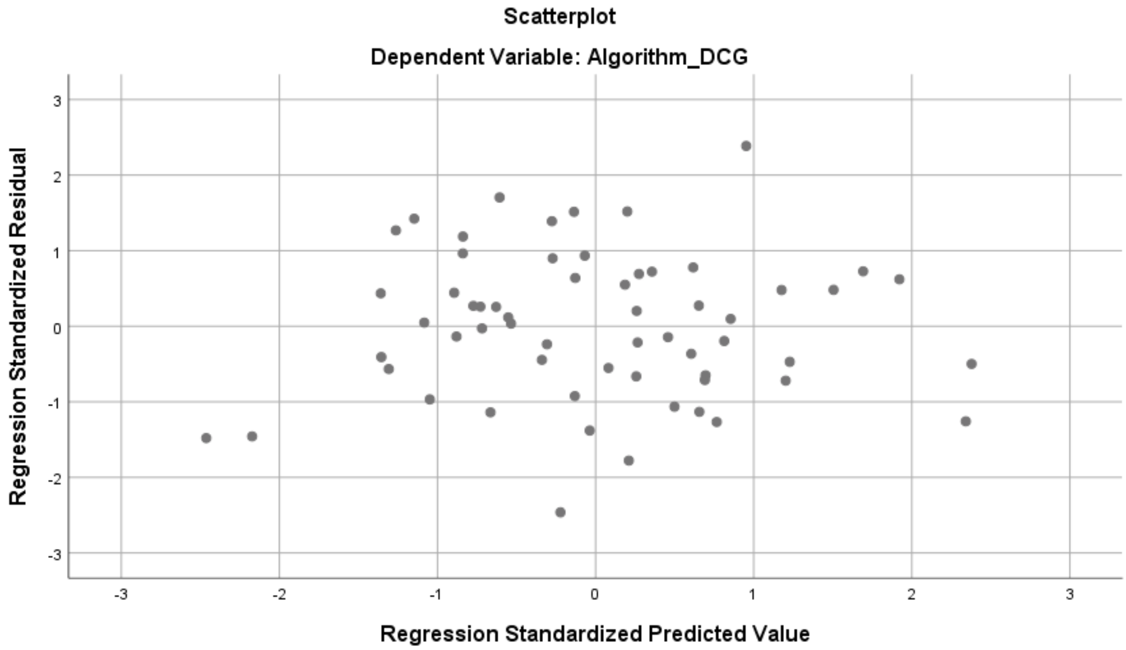

Figure 3. Scatterplot of standardized residuals and predicted values.

Descriptive Statistics

I received a total of 59 responses to my survey. 12 responses were eliminated due to missing data, resulting in 47 responses utilized in the data analysis for this study. Of the 47 usable response sets only six were faculty members and 41 were students. The small response set from faculty members did not warrant breaking down these demographics within the analysis. Similarly, the preferred search engine 43 preferred Google, three preferred DuckDuckGo, and one preferred Yahoo. The small response set that preferred a search engine other than Google did not warrant breaking down these demographics within the analysis. The age of respondents ranged from 18 to 67. Of the usable respondents nine were 18 to 19, 23 were 20 to 29, 10 were 30 to 39, two were 40 to 49, two were 50 to 59, and one was over 60. The browsers utilized to fill out the survey was primarily Chrome with 39 participants. The remaining eight participants utilized

Safari. The browser the participant utilized was included in the questionnaire in case any major outliers were found in the responses to determine if the browser may have caused this. No outliers were found therefore the browser the participants utilized was not needed for the analysis. The responses contained ratings for two sets of seven websites as well as how the participant would order those websites within a search result set. This provided 329 ratings for the keyword “Programming” and “Recipes.” Responses from all participants that completed the survey were loaded into an Access database which was queried to pull random rating and ranking data for each site within the result set for each group of participants. This data were then utilized to calculate the discounted cumulative gain for both the user data and the ranking algorithm data. The ranking algorithm’s ranking data were derived from the location each site was found in when doing a simple search from a fresh browser installation. No outliers existed within the data collected that required any data manipulation. These responses were utilized to calculate the discounted cumulative gain for each site rated. The discounted cumulative gain was calculated by getting the summation of (*User rating*)/Log2(*Position in results* + 1). Table 6 shows the descriptive statistics for the data collected.

Table 6

Descriptive Statistics

Variable	Mean (M)	Standard Deviation (SD)	Bootstrapped 95% Confidence Interval (CI) (M)
Respondent DCG	5.35	.84	[5.19, 5.49]
Algorithm DCG	4.50	.82	[4.35, 4.65]
User Interests	2.5	1.1	[2.31, 2.72]
Search Term	.50	.50	[.41, .58]

Note: N=120.

Inferential Statistics

This study used standard multiple linear regression, $\alpha = .05$ (two-tailed), to examine the effectiveness of the user-driven parameters, and the Search Term with the Respondent DCG in predicting the Ranking Algorithm's DCG value. The independent variables were user interests, search term, and Respondent DCG. The dependent variable was the Algorithm DCG. The null hypotheses and alternative hypothesis were:

Null Hypothesis (H_0): There is no statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms.

Alternative Hypothesis (H_1): There is a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms.

Table 7

Model Summary

Model	R	R^2	Std. Error	Sig. F Change
Model	.923	.852	.320	.000

a. Predictors: (Constant), User Interests, Respondent_DCG, Search Term

The model as a whole was able to significantly predict the ranking algorithm's discounted cumulative gain, $p = .000$ *and* $R^2 = .852$ for the data collected (Table 7). The R^2 value for the data collected indicated that the model could explain 86% of the total variability in the ranking algorithm's discounted cumulative gain (R^2=.852, $F(3,115) = 220.13$, $p < .01$).

The final predictive equation based on the predictor variables (Table 9) was:

Ranking Algorithm DCG = .612 + .586(Respondent DCG) + 1.366(Search Term)

The model was validated utilizing a F-test. The F-test was highly significant for the data collected (Table 8). This testing indicates that the model explains a significant amount of the variance in the Algorithm DCG. This indicates that the regression model is a good fit for the data collected in the study, this means that the independent variables can significantly predict the dependent variable. This means that the regression model is more accurate than the simple intercept model.

Table 8

F-Test Data

Model	Sum of Squares	df	Mean Square	F	Sig.
Predictive Model	67.752	3	22.584	220.138	.000
	11.798	115	.103		
	79.550	118			

Table 9

Regression Analysis

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.612	.210		2.914	.004
	Respondent_DCG	.586	.036	.601	16.354	.000
	User Interests	.026	.026	.036	.990	.324
	Search Type	1.366	.060	.836	22.748	.000

a. Dependent Variable: Algorithm DCG

Search Type * Respondent DCG. Analysis was initially run including the interaction term between Search Type and Respondent DCG. For the data collected the interaction term between Search Type and Respondent DCG had a positive slope of .806 but this was not a significant predictor of the Algorithm DCG because $p > .05$. Since Search Type * Respondent DCG was not a significant predictor and the VIF value was greater than 10 this interaction term was removed from the model and analysis was done again without this interaction term.

Search Type * User Interests. Analysis was initially run including the interaction term between Search Type and User Interests. For the data collected the interaction term between Search Type and User Interests had a negative slope of .093 but this was not a significant predictor of the Algorithm DCG because $p > .05$. Since Search Type * User Interests was not a significant predictor and the VIF value was greater than 10 this interaction term was removed from the model and analysis was done again without this interaction term.

Respondent DCG. For the data collected, Respondent DCG has a positive slope of .586 which indicates that for every point of increase in Respondent DCG there is a .586 increase in the Algorithm DCG. The increase in the Respondent DCG only accounts for a partial change in the Algorithm DCG, the change must also be combined with the constant slope value, .612. This would indicate for every point of increase in Respondent DCG there is a 1.198 increase in the Algorithm DCG. These values are only representative of the overall grouping and does not take into account the subgroups utilized in this study. The respondent DCG was a significant predictor for the Algorithm DCG because $p < .05$. This means that when the overall group's DCG value increases it can be said that it will significantly predict the increase to the Algorithm DCG.

