MEASURING INFORMATION APPETITE:

THE DESIRE TO SPEND TIME WITH INFORMATION, ENGAGING IN CONSUMPTION, DISSEMINATION, AND CREATION OF INFORMATION FOR PERSONAL REASONS

By

SOO-YEON HWANG

A dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

In partial fulfillment of the requirements

For the degree of

Doctor of Philosophy

Graduate Program in Communication, Information and Library Studies

Written under the direction of

Daniel O. O'Connor

And approved by

____

____

____

____

New Brunswick, New Jersey

May 2016

ABSTRACT OF THE DISSERTATION

Measuring Information Appetite:

The Desire to Spend Time with Information, Engaging in Consumption, Dissemination, and Creation of Information for Personal Reasons

by SOO-YEON HWANG

Dissertation Director:

Daniel O. O'Connor

This dissertation stemmed from the question about when individuals are satisfied with information. It attempts to answer the question from the perspective of information appetite, i.e. the desire to spend time with information, expressed in the amount of time spent engaging in personal information practices (PIPs: consumption, dissemination, and creation of information for personal reasons).

Studied here are the American Time Use Survey (ATUS) dataset collected by the U.S. Bureau of Labor Statistics and a survey dataset collected on Amazon Mechanical Turk (MTurk). This dissertation presents analyses and a discussion of findings of the ATUS dataset, a description of the method of collecting the MTurk dataset, and analyses and a discussion of findings of the MTurk dataset.

Two propositions about information appetite (IA) were made, and both were supported by the datasets. The first proposition was that individuals have varying degrees of IA, as measured by the amount of time spent engaging in PIPs, which was supported by the ATUS and the MTurk data. The second proposition was that an individual has different IAs for different topical areas, which was supported by the MTurk data.

The study identified a small group of people who spent much time for PIPs, namely a high IA group. In addition to PIPs, the high IA group spent more time with their interested topics than the regular group did. As for what affects spending time for PIPs, having free time (e.g. not working/not having a job, or evenings and weekends) was found to be an important factor.

The dissertation ends with a discussion of implications, limitations and suggestions for future studies.

Acknowledgement

It has been a long journey to complete the doctoral degree, and it would not have been possible without the help and support from a number of people. First and foremost, the biggest gratitude goes to my advisor Dr. Daniel O'Connor. Without his stalwart support and guidance, I would not have been able to make the journey in whole. Words cannot describe how indebted I am to him. Next I would like to thank Dr. Claire McInerney for her kind support and encouragement over the years. I am also grateful to my two other committee members, Dr. Hartmut Mokros and Dr. Soo Young Rieh, for their helpful feedback throughout the dissertation process.

Special thanks to my parents for their unending support and belief in my success. Finally, I thank God for his blessings and guidance during hard times, and thank everyone who prayed for me.

Table of Contents

List of Tables

List of Figures

Chapter 1 Introduction

Information seekers often exhibit different reactions to a search that, for example, results in twenty pages. Some might be happy with whatever is on the very top of the first page, some might check the first few pages, and some might end up reading through the entire twenty pages of results. One question asked of this is: when are they satisfied with the information they have and would stop searching?

One way this question was tackled in the past is studying relevance in information retrieval. The idea is that people seek information that maximizes the relevance via an evaluation of the match between the search query and the result. This could be viewed as a rational actor model of human behavior: it assumes that people are rational and make optimal choices based on their objectives in order to maximize utility – which is relevance in this case. So, according to this model, users will continue to search until the most relevant result is found, or if there is time constraint, will settle with the result that is the most relevant enough (Zach, 2005).

Another approach is information overload in that people would stop information seeking when they reach a state of information overload at some point (Savolainen, 2006, pp. 116–118). Information overload is an individual's state when there is excessive communication and information input (Rogers, 1986, p. 181). One popular method to determine the state of information overload is by the amount of information and the available time to process: It is a simple rule-based model in that if an individual cannot process all the required information in the given time period, that person is in a state of information overload (Eppler & Mengis, 2004, pp. 326–328). This model may not explain the whole picture, though. There are more causes to information overload, such as personal attributes (e.g. skills, experience, and motivation), the nature of information itself (e.g. diversity, ambiguity, novelty, complexity, etc.), the type of tasks to be completed, the design of the organization one works in (if the information seeking is work-related), and the information technology used (Eppler & Mengis, 2004, pp. 335–336).

In another view of information overload, when an individual feels like being overloaded with information, a selective consumption of information begins: what someone “needs to know” takes priority over what is “nice to know” (Paisley, 1993, p. 232). From the perspective of a behavioral model, however, the boundary between “need to know” and “nice to know” starts to blur. A behavioral model recognizes that actions such as decision-making are often highly influenced by irrational and contextual factors such as emotion (Naqvi, Shiv, & Bechara, 2006). For example, addiction portrays an individual’s inability to admit and/or stick to optimal choices. In terms of information seeking, this explains why some people simply cannot stop consuming information (e.g. reading), spending more time and energy than they can afford.