User interests. User interests has a positive slope of .026 for the data collected. This indicates that for every point of increase in the user interests there is a .026 increase in the Algorithm DCG. These changes to the Algorithm DCG must be looked at in combination with the changes indicated for the Respondent DCG variable, meaning that the change in Algorithm DCG would be 1.198 + .026 = 1.224. The user interests were not a significant predictor for the Algorithm DCG because $p > .05$. This means that it cannot be said that the user interests will significantly predict the Algorithm DCG. Since user interests was not a significant predictor it was not included in the final model.

Search term. Search term has a positive slope of 1.366 for the data collected. This indicates that for every point of increase in the DCG when searching within the user's knowledge area there is a 1.366 increase in the Algorithm DCG. These changes to the Algorithm DCG must be looked at in combination with the changes for the Respondent DCG variable, meaning that the change in Algorithm DCG would be 1.198 + 1.366 = 2.564. The search term was a significant predictor for the Algorithm DCG because $p < .05$. This means that it can be said that the search term will significantly predict the Algorithm DCG.

Analysis summary. The purpose of this study was to determine if there was a statistically significant relationship between user-driven parameters, user ratings, and ranking algorithms. I utilized multiple linear regression to examine the effectiveness of the predictor variables. The assumptions surrounding multiple regression were evaluated and no serious violations were found. The overall model was able to significantly predict the algorithm ranking discounted cumulative gain, R^2=.852, $F(3,115) = 220.13$, $p < .01$ see table 9. While the model as a whole was able to significantly predict the algorithm ranking discounted cumulative gain the user interests variable was not a significant predictor. The Respondent DCG and Search Term were the most significant predictors with p = .000. This would indicate that a significant relationship exists between user ratings and the ranking algorithm when the user is searching within his/her knowledge area.

Theoretical conversation on findings. After analyzing the data collected for my study I was able to show that the overall model could significantly predict the ranking algorithm's DCG. None of the individual groupings alone were a significant predictor for the algorithm's DCG. The participant DCG and search term however did have a significant relationship with the ranging algorithm's DCG. These findings suggest that there may be a correlation between the user's ratings and the ranking algorithm but the user-driven parameters do not play a major role in this correlation.

In studies performed to examine the applications of personalization on search ranking algorithms it has been shown that personalization techniques can provide users with more relevant search results over standard search algorithms (Ramesh & Andrews, 2015). The personalization techniques from these studies often utilized collaborative filtering techniques to group users together in order to increase the accuracy of search results. The biggest problem with collaborative filtering techniques is not having enough data to provide recommendations for new users (Wei et al., 2017). Some studies have utilized a user's interests to get around this lack of data (Guo & Chen, 2016). This study has focused on whether existing ranking algorithms utilized by search engines have a correlation with user ratings and user-driven parameters.

In this study I found that user ratings can be a significant predictor to rankings provided by existing ranking algorithms. The user groups within this study were not significant predictors by themselves for the algorithm rankings. Several studies have found that the addition of personalization factors to search ranking algorithms actually increase the reliability of existing ranking algorithms (Saxena, Agarwal, & Katiyar, 2016). Balakrishnan et al. (2015) found that utilizing user rankings for website significantly outperformed standard search engine ranking algorithms. Specifically Balakrishnan et al. found the CoRRe model utilized to be significant at three different document levels: top-five ($t_{(18)}$ = -3.049, p = 0.016), top-10 ($t_{(18)}$ = -5.158, p = 0.001), and top-15 ($t_{(18)}$ = -5.344, p = 0.01). My study differed from Balakrishnan et al. in the way I looked at the data. In my study I looked at existing result sets and had participants rank sites that were provided by the ranking algorithm instead of generating different result sets based on different criteria. While my study was different, the findings of my study align with Balakrishnan et al. (2015) in that the respondents' calculated discounted cumulative gains were higher than the algorithms calculated discounted cumulative gains, which indicates that the user rankings provide a better representation of a search results set than the algorithm ranking. My model was able to significantly predict 85% of the total variability in the ranking algorithm's discounted cumulative gain (R^2=.852, F(3,115) = 220.13, $p < .01$). Being able to significantly predict the variability indicates there is a correlation between my variables and that the addition of user ratings could potentially help to improve the discounted cumulative gain of ranking algorithms which is similar to the findings of Balarishnon et al.

Grouping users into subgroups is commonly done in an effort to further improve result sets (Guo & Chen, 2016). Suruliandi et al. (2015) looked at grouping users based on both content analysis and user interest groups and found that user groups produced higher quality results then both the search engine algorithms and personalization algorithms based on content analysis. In my study I grouped participants based on their interests pertaining to how they most often utilized the internet. The grouping data in my study indicated that while the addition of these groups did provide some change to the respondent DCG values, the change was not statistically significant. This differs from the findings of Suruliandi et al. (2015). The difference in my findings versus the findings of Suruliandi et al. (2015) does not necessarily mean that the addition of group data cannot provide more accurate results. Differences can easily be explained by the way participants were grouped in the different studies. I chose to group participants based on how they primarily utilize the internet, but drilling down even further and getting more specific regarding user interests could help to improve upon user groupings and possibly result in more statistically significant results. When search term was added within this study I did find a significant relationship for search term. This could indicate that users searching within their own content area view those websites differently than websites outside of their own content area.

In this study I looked at Arrow's impossibility theorem in relationship to ranking algorithms. Arrow outlined five rules that must be met in order to truly represent a groups opinion. Arrow states that it is impossible to meet all five rules when three or more options exist (Ben-Yashar & Nitzan, 2017). As outlined in the literature review, four of these rules are easily met by ranking algorithms, however the his "Complete" rule would essentially be impossible to meet. Some studies have suggested relaxing these rules to meet a conditional Arrow fairness (McComb, Goucher-Lambert, et al., 2017). This study supports the idea that the ranking algorithm actually meets the requirements behind conditional Arrow fairness. The subgroups were not significant predictors of the algorithm rankings, but the overall participant rankings were. This could potentially suggest that the subgroups did not contain enough information to significantly predict the algorithm ranking value, but when combined provided a much more significant prediction. This is in line with conditional Arrow fairness because as more and more data is collected the accuracy of rankings should become more in line with the group opinion as a whole.

Applications to Professional Practice

This study aimed at examining the correlation between user-driven parameters, user ratings, and ranking algorithms. The results of this study will provide content providers with a better understanding of how users view web content as a whole. With a better understanding of this topic the quality of content created may improve and help content providers to focus more on content rather than search engine optimization.

The findings of this study show that the subgroups utilized within this study were not significant predictors of the ranking algorithm by themselves, while the overall participant ratings were significant predictors. It also showed that the search term was also a significant predictor of the ranking algorithm. The overall findings can potentially be very useful to content providers who focus their content on a specific niche. The content created by these providers would most likely be focused entirely on that subgroup of interested users. While it is necessary to focus content to the interested users, it may be beneficial to expand the content to more generic terms to help reach potential new users outside of the subgroups of interest.

In this study, two search terms were utilized. The first term, “Programming,” was selected because every participant should have some general knowledge or interest about this term. The second term, “Recipes,” was selected as a generic term that everyone should know about, but may not necessarily have much interest or knowledge about. The difference in search terms was chosen to look at how searches measured with participants who had a professional interest in one term, but may or may not have a professional interest in the other. The type of search term was a significant predictor for the Algorithm DCG with $p < .05$.

The significance of the search type in this study is an important distinction for content providers to consider when creating web pages. The analysis within this study showed that the search term was a significant predictor of the ranking algorithm. Content providers can potentially utilize this information to their advantage while considering methods for search engine optimization. Understanding more about your target audience and the type of content being provided is important. Targeting generic keywords, like recipes, that users from many different backgrounds may search for, search engine optimization techniques could be more difficult. Just because a page does not make it to the first position of the search results does not mean that users do not like the page better than other pages. This is a good example of why content providers should be concerned more with the quality of their pages instead of algorithm rankings.

Search rankings are important to content providers, because the higher in a result set a page is the more likely users are to find that page. The position a page appears in a result set is not necessarily an indicator of the page's quality. Content providers need to understand this, while it is important to try to get higher in a result set, focusing on keywords within their content is not necessarily going to help in the long run. User opinions are what will truly define the success of a web page. Being the first site listed on a search engine may help drive people to a web site, but it is the content of the site that will help to bring them back.

Implications for Social Change

This study was done to help provide a better understanding of the correlation between user-driven parameters, user ratings, and ranking algorithms to content providers. The results of this study showed that the user-driven parameters alone were not a significant predictor of ranking algorithms, while the user ratings were a significant predictor. Understanding that the user ratings are actually a significant predictor regardless of the user-driven parameters can potentially lead to a significant change in how content is derived. Some content providers, for niche content, may focus less time on utilizing niche keywords in an effort to push their websites higher on the search results page and begin focusing more time on the content regardless of keywords.

When the content providers broaden the scope of keywords utilized their content would begin to fall into more searches, exposing their content to a wider range of users searching for similar content. A couple of benefits could occur as a result of broadening the scope of keywords. User's searching for more generic terms could potentially be exposed to more content that would generally require more specific search terms. This exposure with more generic terms could help users expand their knowledge new content more quickly. The need for users to expand upon their search term to be more specific could be greatly reduced allowing them to conduct their search within a single result set. These possible changes to how users can find information within a search result set can lead to users finding new topics of interest and expanding their knowledge more quickly. The faster expansion of knowledge into new topics can lead to innovations within countless topics. Innovations can in turn lead to improvements to how people live everyday life.

Recommendations for Further Research

This study was limited by the amount of data and the location of participants. Ranking algorithms require large amounts of data and this amount of data was not readily available for this study. To collect data, I chose the search terms and randomly selected websites to have participants rate. The participants were IT students and faculty in Huntington, WV, limiting to a specific field was simply done in order to be able to choose a search term that was relevant to the participants' field of interest.

Further studies could expand upon the search terms utilized and use a larger area for participants. The limitation of using a specific field could be eliminated for future studies and derive sites from various fields showing relevant search terms to participants based off of a series of interest questions. In addition, this study could utilize a larger number of participants to determine if different topics could produce different results.

Future studies could also expand on the groupings utilized in this study. The groups within this study were generic based upon how participants primarily utilized the internet. Expanding upon these groups to more specific user groupings could produce very different results. Another possible approach would be to utilize the same groupings with a larger participant pool and further divide those groups based upon how participants rated each site similar to how many collaborative filtering algorithms work.

Future researchers can also utilize this study as a source to try to develop new ranking algorithms. This study showed that the user-driven parameters were not significant predictors of existing ranking algorithms. The development of new ranking algorithms that focus on user-driven parameters is still a potential research topic. Future studies could try to modify algorithms to try to base results on user-driven parameters in an effort to improve upon result sets for specific users.

Reflections

I have spent my entire professional career working in higher education, as a result furthering my education became a goal that I truly wanted to achieve. This entire process has been an amazing experience and has helped me find new interests and directions for my career. Finding and reading the many different studies and looking at them and how the different topics can be utilized together has truly changed the way I think. In my current job I have found myself doing a great deal more research and thinking more outside the box to try to improve the quality of service I provide.

I have learned a great deal throughout this process about myself and worked very hard on improving some professional shortcomings that I had, such as organizational and time management skills. This has been a long process for me, as I had several personal issues arise that have required me to shift focus away from my study and to my family and friends. Despite this I would not trade this experience for anything because I truly feel it has helped me improve both my professional and personal life.

Summary and Study Conclusions

The analysis did not show a statistically significant between user-driven parameters, user ratings, and ranking algorithms. The user ratings variable did show a statistically significant relationship, but the user-driven parameters did not. The overall user ratings, regardless of user interests, are where content providers should focus their efforts when designing web content.

The results of this study do not readily suggest any changes should be made to how content providers implement search engine optimization techniques. This study did not find a statistically significant relationship between algorithm rankings and user-driven parameters but did show a statistically significant relationship between user ratings and algorithm ranking. This finding indicates that search ranking algorithms may be more predictable among overall user usage as opposed to predicting rankings based on subsets of users. Content providers can utilize this information to help improve their overall site quality by broadening keywords utilized to focus more towards a larger group of users as opposed to subsets of users.

This would indicate a slight shift in how some content providers think. Many content providers currently focus on keyword sets in an effort to get their websites higher within a result set. While this is important, this study showed that there is a relationship between overall user ratings and the algorithm ranking. Gaining a better understanding of what this relationship is an important next step in order to provide content providers with a complete picture of how they can improve upon their content and still maintain a high ranking among search result pages.

References

Abdel-Hafez, A., Xu, Y., & Jøsang, A. (2015, January). A normal-distribution based rating aggregation method for generating product reputations. *Web Intelligence, 13*(1), 43-51. doi:10.3233/WEB-150306

Abowitz, D., & Toole, M. (2010). Mixed method research: Fundamental issues of design, validity, and reliability in construction research. *Journal of Construction Engineering and Management*, 1-23. doi:10.1061/(ASCE)CO.1943-7862.0000026

Acharya, A. S., Prakash, A., Saxena, P., & Nigam, A. (2013). Sampling: Why and how of it. *Indian Journal of Medical Specialties, 4*(2), 330-333. doi:10.7713/ijms.2013.0032

Aghaiipour-Chafuchahi, H., & Ahmadi-Abkenari, F. (2015). PSEFiL: A personalized search engine with filtered links. *International Journal of Computer Applications, 110*(6), 34-40. doi:10.5120/19323-0961.

Almulla, M., Yahyaoui, H., & Al-Matori, K. (2015). A new fuzzy hybrid technique for ranking real world web services. *Knowledge-Based Systems, 77*, 1-15. doi:10.1016/j.knosys.2014.12.021

Al-Shamri, M. Y. H. (2016). User profiling approaches for demographic recommender systems. *Knowledge-Based Systems*, *100*, 175-187. doi:10.1016/j.knosys.2016.03.006

Aravamudhan, N., & Krishnaveni, R. (2015). Establishing and reporting content validity evidence of training and development capacity building scale (TDCBS).

Management: Journal of Contemporary Management Issues, 20(1), 131-158. Retrieved from https://hrcak.srce.hr/file/208893

Ardern, C., Nie, J., Perez, D., Radhu, N., & Ritvo, P. (2013). Impact of participant incentives and direct and snowball sampling on survey response rate in an ethnically diverse community: Results from a Pilot study of physical activity and the built environment. *Journal of Immigrant and Minority Health, 15*(1), 207-214. doi:10.1007/s10903-011-9525-y

Arrow, K. (2012). *Social choice and individual values.* Mansfield Centre, CT: Martino Publishing.