In light of the above and the advent of mobile information communication technologies, the question of when individuals are satisfied enough with information takes a new twist. We now live in an age of almost infinite amount of information at our fingertips thanks to portable, ubiquitous, and networked technologies. Therefore, when there is no limit to the amount of information one can consume, studying when users feel content with information, their “information appetite” in other words, offers a new perspective for understanding information behavior.

Appetite is defined as “the desire to eat” according to the Encyclopedia Britannica (“Appetite (diet),” 2015). Britannica also says the following about appetite:

> Appetite is often associated with the desire to eat particular foods based on their smell, flavour, appearance, and appeal; this is a primary factor separating appetite from the primary motive of hunger. In addition, a person may be totally filled with food from a meal and still have an “appetite” for dessert. Furthermore, appetite may be increased or diminished depending on pleasant or unpleasant experiences associated with certain foods.

From the quote above, it is easy to find the parallel between food and information in terms of appetite. Just as appetite is different from hunger, information appetite is not the same as hunger for information: it is triggered more by pleasurable appeal (e.g. personal hobby) than by required needs (e.g. work assignment). In addition, one can argue that those with voracious information appetite on a certain topic would keep consuming information on the topic, sometimes spending more time than they can afford, unlike those with lower appetite levels. Moreover, an individual’s information appetite on a particular topic may increase or decrease over time for various reasons.

Studying information appetite offers several new academic and commercial research opportunities. First, it may explain the individual differences in emotional response to information overload in that arrival of new information causes anxiety in some, but it fills the need for stimulus with others (Bawden & Robinson, 2009, p. 185; Case, 2012, p. 119). Also, it would be interesting to see if there is any connection between information appetite and productivity: would those with high information appetite make more productive workers, or would they be viewed more as procrastinators? In terms of information gatekeepers such as those who write stock newsletters or bloggers offering product recommendations, would their information appetite affect what they choose to write about? One may also consider the effect of one’s information appetite on selective exposure (i.e. the tendency to favor information that reinforces pre-existing viewpoints and avoid information that contradicts them) to political campaigns, for instance. For online advertising agencies and ad-supported services, finding out whether information appetite affects a user's retention rate – the amount of time spent on a website – would yield valuable marketing data.

Therefore, information appetite lends a new insight into the study of information behavior. Some scholars posit that an individual has an implicit knowledge of his or her cognitive capacity of information load (Case, 2012, pp. 97–98). In other words, just as people start eating when they are hungry and stop when they are full, they would do the same with information. However, it is also well known that when it comes to eating food, not everyone behaves rationally every time. Decision making is often influenced by irrational and contextual factors such as emotion (Naqvi et al., 2006) as shown in emotional eating – eating for emotional reasons, not for physical hunger. Likewise, information appetite may explain why some individuals tend to spend more time and energy to seek out information than others do.

Furthermore, for those with higher information appetite, the amount of available information (as discussed in determining the state of information overload) would be irrelevant because the quantity of information to be consumed is virtually limitless in today's information environment. Such individuals would be able to keep consuming and consuming information until time runs out. What is important here is that although a quantity limit is subjective, time limit is not: time is the only finite currency of one's life because everyone is allotted the same 24 hours a day to spend (Kimmel, 2008). In this context, if someone spends his or her time devouring information instead of doing something else, it means that information is more important to that person than anything else at that time. Hence, time is a good measure to compare the voracity of one individual's information appetite to the next. Just as Encyclopedia Britannica defines appetite in diet as the desire to eat, I would define information appetite as "the desire to spend time with information." That is, the more time an individual spends seeking information, the higher the information appetite the individual has.

In this study, I will focus on the following propositions about information appetite:

Proposition 1. Individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices (PIPs). PIPs encompass consumption, dissemination, and creation of information for personal reasons.

Proposition 2. An individual has different information appetites for different topical areas.

The figure below shows a preliminary model.

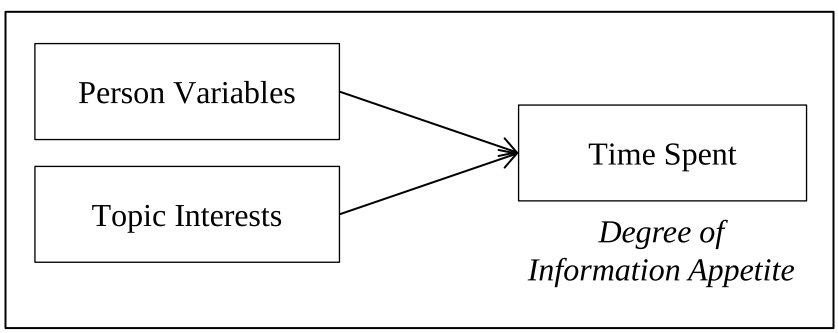


Figure 1.1 Preliminary model

Chapter 2 Review of Relevant Works

2.1 Information Overload

Although having become a popular term with the digital revolution, the concept of information overload has been around since the days of books: Diderot wrote in 1755 that "the number of books will grow continually, and one can predict that a time will come when it will be almost as difficult to learn anything from books as from the direct study of the whole universe" (Diderot, 1964, p. 299). As Diderot anticipated, the amount of information available in the world kept growing, but with the advent of the Internet and the world wide web, it has exploded exponentially.