Artino, A. R. Jr., La Rochelle, J. S., Dezee, K. J., & Gehlbach, H. (2014). Developing questionnaires for educational research: AMEE Guide No. 87. *Medical Teacher, 36*(6), 463-474. doi:10.3109/0142159X.2014.889814

Balakrishnan, V., Ahmadi, K., & Ravana, S. D. (2015). Improving retrieval relevance using users' explicit feedback. *Aslib Journal of Information Management*, *68*(1), 76-98. doi:10.1108/AJIM-07-2015-0106

Banjanovic, E. S., & Osborne, J. W. (2016). Confidence intervals for effect sizes: Applying bootstrap resampling. *Practical Assessment, Research & Evaluation*, *21*(5), 1-20. Retrieved from http://pareonline.net

Barboucha, H., & Nasri, M. (2015). Simulator for calculating Pagerank by javascript. *International Journal of Computer Applications*, *122*(22), 36-40. doi:10.5120/21860-5188

Bateson, M., McPeake, J., & O'Neill, A. (2014). Electronic surveys: How to maximise success. *Nurse Researcher, 21*(3), 24-26. doi:10.7748/nr2014.01.21.3.24.e1205

Becker, T. E., Atinc, G., Breaugh, J. A., Carlson, K. D., Edwards, J. R., & Spector, P. E. (2016). Statistical control in correlational studies: 10 essential recommendations for organizational researchers. *Journal of Organizational Behavior, 37*(2), 157-167. doi:10.1002/job.2053

Benabbou, N., Perny, P., & Viappiani, P. (2017). Incremental elicitation of Choquet capacities for multicriteria choice, ranking and sorting problems. *Artificial Intelligence, 246,* 152-180. doi:10.1016/j.artint.2017.02.001

Ben-Yashar, R., & Nitzan, S. (2017). Is diversity in capabilities desirable when adding decision makers? *Theory and Decision, 82*(3), 395-402. doi:10.1007/s11238-016-9570-z

Bodoff, D., & Ho, S. Y. (2016). Effectiveness of website personalization: does the presence of personalized recommendations cannibalize sampling of other items? *International Journal of Electronic Commerce, 20*(2), 208-235. doi:10.1080/10864415.2016.1087821

Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology, 100*(2), 431. doi:10.1037/a0038047

Bostan, S., & Ghasemzadeh, M. (2014). Personalization of search engines, based-on comparative analysis of user behavior. *Journal of Advances in Computer*

Research, *6*(2), 65-72. Retrieved from http://jacr.iausari.ac.ir/article_12698_1fb22bf044c86bc0b65a2168a57c9416.pdf

Bowden, C., & Galindo-Gonzalez, S. (2015). Interviewing when you're not face-to-face: The use of email interviews in a phenomenological study. *International Journal of Doctoral Studies*, *10*(12), 79-92. Retrieved from http://ijds.org/Volume10/IJDSv10p079-092Bowden0684.pdf

Brin, S., & Page, L. (2012). Reprint of: The anatomy of a large-scale hypertextual web search engine. *Computer Networks, 56*(18), 3825-3833. doi:10.1016/j.comnet.2012.10.007

Bro, R., & Smilde, A. K. (2014). Principal component analysis. *Analytical Methods, 6*, 2812-2831. doi:10.1039/c3ay41907j

Brutus, S., Aguinis, H., & Wassmer, U. (2012). Self-reported limitations and future directions in scholarly reports analysis and recommendations. *Journal of Management*, *39*(1), 48-75. doi:10.1177/0149206312455245

Bryman, A. (2001). The debate about quantitative and qualitative research: a question of method or epistemology? *The British Journal of Sociology*, *35*(1), 75-92. doi:10.2307/590553

Campbell, D. E., Parks, C., & Wells, J. D. (2015). Understanding the role of theory on instrument development: an examination of strengths and weaknesses of discriminant validity analysis techniques. *Journal of Information Technology Theory and Application*, *16*(1), 25. Retrieved from http://aisel.aisnet.org/jitta/vol16/iss1/3

Chen, K., Ji, X., & Wang, H. (2017). A search index-enhanced feature model for news recommendation. *Journal of Information Science, 43*(3), 328-341. doi:10.1177/0165551516639801

Chen, L., Chen, G., & Wang, F. (2015). Recommender systems based on user reviews: The state of the art. *User Modeling and User-Adapted Interaction, 25*(2), 99-154. doi:10.1007/s11257-015-9155-5

Cina, G., & Endriss, U. (2016). Proving classical theorems of social choice theory in modal logic. *Autonomous Agents and Multi-Agent Systems, 30*(5), 963-989. doi:10.1007/s10458-016-9328-6

Coban, C., & Sanver, M. R. (2014). Social choice without the Pareto principle under weak independence. *Social Choice and Welfare, 43*(4), 953-961. doi:10.1007/s00355-014-0812-0

da Silva, E. Q., Camilo-Junior, C., Pascoal, L. M., & Rosa, T. (2016). An evolutionary approach for combining results of recommender systems techniques based on collaborative filtering. *Expert Systems with Applications, 53*, 204-218. doi:10.1016/j.eswa.2015.12.050

Dascalu, M., Trausan-Matu, S., McNamara, D. S., & Dessus, P. (2015). ReaderBench: Automated evaluation of collaboration based on cohesion and dialogism. *International Journal of Computer-Supported Collaborative Learning, 10*(4), 395-423. doi:10.1007/s11412-015-9226-y

Demeester, T., Aly, R., Hiemstra, D., Nguyen, D., & Develder, C. (2016). Predicting relevance based on assessor disagreement: analysis and practical applications for search evaluation. *Information Retrieval Journal*, *19*(3), 284-312. doi:10.1007/s10791-015-9275-x

Dempsey, L., Dowling, M., Larkin, P., & Murphy, K. (2016). Sensitive interviewing in qualitative research. *Research in Nursing & Health*, *39*(6), 480-490. doi:10.1002/nur21743

de Sáa, S., Gil, M. Á., González-Rodríguez, G., López, M. T., & Lubiano, M. A. (2015). Fuzzy rating scale-based questionnaires and their statistical analysis. *IEEE Transactions on Fuzzy Systems*, *23*(1), 111-126. doi:10.1109/TFUZZ.2014.2307895

Dong, X., Li, R., He, H., Gu, X., Sarem, M., Qiu, M., & Li, K. (2017). EDS: An efficient data selection policy for search engine storage architectures. *Future Generation Computer Systems*, 220-231. doi:10.1016/j.future.2016.02.014

Eisend, M. (2015). Have we progressed marketing knowledge. A meta-meta-analysis of effect sizes in marketing research. *Journal of Marketing*, *79*(3), 23-40. doi:10.1509/jm.14.0288

Enayatollahi, I., Bazzazi, A. A., & Asadi, A. (2014). Comparison between neural networks and multiple regression analysis to predict rock fragmentation in open-pit mines. *Rock Mechanics and Rock Engineering*, *47*(2), 799-807. doi:10.1007/s00603-013-0415-6

Faden, R., Kass, N., Goodman, S., Pronovost, P., Tunis, S., & Beauchamp, T. (2013). An ethics framework for a learning health care system: A departure from traditional research ethics and clinical ethics. *Hastings Center Report, 43*(1), 16-27. doi:10.1002/hast.134

Fang, J. (2016). Mining analysis on user search behavior based on hadoop. *International Journal of Database Theory and Application, 9*(3), 81-86. doi:10.14257/ijdta.2016.9.3.09

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analysis using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149-1160. doi:10.3758/brm.41.4.1149

Feng, S., Wong, Y. K., Wong, L. Y., & Hossain, L. (2019). The Internet and facebook usage on academic distraction of college students. *Computers & Education*, *134*, 41-49. doi:10.1016/j.compedu.2019.02.005

Fotso, S., & Fono, L. (2015). Arrow-type results under fuzzy preferences based on filter and ultrafilter. *Fuzzy Sets and Systems*, *266*, 101-113. doi:10.1016/j.fss.2014.05.014

Frels, R., & Onwuegbuzie, A. (2013). Administering qualitative instruments with qualitative interviews: A mixed research approach. *Journal of Counseling & Development*, *91*(2), 184-194. doi:10.1002/j.1556-6676-2013.00085.x