Information overload is a state one falls into when there is excessive communication and information input to be dealt with. To prevent being overloaded, one makes various adjustments in dealing with the amount of information at hand (Jacoby, 1984; J. G. Miller, 1960). In addition, too frequent communications over digital communication channels (i.e. "chronic interruptions" (J. Robinson, 2014)) may also cause information overload (e.g. "tweet overload" (Sasaki, Kawai, & Kitamura, 2015)).

Rogers (1986) said that information overload "is the state of an individual or system in which excessive communication inputs cannot be processed, leading to breakdown" (p. 181). Eppler and Mengis' literature review revealed more varied definitions of information overload, as shown in table 2.1.

Definitions	Components/dimensions	References
The decision maker is considered to have experienced information overload at the point where the amount of information actually integrated into the decision begins to decline. Beyond this point, the individual's decisions	• Inverted U-curve: relationship between amount of information provided and amount of	Chewning and Harrell (1990), Cook (1993), Griffeth et al. (1988), Schroder et al. (1967), Swain and Haka (2000)

reflect a lesser utilization of the available information.	information integrated by decision maker • Information utilization	
Information overload occurs when the volume of the information supply exceeds the limited human information processing capacity. Dysfunctional effects such as stress and confusion are the result.	• Volume of information supply (information items versus - chunks) • Information processing capacity • Dysfunctional consequences	Jacoby et al. (1974), Malhotra (1982), Meyer (1998)
Information overload occurs when the information-processing requirements (information needed to complete a task) exceed the information-processing capacity (the quantity of information one can integrate into the decision-making process).	• Information-processing capacity • Information-processing requirements	Galbraith (1974), Tushman and Nadler (1978)
Information overload occurs when the information-processing demands on time to perform interactions and internal calculations exceed the supply or capacity of time available for such processing.	• Time demands of information processing; available time versus invested time • Number of interactions (with subordinates, colleagues, superiors) • Internal calculations (i.e., thinking time)	Schick, et al. (1990), Tuttle and Burton (1999)
Information overload occurs when the information-processing requirements exceed the information-processing capacity. Not only is the amount of information (quantitative aspect) that has to be integrated crucial but also the characteristics (qualitative aspect) of information.	• Information-processing requirements • Information-processing capacity • Quantitative and qualitative dimensions of information (multidimensional approach)	Keller and Staelin (1987), Schneider (1987), Owen (1992), Iselin (1993)
Information overload occurs when the decision maker estimates he or she has to handle more information than he or	• Subjective component: opinion, job and	Abdel-Khalik (1973), Iselin (1993), O’Reilly

she can efficiently use.	communication satisfaction • Situational factors and personal factors	(1980), Haksever and Fisher (1996)
Amount of reading matter ingested exceeds amount of energy available for digestion; the surplus accumulates and is converted by stress and overstimulation into the unhealthy state known as information overload anxiety.	• Subjective cause component: energy • Symptom: stress, overstimulation • Subjective effect: information overload • anxiety	Wurman (1990), Wurman (2001), Shenk (1997)

Table 2.1 Definitions of information overload (Eppler & Mengis, 2004, p. 328)

In the definitions above, the information-processing capacity of an individual is featured in multiple definitions. Eppler and Mengis (2004) explained that it is “the quantity of information one can integrate into the decision-making process within a specific time period” (p. 326). What is important here is the time constraint. They suggested a formula of information overload as “information processing requirements > information processing capacities” (p. 326), measured in terms of how much time is available to process information and the amount of information that could be processed in the available period of time. In this view, if all the information could not be processed in the available processing time, information overload occurs.

Time constraint, therefore, is an important factor of information overload. Savolainen (2007) also acknowledged it and defined information overload as “a subjective experience of the insufficiency of time needed to make effective use of information resources available in specific situations” (p. 612). His work was a study of information overload from an information science perspective. His interviews of participants revealed two major strategies to cope with information overload: filtering and withdrawal. The filtering strategy is filtering out useless information from the sources, focusing on the content to avoid; and the withdrawal strategy is minimizing the daily information sources, focusing on the information supply to avoid in order not to be overwhelmed with information. He found that people use a mixture of both strategies to cope with information overload.

Reactions to information overload may be varied. When a new piece of information arrives in the inboxes of individuals with information overload, some may feel anxiety whereas some others may feel stimulated. Anxiety stems from the worry that the volume of information to process is too great, and in the midst of selecting which one to process and which one to pass, one might miss an important piece of information (Edmunds & Morris, 2000, p. 19). Savolainen's study (2007) also acknowledged this phenomenon, but he argued that such feelings tended to be brief without any lasting harm, and often justified by the lack of time.

As for stimulation, the argument is that when new information fails to appear it could cause boredom (Case, 2012, p. 119). This would be especially true for those with high stimulation threshold such as sensation seekers (Zuckerman, 2009).

However, there is a disagreement as to whether information overload exists in reality. Jacoby (1984) argued that individuals can be overloaded in theory, but may not be overloaded in practice because they use various strategies to prevent themselves from being overloaded in the first place. He pointed out that however much information is available, individuals selectively choose the amount and kinds of information to use, though sometimes resulting in missing key information due to such a compromise. The two strategies revealed in Savolainen's study (2007) coincide with his argument.

2.2 Desire

Dholakia’s (2014) review of desire in consumer behavior research revealed three distinct concepts of desire. First is that desire motivates decision making and goal-directed efforts. In this viewpoint, desire operates in two ways: as a goal desire as in the goal to be pursued, and as an implementation desire as in the specific means to reach the goal. Second is that desire is a counter-force that opposes self-control. If a person follows desire, in this viewpoint, it may result in short-term pleasure but long-term negative consequences. Therefore, when desire is experienced, self-control kicks in to prevent a negative outcome. Third is that desire is a sudden urge or craving to buy merchandise impulsively, immediately produced upon encountering a temptation. If left unchecked, this could lead to compulsive buying or hoarding.

The model of goal-directed behaviour represents the first view. In this model, desire is “a state of mind whereby an agent has a personal motivation to perform an action or to achieve a goal” (Perugini & Bagozzi, 2004, p. 71) and provides a direct pathway to intentions and behaviors (Perugini & Bagozzi, 2001). Perugini and Bagozzi (2004) differentiate desire and intention as “desiring a goal and intending to achieve it” (p. 69). That is, unlike desire which is more abstract and less specific than intention, intention leads to action. Also, compared to intention, people consider desire in a longer term and think less of its feasibility. As shown in the figure below, desire is an important predictor of intention which, in turn, is an important predictor of behavior. As there are goal intentions and implementation intentions (Gollwitzer, 1999), there are goal desires and implementation desires that are precursors to the intentions respectively (Dholakia, 2014).

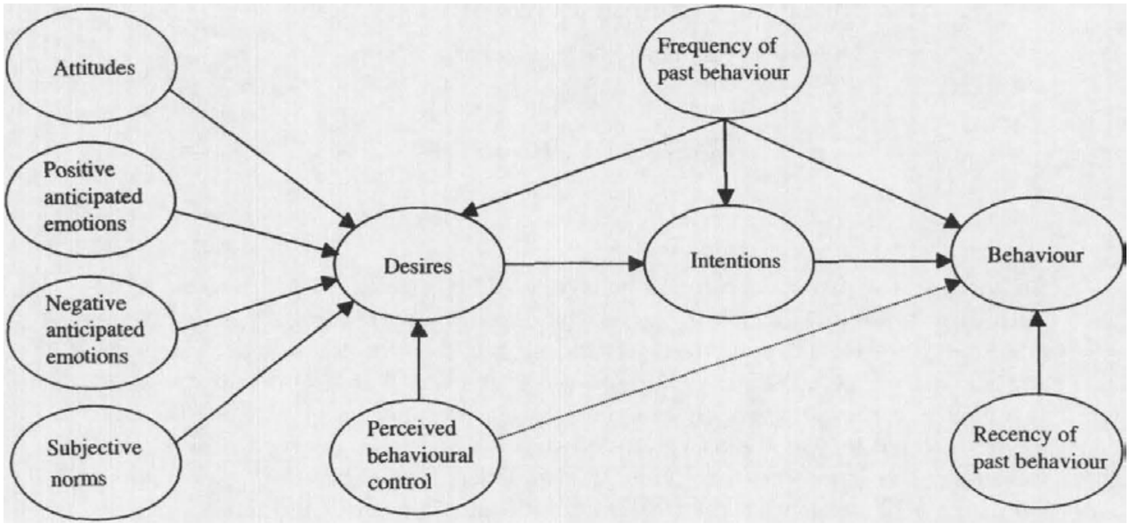


Figure 2.1 The model of goal-directed behaviour (Perugini & Bagozzi, 2001, p. 80)

In the second view of desire, the focus is on self-control, which is a constant struggle between willpower and desire. Desire is something that needs to be resisted, otherwise negative consequences would occur. Whether all desires are problematic, however, is questionable. It is those desires that conflict with other goals that tend to cause resistance (Hofmann, Baumeister, Förster, & Vohs, 2012).

The third view of desire casts it as a sudden buying impulse to acquire more. Consumer research has focused on how to encourage buying impulse so that it can lead to more profit, and then shifted to regarding the impulse as a psychological trait of consumers (Dholakia, 2014).

A buying impulse has the following properties (Dholakia, 2014, p. 16):

1. It is experienced immediately when the consumer is exposed to the desirable object.
2. It is intense and short-lived, manifesting as a burst of motivation, and impelling the consumer to buy within a very short time period, usually a few moments.
3. It may produce a state of psychological disequilibrium marked by emotional conflict.
4. It discourages inaction or the status quo.
5. It may occur even in the absence of a deficit state such as physiological deprivation or withdrawal.

This third viewpoint will be considered in this study. Compulsive acquisition syndrome will be discussed later in this chapter under personality traits relevant to information appetite.