Gautam, A., & Bedi, P. (2017). Developing content-based recommender system using Hadoop MapReduce. *Journal of Intelligent & Fuzzy Systems, 32*(4), 2997-3008. doi:10.3233/JIFS-169243

Ghose, A., Ipeirotis, P., & Li, B. (2014). Examining the impact of ranking on consumer behavior and search engine revenue. *Management Science, 60*(7), 1632-1654. doi:10.1287/mnsc.2013.1828

Gibbard, A. F. (2014). Intransitive social indifference and the Arrow dilemma. *Review of Economic Design*, *18*(1), 3-10. doi:10.1007/s10058-014-0158-1

Giomelakis, D., & Veglis, A. (2016). Investigating search engine optimization factors in media websites: The case of Greece. *Digital Journalism*, *4*(3), 379-400. doi:10.1080/21670811.2015.1046992

Gleich, D. (2015). PageRank beyond the web. *SIAM Review*, *57*(3), 321-363. doi:10.1137/140976649

Goode, S., Lin, C., Tsai, J., & Jiang, J. (2015). Rethinking the role of security in client satisfaction with Software-as-a-Service (SaaS) providers. *Decision Support Systems*, *70*, 73-85. doi:10.1016/j.dss.2014.12.005

Grant, A. (2014). Troubling 'lived experience': A post-structural critique of mental health nursing qualitative research assumptions. *Journal of Psychiatric and Mental Health Nursing*, *21*(6), 544-549. doi:10.1111/jpm.12113

Green, C. A., Duan, N., Gibbons, R. D., Hoagwood, K. E., Palinkas, L. A., & Wisdom, J. P. (2015). Approaches to mixed methods dissemination and implementation research: Methods, strengths, caveats, and opportunities. *Administration and Policy in Mental Health and Mental Health Services Research*, *42*(5), 508-523. doi:10.1007/s10488-014-0552-6

Guo, G., Zhang, J., & Yorke-Smith, N. (2015). Leveraging multiviews of trust and similarity to enhance clustering-based recommender systems. *Knowledge-Based Systems, 74,* 14-27. doi:10.1016/j.knosys.2014.10.016

Guo, H., & Chen, Y. (2016). User interest detecting by text mining technology for microblog platform. *Arabian Journal for Science and Engineering, 41*(8), 3177-3186. doi:10.1007/s13369-016-2137-z

Gupta, D., & Chavhan, N. (2014). SmartPMS: User profiling based re-ranking approach for mobile search engine. *International Journal of Advanced Research in Computer Science and Management Studies, 2*(4), 135-142. Retrieved from http://www.ijarcsms.com/docs/paper/volume2/issue4/V2I4-0054.pdf

Halperin, I., Pyne, D. B., & Martin, D. T. (2015). Threats to internal validity in exercise science: a review of overlooked confounding variables. *International Journal of Sports Physiology and Performance*, *10*(7), 823-829. doi:10.1123/ijspp.2014-0566

Hanisch, D. N., & Rau, S. B. (2014). Application of metric conjoint analysis in family business research. *Journal of Family Business Strategy*, *5*(1), 72-84. doi:10.1016/j.jfbs.2014.01.003

Hopkins, L., & Ferguson, K. E. (2014). Looking forward: The role of multiple regression in family business research. *Journal of Family Business Strategy*, *5*(1), 52-62. doi:10.1016/j.jfbs.2014.01.008

Hoxha, L. (2017). Relationship between project managers' age, years of experience, and project success (Doctoral dissertation, Walden University). Retrieved from https://scholarworks.waldenu.edu/cgi/viewcontent.cgi?article=6575&context=dissertations

Igersheim, H. (2017). The death of welfare economics: History of a controversy. *History of Political Economy*, *2017*(03), 1-37. doi:1-.2139/ssrn.2901574

Ignatov, D., Nikolenko, S., Abaev, T., & Poelmans, J. (2016). Online recommender system for radio station hosting based on information fusion and adaptive tag-aware profiling. *Expert Systems with Applications, 55,* 546-558. doi:10.1016/j.eswa.2016.02.020

Jacobs, J., de Poel, I., & Osseweijer, P. (2014). Clarifying the debate on selection methods for engineering: Arrow's impossibility theorem, design performances, and information basis. *Research in Engineering Design*, *25*(1), 3-10. doi:10.1007/s00163-013-0160-6

Jamshidian, M., Jalal, S. J., & Jansen, C. (2014). MissMech: An R package for testing homoscedasticity, multivariate normality, and missing completely at random (MCAR). *Journal of Statistical Software*, *56*(6), 1-31. doi:10.18637/jss.v056.i06

Jayashree, R., & Christy, A. (2015). Improving the enhanced recommended system using bayesian approximation method and normalized discounted cumulative gain. *Procedia Computer Science*, *50*, 216-222. doi:10.1016/j.procs.2015.04.057

Jerath, K., Ma, L., & Park, Y. H. (2014). Consumer click behavior at a search engine: The role of keyword popularity. *Journal of Marketing Research*, *51*(4), 480-486. doi:10.1509/jmr.13.0099

Jilani, T., Fatima, U., Baig, M. M., & Mahmood, S. (2015). A survey and comparative study of different PageRank algorithms. *International Journal of Computer Applications*, *120*(24), 24-30. doi:10.5120/21410-4444

Johnson, D. (2017). How attitude toward the behavior, subjective norm, and perceived behavioral control affects information security behavior intention (Doctoral dissertation, Walden University). Retrieved from https://pdfs.semanticscholar.org/8fe6/49be6cfd1dad7672a72146d626e2c42ebeae.pdf?_ga=2.98075034.480462793.1575689117-1892205603.1573149922

Judkins-Cohn, T. M., & Kielwasser-Withrow, K. (2014). Ethical principles of informed consent: Exploring nurses' dual role of care provider and researcher. *The Journal of Continuing Education in Nursing, 4545*(11), 35-42. doi:10.3928/00220124-20131223-03

Kapeller, J., Schutz, B., & Steinerberger, S. (2013). The impossibility of rational consumer choice. *Journal of Evolutionary Economics, 23*(1), 39-60. doi:10.1007/s00191-012-0268-2

Khazaal, Y., Van Singer, M., Chatton, A., Achab, S., Zullino, D., Rothen, S., & Thorens, G. (2014). Does self-selection affect samples' representativeness in online surveys? An investigation in online video game research. *Journal of Medical Internet Research*, *16*(7), 1-23. doi:10.2196/jmir.2759

Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R package for assessing multivariate normality. *The R Journal*, *6*(2), 151-162. Retrieved from https://journal.r-project.org/archive/2014/RJ-2014-031/RJ-2014-031.pdf

Kratochwill, T. R., & Levin, J. R. (2014). Meta-and statistical analysis of single-case intervention research data: Quantitative gifts and a wish list. *Journal of School Psychology*, *52*(2), 231-235. doi:10.1016/j.jsp.2014.01.003

Kritzinger, W. T., & Weideman, M. (2017). Parallel search engine optimization and pay-per-click campaigns: A comparison of cost per acquisition. *South African Journal of Information Management*, *19*(1), 1-13. doi:10.4102/sajim.v19i1.820

Kumar, R., & Prakash, J. (2015). Web crawling through shark-search using PageRank. *Procedia Computer Science*, *48,* 210-216. doi:10.1016/j.procs.2015.04.172

Kyvik, S. (2013). The academic researcher role: Enhancing expectations and improved performance. *Higher Education, 65*(4), 525-538. doi:10.1007/s10734-012-9561-0

Laage, T., Loewy, J. W., Menon, S., Miller, E. R., Pulkstenis, E., Kan-Dobrosky, N., & Coffey, C. (2017). Ethical considerations in adaptive design clinical trials. *Therapeutic Innovation & Regulatory Science*, *51*(2), 190-199. doi:10.1177/2168479016667766