2.3 Personal Information Practices

Time is an interesting construct because it is a limited resource with absolutely no discrimination: everyone has the same 24 hours a day to spend regardless of one's gender, age, wealth, or social status (Kimmel, 2008). Therefore, spending time to seek information rather than doing something else means that information seeking is more appealing to the person than anything else. To understand time consumption for personal information practices (PIPs) in everyday life, data from the American Time Use Survey (ATUS) will be used, along with additionally collected data which will be discussed later.

ATUS is an ongoing time diary study conducted by the U.S. Bureau of Labor Statistics (BLS) since 2003. Samples of the ATUS are randomly drawn from those who participated in the Current Population Survey, which is jointly sponsored by the BLS and the U.S. Census Bureau, and are adjusted to represent geographically the population of the United States (U.S. Bureau of Labor Statistics, 2015b). Each record in the dataset corresponds to an individual reporting how they spent their time in a 24 hour period. In the years 2003 – 2014, there were a total of 159,937 participants, averaging about 13,328 participants per year.

Among over 470 coded activities in the ATUS dataset, eight activities were identified as relevant to PIPs: taking class for personal interest, research or homework for class (for personal interest), researching purchases, television and movies (not religious), listening to the radio, computer use for leisure (excluding games), reading for personal interest, and writing for personal interest. Examples of these eight coded activities are in Appendix A (Hofferth, Flood, & Sobek, 2013; U.S. Bureau of Labor Statistics, 2015a).

TV and radio are included because according to St. Jean and others' (2012) study of online information activities, multimedia use was found to be one of the five major groups of information behaviors. Another major group was content creation. Therefore, the following list of PIPs was compiled for the purpose of this study. These activities cover all gamut of PIP: consumption, dissemination, and creation of information for personal reasons.

- Taking/studying for a class on a subject of personal interest
- Researching purchases
- Computer use for personal interest (excluding games)
- Reading/doing research for personal interest
- Writing/blogging for personal interest
- Listening/watching radio, podcast, TV, or videos for personal interest
- Creating multimedia contents for personal interest

One thing to note is that computer use might seem redundant because all the other activities can be done on a computer. However, they can be done offline, non-electronically as well. Therefore, the question will be asked with an instruction for the participants not to treat each activity as mutually exclusive: that is, if someone spent 30 minutes to do personal research on computer, the 30 minutes would count towards both Computer use and Reading/doing research. This non-mutual exclusivity is another reason why a traditional survey will be conducted instead of time diary to collect time use data because duplicate entries of activities are harder to collect and manage in time diaries.

When the ATUS data was analyzed, it was found that the participation rate, i.e. the percentage of those who spent time with information for a personal reason, was extremely low except for TV and movies, as shown in table 2.2.

PIP Activity	Participation rate (%)[1]	Time spent by person (minutes)[2]				
		n	Min.	Max.	Average	Std. Dev.
Taking class for personal interest	0.7%	1111	5	760	115.38	91.009
Research or homework for class (for personal interest)	0.2%	380	1	840	109.53	107.099
Researching purchases	0.1%	197	2	305	61.47	54.548
Television and movies (not religious)	79.9%	127733	1	1430	212.30	168.599
Listening to the radio	1.7%	2783	1	1055	91.17	109.516
Computer use for leisure (excluding games)	10.3%	16464	1	1080	92.81	93.496
Reading for personal interest	25.6%	40865	1	1263	90.37	90.591
Writing for personal interest	0.3%	422	1	1061	103.72	129.788
All eight activities combined	86.6%	138439	1	1439	237.06	180.634

Table 2.2 Descriptive statistics of the eight PIP activities in the ATUS dataset

One possible scenario is that the ATUS might have failed to capture what people actually do in idle times such as waiting in line or going to a place on public transit because those idle times were recorded as individual events. With mobile devices being so ubiquitous, idle times are rarely spent idle anymore; therefore, perhaps those who cannot afford to carve out a separate time for PIPs may be satisfying their needs using mobile phones or tablets during the idle times. This is another reason why the participants will be asked not to treat the time use activities as mutually exclusive.

Table 2.2 also shows that 86.6% of the sample population engaged in at least one of the eight PIP activities in a day. Among these, 58% were involved in one activity, 25.1%

[1] The percentage of population that spent time on the activity. Participation rate percentage is calculated as Doers / All persons * 100 (Harvey, 1999, pp. 32–36).

[2] *n* = Those who participated in the activity (Total sample size: 159,937)

in two activities, 3.4% in three activities, and 0.1% in four or five activities. This means that individuals have a definite preference when it comes to the type of PIP activity they engage in, and that the appetite in one might not translate into the appetite in another. It would be interesting to see if this phenomenon is replicated in additional data collection for this study.

HP1. Information appetite in one type of activity does not predict information appetite in other types of activities.

The influence of personal characteristics on the time spent will be explored. These independent variables include age, gender, and education level.

HP2. Individual characteristics such as age, gender, and education level affect the time spent for PIPs.

Also, the time spent for PIPs is hypothesized to depend upon how much free time an individual has. Work, domestic and family responsibilities take away what time one can and/or is willing to spend on PIPs.

HP3. Time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities.

2.4 Time as a Matter of Context

Savolainen (2006) described three roles of time in the information studies literature. First is time as a contextual attribute, integrated in the social and cultural situation or context of information seeking. Second is time as a qualifier of access to information: for example, time requirements (e.g. how soon information is needed) or time limit (e.g. time pressure hindering information seeking). Third is time as a qualifier of the information seeking process, as in information seeking behavior models that imply temporal factors.

Of these three, perhaps the most studied is the second aspect of time: time requirements and pressure affecting information seeking behavior (Allen, 2011; Chen & Rieh, 2009; Crescenzi, Capra, & Arguello, 2013; Liu, Yang, Zhao, Jiang, & Zhang, 2014; Slone, 2007). In such studies, time is seen as a frozen element, with no past or future. However, time is constantly moving and evolving, from past to present to future (Savolainen, 2006), and when considering the gamut of personal information practices – creating, consuming, and sharing information – over one's lifetime, this moving aspect of time must be considered.

Sonnenwald and Iivonen (1999) acknowledged the moving aspect of time in their human information behavior framework, and categorized time periods in three ways: a short episode, a longer interval, and an eon which is a long continuous period of time. Hartel (2010) shared a similar insight in her study of gourmet cooking hobbyists. In her framework, there are three temporal arcs in a hobby: career (years/lifetime), subject (weeks/months), and episode (a short episode). She summarizes that:

> There is the long-running career arc that lasts for years or decades and represents the cook's lifetime experience of the hobby. The subject arc is made up of shorter periods (weeks or months) in which a topic organizes culinary activity. Then, the episode arc entails distinct hands-on cooking projects that happen closer to real-time.

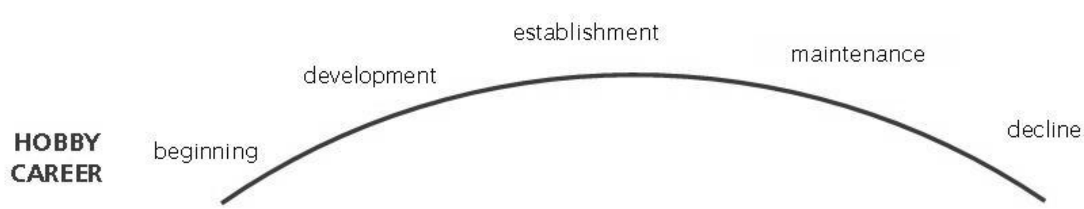


Figure 2.2 The hobby career arc (Hartel, 2010)

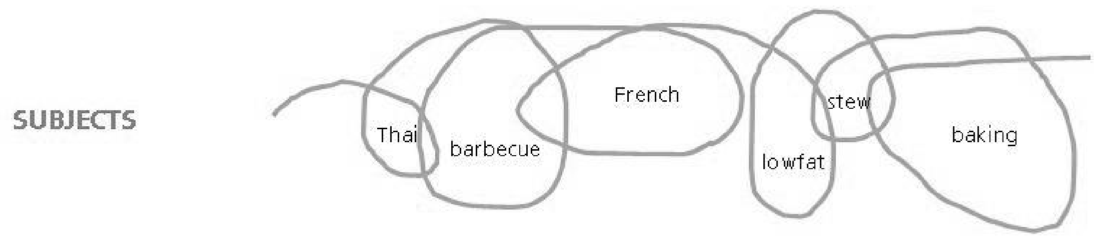


Figure 2.3. The subject arc (Hartel, 2010)

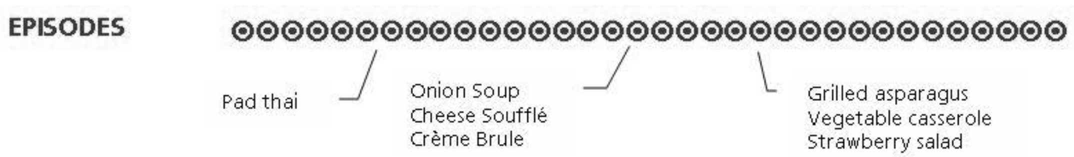


Figure 2.4. The episode arc (Hartel, 2010)

She argued that her framework reflects human information experience, as did Sonnenwald and Iivonen (1999). This viewpoint merits attention. We experience some information episodes that are short bursts (e.g. a quick search on a topic), some that last months (e.g. planning a vacation), and some that last decades (e.g. a deeply-interested topic such as a hobby). Along with information practices, it is assumed that information appetite in a topic ebbs and flows on the same arcs as well.

To account for this moving aspect of time, participants will be asked for how long they have been interested in the topics on which they had spent time with information. In addition to measuring information appetite using the PIP activities, information appetite (i.e. time spent) in multiple topics will be measured. The variables to investigate are the topical information appetite, the topic's temporal context and their relationships to the activity-based information appetite.

RQ1. What are the relationships between the activity-based information appetite and the topical information appetite along with the topic's temporal context?

2.5 Personality Markers

Many of information behavior studies focus on the situation variable. For example, the more studied topic of the time context as discussed before, time requirements and pressure affecting information seeking behavior, is looking at time as a situation variable. Another important focus in this study, however, is the person variable. Therefore, some person variables related to information appetite need to be considered.