Lantos, J., & Spertus, J. (2014). The concept of risk in comparative-effectiveness research. *New England Journal of Medicine, 371*, 2129-2130. doi:10.1056/NEJMhle1413301

LaRose, R., & Tsai, H.-y. S. (2014). Completion rates and non-response error in online surveys: Comparing sweepstakes and pre-paid cash incentives in studies of online behavior. *Computers in Human Behavior, 34*, 110-119. doi:10.1016/j.chb.2014.01.017

Li, G., Zhang, Z., Wang, L., Chen, Q., & Pan, J. (2017). One-Class collaborative filtering based on rating prediction and ranking prediction. *Knowledge-Based Systems, 124*, 46-54. doi:10.1016/j.knosys.2017.02.034

Liu, L., Chen, X., Liu, M., Jia, Y., Zhong, J., Gao, R., & Zhao, Y. (2016). An influence power-based clustering approach with PageRank-like model. *Applied Soft Computing, 40*, 17-32. doi:10.106/j.asoc.2015.10.050

Liu, L., Song, D., Tang, J., Liao, L., Li, X., & Du, J. (2016). FriendBurst: Ranking people who get friends fast in a short time. *Neurocomputing, 210*, 116-129. doi:10.1016/j.neucom.2015.11.124

Lucero, J., Wallerstein, N., Duran, B., Alegria, M., Greene-Moton, E., Israel, B., & Schulz, A. (2016). Development of a mixed methods investigation of process and outcomes of community-based participatory research. *Journal of Mixed Methods Research*, 1-20. doi:10.1177/1558689816633309

Luh, C., & Huang, T. (2016). Estimating Google's search engine ranking function from a search engine optimization perspective. *Online Information Review, 40*(2), 239-255. doi:10.1108/OIR-04-2015-0112

Maddux, R. (2014). Arrow's theorem for incomplete relations. *Journal of Logical and Algebraic Methods in Programming, 83*(2), 235-248. doi:10.1016/j.jlap.2014.02.012

Mahendru, M., Singh, S., & Sharma, G. D. (2014). Emergence of search engine optimization as an advertising tool. *Global Journal of Enterprise Information System*, *6*(1), 19-28. doi:10.18311/gjeis/2014/3104

Maher, J., Markey, J., & Ebery-May, D. (2013). The other half of the story: Effect size analysis in quantitative research. *CBE-Life Sciences Education, 12*(3), 345-351. doi:10.1187/cbe.13-04-0082

Makovi, M. (2016). Interest groups and the impossibility of democratic socialism: Hayek, Jewkes, and the Arrow theorem. *Munich Personal RePEc Archive*. Retrieved from https://mpra.ub.uni-muenchen.de/70173/

Mala, K., & Hema, M. A. (2014). An efficient query processing for personalized web search engine. *International Journal of Advanced Research in Computer Science and Management Studies, 2*(5), 132-139. Retrieved from http://www.ijarcsms.com/docs/paper/volume2/issue5/V2I5-0013.pdf

Maniquet, F., & Mongin, P. (2015). Approval voting and Arrow's impossibility theorem. *Social Choice and Welfare, 44*(3), 519-532. doi:10.1007/s00355-014-0847-2

Marshall, P. A., Adebamowo, C. A., Adeyemo, A. A., Ogundiran, T. O., Strenski, T., Zhou, J., & Rotimi, C. N. (2014). Voluntary participation and comprehension of informed consent in a genetic epidemiological study of breast cancer in Nigeria. *BMC Medical Ethics*, *15*(1), 38. doi:10.1186/1472-6939-15-38

Mavridis, T., & Symeonidis, A. (2015). Identifying valid search engine ranking factors in a web 2.0 and web 3.0 context for building efficient SEO mechanisms. *Engineering Applications of Artificial Intelligence, 41*, 75-91. doi:10.1016/j.engappai.2015.02.002

McComb, C., Cagan, J., & Kotovsky, K. (2017). Optimizing design teams based on problem properties: Computational team simulations and an applied empirical test. *Journal of Mechanical Design, 139*(4), 1-12. doi:10.1115/1.4035793

McComb, C., Goucher-Lambert, K., & Cagan, J. (2017). Impossible by design? Fairness, strategy, and Arrow's impossibility theorem. *Design Science, 3*(2), 1-26. doi:10.1017/dsj.2017.1

McCusker, K., & Gunaydin, S. (2015). Research using qualitative, quantitative or mixed methods and choice based on the research. *Perfusion, 30*, 537-542. doi:10.1177/0267659114559116

McKechnie, S., & Prithwiraj, N. (2016). Effect of new-to-market e-store features on first time browsers. *International Journal of Human-Computer Studies, 90*, 14-26. doi:10.1016/j.ijhcs.2016.03.002

Meusel, R., Ritze, D., & Paulheim, H. (2016). Towards more accurate statistical profiling of deployed schema.org microdata. *ACM Journal of Data and Information Quality, 8*(1), 3. doi:10.1145/2992788

Moradi, P., & Ahmadian, S. (2015). A reliability-based recommendation method to improve trust-aware recommender systems. *Expert Systems with Applications, 42*(21), 7386-7398. doi:10.1016/j.eswa.2015.05.027

Morreau, M. (2015). Theory choice and social choice: Kuhn vindicated. *Mind: A Quarterly Review of Philosophy, 124*(493), 239-262. doi:10.1093/mind/fzu176

Najafabadi, M. K., Mahrin, M. N., Chuprat, S., & Sarkan, H. M. (2017). Improving the accuracy of collaborative filtering recommendations using clustering and association rules mining on implicit data. *Computers in Human Behavior, 67*, 113-128. doi:10.1016/j.chb.2016.11.010

Ninjbat, U. (2015). Impossibility theorems are modified and unified. *Social Choice and Welfare*, *45*(4), 849-866. doi:10.1007/s00355-015-0887-2

Okasha, S. (2015). On Arrow's theorem and scientific rationality: reply to Morreau and Stegenga. *Mind: A Quarterly Review of Philosphy*, *124*(493), 279-294. doi:10.1093/mind/fzu177

Orlikowski, W., & Scott, S. (2015). The Algorithm and the crowd: Considering the materiality. *MIS Quarterly, 39*(1), 201-206. doi:10.25300/MISQ/2015/39.1.09

Orso, V., Ruotsalo, T., Leino, J., Gamberini, L., & Jacucci, G. (2017). Overlaying social information: The effects on users' search and information-selection behavior. *Information Processing & Management*, *53*(6), 1269-1286. doi:10.1016/j.ipm.2017.06.001

Palotti, J., Hanbury, A., Muller, H., & Kahn, C. E. (2016). How users search and what they search for in the medical domain. *Information Retrieval Journal*, *19*(1-2), 189-224. doi:10.1007/s10791-015-9269-8

Parsons, J., & Ralph, P. (2014). Generating effective recommendations using viewing-time weighted preferences for attributes. *Journal of the Association for Information Systems*, *15*(8), 484. Retrieved from https://pdfs.semanticscholar.org/d031/b9f81aae30f8ec272ac563650728d787256d.pdf

Pazzani, M. J. (1999). A framework for collaborative, content-based and demographic filtering. *Artificial Intelligence Review*, *13*(5-6), 393-408. Retrieved from https://link.springer.com/content/pdf/10.1023/A:1006544522159.pdf

Pirouz, M., & Zhan, J. (2016). Optimized relativity search: Node reduction in personalized page rank estimation for large graphs. *Journal of Big Data, 3*(12), 1-17. doi:10.1186/s40537-016-0047-2

Ramesh, N., & Andrews, J. (2015). Personalized search engine using social networking activity. *Indian Journal of Science and Technology, 8*(4), 301-306. doi:10.17485/ijst/2015/v8i4/60376

Raschke, R., Krishen, A., Kachroo, P., & Maheshwari, P. (2013). A combinatorial optimization based sample identification method for group comparisons. *Journal of Business Research, 66*, 1267-1271. doi:10.1016/j.jbusres.2012.02.024

Rice, S., Winter, S. R., Doherty, S., & Milner, M. (2017). Advantages and disadvantages of using internet-based survey methods in aviation-related research. *Journal of Aviation Technology and Engineering*, *7*(1), 58-65. doi:10.7771/2159-6670.1160

Rindfuss, R. R., Choe, M. K., Tsuya, N. O., Bumpass, L. L., & Tamaki, E. (2015). Do low survey response rates bias results? Evidence from Japan. *Demographic Research*, *32*, 797. doi:10.4054/DemRes.2015.32.26

Ritesh, J. (2015). A note on the Arrow's impossibility theorem. *Economic Annals, 60*(207), 39-48. doi:10.2298/EKA1507039J

Roberts, L., & Allen, P. (2015). Exploring ethical issues associated with using online surveys in educational research. *Educational Research and Evaluation, 21*(2), 95-108. doi:10.1080/13803611.2015.1024421

Roosta, M., Ghaedi, M., Daneshfar, A., Sahraei, R., & Asghari, A. (2014). Optimization of the ultrasonic assisted removal of methylene blue by gold nanoparticles loaded on activated carbon using experimental design methodology. *Ultrasonics Sonochemistry*, *21*(1), 242-252. doi:10.1016/j.ultsonch.2013.05.014

Rouf, A. S., Grech, A., & Allman-Farinelli, M. (2017). Assessing the efficacy and external validity of interventions promoting calcium or dairy intake in young adults: A systematic review with meta-analysis. *Critical Reviews in Food Science and Nutrition,* 1-17. doi:10.1080/10408398.2017.1336508

Saari, D. (2016). From Arrow's theorem to 'dark matter'. *British Journal of Political Science, 46*(1), 1-9. doi:10.1017/S000712341500023X

Salam, A., Siraj, H., Hasan, N., Jin, T., Roslan, R., Othman, B., & Nazam, M. (2015). Internet usage and academic performance: A study in a malaysian public university. *International Medical Journal*, *22*(2), 83-86. Retrieved from https://www.researchgate.net/publication/275833912_Internet_Usage_and_Academic_Performance_A_Study_in_a_Malaysian_Public_University

Saxena, N., Agarwal, S., & Katiyar, V. (2016). Personalized web search using user identity. *International Journal of Computer Applications*, *147*(12), 14-17. doi:10.5120/ijca2016911267

Schroeder, R. (2014). Does google shape what we know? *Prometheus, 32*(2), 145-160. doi:10.1080/08109028.2014.984469

Sehgal, S., Chaudry, S., Biswas, P., & Jain, S. (2016). A new genre of recommender systems based on modern paradigms of data filtering. *Procedia Computer Science, 92*, 562-567. doi:10.1016/j.procs.2016.07.383

Shafiq, O., Alhajj, R., & Rokne, J. G. (2015). On personalizing web search using social network analysis. *Information Sciences, 314,* 55-76. doi:10.1016/j.ins.2015/02.029

Simon, S. S. (2013). The essentials of employee engagement in organizations. *Journal of Contemporary Research in Management*, *6*, 386-408. doi:10.5465/AMR.2006.20208687

Singh, A., Sharma, A., & Dey, N. (2015). Semantics and agents oriented web personalization: State of the art. *International Journal of Service Science, Management, Engineering, and Technology (IJSSMET)*, *6*(2), 35-49. doi:10.4018/jssmet.2015040103

Singh, D., & Sethi, A. (2016). An integrated approach to search engine evolution. *International Journal of Engineering Sciences*, *17*, 538-548. Retreived from http://ijoes.vidyapublications.com/paper/Vol17/81-Vol17.pdf

Singh, N., & Varshney, M. (2013). Query recommendation employing query logs in search optimization. *International Journal of Advanced Networking and Applications*, *5*(3), 1917-1921. Retrieved from https://pdfs.semanticscholar.org/238c/c737e4a52e235ca5a54f6b6fd2f6d11d1974.pdf

Sohn, B. K., Thomas, S. P., Greenberg, K. H., & Pollio, H. R. (2017). Hearing the voices of students and teachers: a phenomenological approach to educational research. *Qualitative Research in Education*, *6*(2), 121-148. doi:10.17583/qre.2017.2374

Son, L. H. (2014). HU-FCF: A hybrid user-based fuzzy collaborative filtering method in recommender systems. *Expert Systems with Applications*, *41*(15), 6861-6870. doi:10.1016/j.eswa.2014.05.001

Spector, P., & Meier, L. (2014). Methodologies for the study of organizational behavior processes: How to find your keys in the dark. *Journal of Organizational Behavior, 35*(8), 1109-1119. doi:10.1002/job.1966

Spitzmuller, J., & Warnke, I. H. (2011). Discourse as a “linguistic object”: Methodical and methodological delimitations. *Critical Discourse Studies,* 8, 75-94. doi:10.1080/17405904.2011.558680

St. Clair, T., Cook, T. D., & Hallberg, K. (2014). Examining the internal validity and statistical precision of the comparative interrupted time series design by comparison with a randomized experiment. *American Journal of Evaluation*, *35*(3), 311-327. doi:10.1177/1098214014527337

Stang, J. (2015). Ethic in action: Conducting ethical research involving human subjects: a primer. *Journal of the Academy of Nutrition and Dietetics, 115*(12), 2019-2022. doi:10.1016/j.jand.2015.10.006

Stegenga, J. (2015). Theory choice and social choice: Okasha versus Sen. *Mind*, *124*(493), 263-277. doi:10.1093/mind/fzu

Sugumaran, P., Ravi, K. K., & Shanmugam, T. (2017). A Novel algorithm for enhancing search results by detecting dissimilar patterns based on correlation method. *The International Arab Journal of Information Technology, 14*(1), 60-69. Retrieved from https://www.researchgate.net/publication/309457304_A_Novel_Algorithm_for_Enhancing_Search_Results_by_Detecting_Dissimilar_Patterns_Based_on_Correlation_Method

Suruliandi, A., Rajkumar, T. D., & Selvaperumal, P. (2015). Validating the performance of personalization techniques in search engine. *ICTACT Journal On Soft Computing*, *5*(3), 965-970. doi: 10.21917/ijsc.2015.0135

Thamhain, H. J. (2014). Assessing the effectiveness of quantitative and qualitative methods for R&D project proposal evaluations. *Engineering Management Journal*, *26*(3), 3-12. doi:10.1080/10429247.2014.11432015

Udom, A. (2017). Virtual Team Success: The impact of leadership style and project management experience (Doctoral dissertation, Walden University). Retrieved from https://pdfs.semanticscholar.org/5137/b88070e51603093e294070d200355784d360.pdf?_ga=2.139412790.480462793.1575689117-1892205603.1573149922

U.S. Department of Health and Human Services. (1979). *The Belmont Report*. Retrieved from http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.html#xrespect

Venkatesh, V., Brown, S., & Bala, H. (2013). Bridging the qualitative-quantitative divide: Guidelines for conducting mixed methods research in information systems. *MIS Quarterly, 37*(1), 21-54. doi:10.25300/MISQ/2013/37.1.02

Verma, A., Jha, M., & Mitra, M. (2016). Digital awareness and internet usage by school students. *Journal of Psychosocial Research, 11*(2), 259-270. Retrived from https://search.ebscohost.com/login.aspx?direct=true&db=a9h&AN=121178913&site=eds-live&scope=site

Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, *69*, 29-39. doi:10.1016/j.eswa.2016.09.040

Westerman, M. (2014). Examining arguments against quantitative research: "Case studies" illustrating the challenge of finding a sound philosophical basis for a human sciences approach to psychology. *New Ideas in Psychology, 32*, 42-58. doi:10.1016/j.newideapsych.2013.08.002