2.5.1 Intellectual curiosity

The question of when people stop searching can be rephrased as the question of when people continue to search or not. The drive to do something, or in this case not to do something, is referred to as motivation in psychology. It is defined as the "choice, effort, and persistence to engage in a particular activity or task" ("Motivation: Research starters topic," 2015).

Motivation in information seeking behavior studies have been mostly studied as information needs. Case (2012) explained that needs are "inner motivational state that brings about thought and action" (p. 78). Information need is one of the causes of information seeking, and is often defined as such. A reason why it does not have a clearer definition in the field is because it is something inside one's head and what the researchers can observe and study is information seeking behaviors, from which the needs must be inferred (Case, 2012, p. 87).

Hence, it is imperative to consider the psychological definition of motivation. There are two types of motivation: intrinsic motivation which is doing something because of inherent interest or enjoyment, and extrinsic motivation which is doing something because of a certain separable outcome (Ryan & Deci, 2000). Information appetite works as an inherent drive to seek out more information. To study this intrinsic motivation for continued information seeking, intellectual curiosity of the research participants will be investigated.

Intellectual curiosity, or epistemic curiosity, is viewed as a personality trait that represents the curiosity (or desire) for knowledge (Berlyne, 1954). It is considered one of the predictors of academic performances (von Stumm, Hell, & Chamorro-Premuzic, 2011), and is also counted as one of the desirable personality traits in job seekers (Casserly, 2012; Shanker, 2015).

One of the recent measurements of intellectual curiosity is called the epistemic curiosity (EC) scale (J. A. Litman, 2008; J. A. Litman & Spielberger, 2003). This scale is particularly relevant to information behavior due to the way Litman defined two different aspects of it (2008, 2010): I-type EC (I from interest) and D-type EC (D from deprivation). The I-type EC is about positive feelings of interest: learning something new is for pleasure, not necessarily for an obligation, and is a rewarding act. In information behavior, this is akin to browsing. On the other hand, the D-type EC is about lacking information, being uncomfortable about such a deprived state, and gathering new information in order to escape the negative state. This is analogous to information seeking in response to a task/problem or a recognized gap/uncertainty in one's state of knowledge (Belkin, 2005).

Although the D-type EC sounds like extrinsic motivation, the scale itself (J. Litman, n.d.) shows that the D-type EC is still intrinsic in nature. In other words, it deals with intrinsic values that will come into play when one is faced with an extrinsic motivating factor. Hence, for information appetite, it is necessary to study both types of EC and how they correlate with the time spent for PIPs.

HP4. Those with a higher epistemic curiosity score will spend more time for PIPs.

2.5.2 Compulsive acquisition

Another personality dimension that relates to information appetite is compulsive hoarding, especially acquisition. There are three factors to compulsive hoarding as measured by Frost, Steketee, and Grisham (2004): clutter, difficulty discarding, and acquisition. Among these, of particular interests is the acquisition dimension and how it relates to information appetite. Would those with extreme information appetite exhibit compulsive acquisition symptoms? If so, maybe what matters to them is less what they acquire but more the act of acquiring information. Also, although participants are not asked how much information they consumed because the measure of how much can be very subjective and individual, those with a higher compulsive acquisition score can be assumed to have consumed more information than those with a lower score.

Listed below is the Acquisition subscale items from the Saving Inventory-Revised (SI-R) (Frost et al., 2004, p. 1167) that will be used in data collection for this study. SI-R has been developed specifically for hoarding, all three subscales (clutter, difficulty discarding and acquisition) are found reliable, and it has demonstrated validity as well (Pertusa et al., 2010, p. 380).

- Feel compelled to acquire?
- Lack of control over urges to acquire
- Distress if can’t acquire
- Strength of urge to acquire
- Frequency of buying unneeded items
- Financial difficulties from saving/buying
- Distress over acquiring habits

RQ2. How do the participants' reported compulsive acquisition symptoms relate to the time they spent for PIPs?

2.6 Model of Information Appetite

The diagram below shows the model for this study.