Winship, C., & Western, B. (2016). Multicollinearity and model misspecification. *Sociological Science*, *3*(27), 627-649. doi:10.15195/v3.a27

Woodside, A. G. (2013). Moving beyond multiple regression analysis to algorithms: Calling for adoption of a paradigm shift from symmetric to asymmetric thinking in data analysis and crafting theory. *Journal of Business Research, 66*(4), 463-472. doi:10.1016/j.jbusres.2012.12.021

Wu, W., Jia, F., & Enders, C. (2015). A comparison of imputation strategies for ordinal missing data on likert scale variables. *Multivariate Behavioral Research, 50*(5), 484-503. doi:10.1080/00273171.2015.1022644

Yan, X., Zhou, W., Gao, Y., Zhang, Z., Han, J., & Fu, G. (2014). PADM: PageRank-based anomaly detection method of log sequences by graph computing. *IEEE*, 700-703. doi:10.1109/CloudCom.2014.70

Yang, X., Guo, Y., Liu, Y., & Steck, H. (2014). A survey of collaborative filtering based social recommender systems. *Computer Communications, 41,* 1-10. doi:10.1016/j.comcom.2013.06.009

Yeung, C. (2016). Do recommender systems benefit users? A modeling approach. *Journal of Statistical Mechanics Theory and Experiment, 4*. 1-15. doi:10.1088/1742-5468/2016/04/043401

Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Thousand Oaks, CA: SAGE Publications.

Yoo, W., Mayberry, R., Bae, S., Singh, K., He, Q. P., & Lillard J. W. Jr. (2014). A study of effects of multicollinearity in the multivariable analysis. *International Journal of Applied Science and Technology, 4*(5), 9. Retrieved from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4318006/

Yu, H., Jiang, S., & Land, K. C. (2015). Multicollinearity in hierarchical linear models. *Social Science Research, 53,* 118-136. doi:10.1016/j.ssresearch.2015.04.008

Yu, N. N. (2015). A quest for fundamental theorems of social choice. *Social Choice and Welfare, 44*(3), 553-548. doi:10.1007/s00355-014-0846-3

Zhang, H. R., & Min, F. (2016). Three-way recommender systems based on random forests. *Knowledge-Based Systems*, *91*, 275-286. doi:10.1016/j.knosys.2015.06.019

Zhao, S., Zhang, Y., Cheng, B., & Chen, J. L. (2014). A feedback-corrected collaborative filtering for personalized real-world service recommendation. *International Journal of Computers Communications & Control*, *9*(3), 356-369. doi:10.15837/ijccc.2014.3.1085

Zhitomirsky-Geffet, M., Bar-Ilan, J., & Levene, M. (2016). Testing the stability of "wisdom of crowds" judgments of search results over time and their similarity with the search engine rankings. *Aslib Journal of Information Management*, *68*(4), 407-427. doi: 10.1108/AJIM-10-2015-0165

Zhou, M., & Shao, Y. (2014). A powerful test for multivariate normality. *Journal of Applied Statistics*, *41*(2), 351-363. doi:10.1080/02664763.2013.839637

Appendix A: Questionnaire

Demographic Section

The following questions are used for demographic purposes only all responses will be kept completely confidential.

1. Are you a Faculty member or a student?
    a. Faculty
    b. Student
2. What is your age?
3. Which of the following search engines do you prefer to use when searching the web?
    a. Google
    b. Yahoo
    c. Other (if other please specify)
4. Do you utilize the internet for social media?
    a. Everyday
    b. A few times a week
    c. A few times month
    d. Never
5. Do you utilize the internet for research?
    a. Everyday
    b. A few times a week
    c. A few times a month
    d. Never
6. Do you utilize the internet for news?
    a. Everyday
    b. A few times a week
    c. A few times a month
    d. Never
7. Do you utilize the internet for games?
    a. Everyday
    b. A few times a week
    c. A few times a month
    d. Never
8. Do you utilize the internet for media, such as videos or music?
    a. Everyday
    b. A few times a week
    c. A few times a month
    d. Never
9. Of the following, which would you say is your primary usage of the internet?
    a. Social Media
    b. Research

    c. News
    d. Games
    e. Media, such as videos or music
10. What browser are you currently using to fill out this survey?
    a. Chrome
    b. Firefox
    c. Internet Explorer
    d. Safari
    e. Other

The following sites were found using the search term given. Please rate each site using a 1 to 5 scale where 1 is strongly dislike and 5 is strongly like. For your answer please consider the term provided when ranking.

(For the actual survey the following questions will be randomized so the sites are not presented in any particular order)

1. **Search Term:** Computer Programming
   Site: https://en.wikipedia.org/wiki/Computer_programming
2. **Search Term:** Computer Programming
   Site: http://guyhaas.com/bfoit/itp/Programming.html
3. **Search Term:** Computer Programming
   Site: http://www.programmingbasics.org/en/beginner/gettingstarted.html
4. **Search Term:** Computer Programming
   Site: https://en.wikiversity.org/wiki/Introduction_to_Programming/About_Programming
5. **Search Term:** Computer Programming
   Site: https://www.quora.com/topic/Computer-Programming
6. **Search Term:** Computer Programming
   Site: http://www.sciencedirect.com/science/journal/01676423?sdc=1
7. **Search Term:** Computer Programming
   Site: http://www.computerscience.org/careers/computer-programmer/
8. **Search Term:** Recipes
   Site: http://allrecipes.com/
9. **Search Term:** Recipes
   Site: http://www.recipe.com/
10. **Search Term:** Recipes
    Site: http://www.food.com/recipe
11. **Search Term:** Recipes
    Site: http://www.seriouseats.com/recipes
12. **Search Term:** Recipes
    Site: http://www.health.com/recipes
13. **Search Term:** Recipes
    Site: https://minimalistbaker.com/recipes/

14. **Search Term:** Recipes
    Site: http://www.foodandwine.com/recipes

The next set of questions place the websites in the order you feel would be most appropriate for a search engine to return the sites for the given search term.

1. **Search Term:** Computer Programming
    a. **Site:** https://en.wikipedia.org/wiki/Computer_programming
    b. **Site:** http://guyhaas.com/bfoit/itp/Programming.html
    c. **Site:** http://www.programmingbasics.org/en/beginner/gettingstarted.html
    d. **Site:** https://en.wikiversity.org/wiki/Introduction_to_Programming/About_Programming
    e. **Site:** https://www.quora.com/topic/Computer-Programming
    f. **Site:** http://www.sciencedirect.com/science/journal/01676423?sdc=1
    g. **Site:** http://www.computerscience.org/careers/computer-programmer/
2. **Search Term:** Recipes
    a. **Site:** http://allrecipes.com/
    b. **Site:** http://www.recipe.com/
    c. **Site:** http://www.food.com/recipe
    d. **Site:** http://www.seriouseats.com/recipes
    e. **Site:** http://www.health.com/recipes
    f. **Site:** https://minimalistbaker.com/recipes/
    g. **Site:** http://www.foodandwine.com/recipes

Appendix B: Invitation Email

Hello, My name is Gary Michael Taylor, a doctoral student at Walden University. I graduated from Marshall University with a Bachelor of Science in Computer Information Technology in 2006 and then obtained my Masters of Science in Software Engineering from West Virginia University in 2010. I have worked in higher education since I graduated in 2006 at various organizations including Marshall University, Mountwest Community & Technical College, and most recently INTO Marshall University.

I would like to invite you to participate in a research study to help determine how user's rankings of websites differ from existing search ranking algorithms. The findings of this study can potentially help designers to consider new techniques for search ranking algorithms which could lead to improvements in the quality of search results. By improving the quality of search results you could potentially find information you are looking for faster.

Thanks in advance for your participation,

Gary Michael Taylor