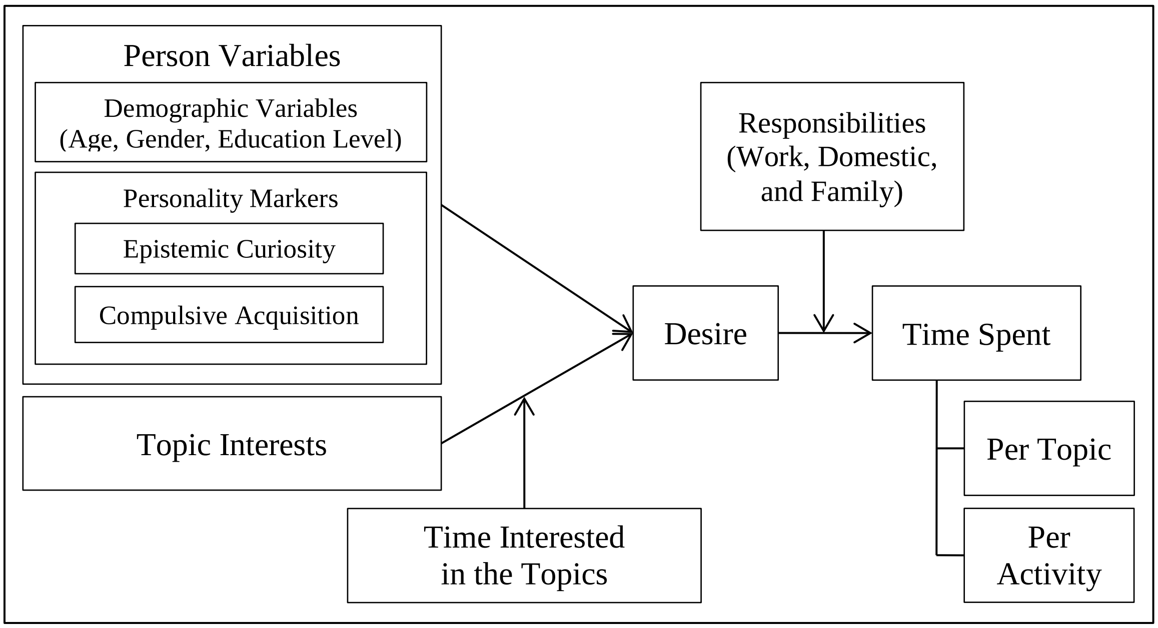


Figure 2.5 Model of information appetite

Below are the hypotheses and research questions proposed in this chapter.

HP1. Information appetite in one type of activity does not predict information appetite in other types of activities.

HP2. Individual characteristics such as age, gender, and education level affect the time spent for PIPs.

HP3. Time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities.

RQ1. What are the relationships between the activity-based information appetite and the topical information appetite along with the topic's temporal context?

HP4. Those with a higher epistemic curiosity score will spend more time for PIPs.

RQ2. How do the participants' reported compulsive acquisition symptoms relate to the time they spent for PIPs?

Chapter 3 Study 1: American Time Use Survey Analysis

In this chapter, American Time Use Survey (ATUS) data has been analyzed to test the first proposition (Individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices (PIPs)), hypothesis 2 (Individual characteristics such as age, gender, and education level affect the time spent for PIPs), and hypothesis 3 (Time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities.)

ATUS is an ongoing time diary study conducted by the U.S. Bureau of Labor Statistics since 2003. The dataset has been extracted using the ATUS-X American Time Use Study Extract Builder service (Hofferth et al., 2013). Data collected from 2003 through 2014 were used (*N* = 159,937).

Here are some sample population characteristics in the ATUS dataset.

- Gender: Male 43.7%, Female 56.3%
- Average age: 47 years old
- About 57% have some college education or hold a college degree.
- About 63% are employed, and 6% hold more than one job.

The data analysis methods were the following. First, cluster analysis was performed to segregate the high information appetite (IA) group, and various analyses were performed to describe the characteristics of the high IA group. Then, multiple regression, factor analysis, and canonical correlation analyses were conducted to investigate the effects of personal and situational variables on the time spent for PIPs.

3.1 Information Appetite Levels

As discussed in 2.3, eight activities have been identified as relevant to PIPs: taking class for personal interest, research or homework for class (for personal interest), researching purchases, television and movies (not religious), listening to the radio, computer use for leisure (excluding games), reading for personal interest, and writing for personal interest. A PIP variable (total time spent in PIP) has been computed from the sum of time spent on all eight PIP activities in the dataset. A Two-Step cluster analysis of the PIP variable resulted in two clusters.

The Two-Step Cluster method was used for two reasons: First, it is based on the Hierarchical cluster method, but unlike Hierarchical cluster analysis it can handle large data files. Second, unlike K-Means cluster analysis where the number of clusters has to be specified manually Two-Step cluster analysis can automatically determine the optimal number of clusters. As for the distance measure, the Euclidean measure was chosen because it is appropriate with continuous variables. The log-likelihood distance measure assumes that continuous variables are normally distributed, but the PIP variable data is positively skewed (skewness: 1.287).

Cluster	*N*	% Distribution	Min	Max	Mean	Std. Deviation	Label
1	144355	90.3%	0	473	160.23	126.350	Regular
2	15582	9.7%	474	1439	621.74	129.887	High IA
Combined	159937	100%	0	1439	205.20	186.498	

Table 3.1 Cluster analysis on the sum of time spent on all PIP activities

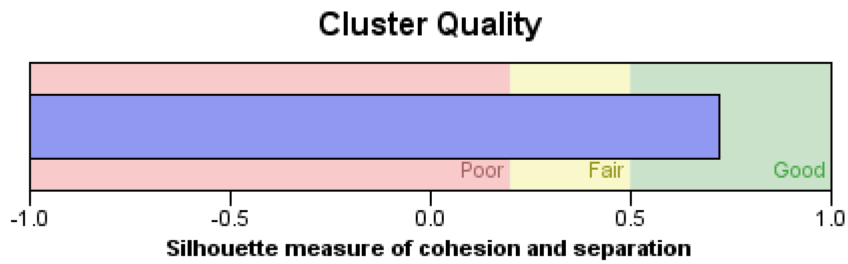


Figure 3.1 Cluster quality measure (Average silhouette = 0.7)

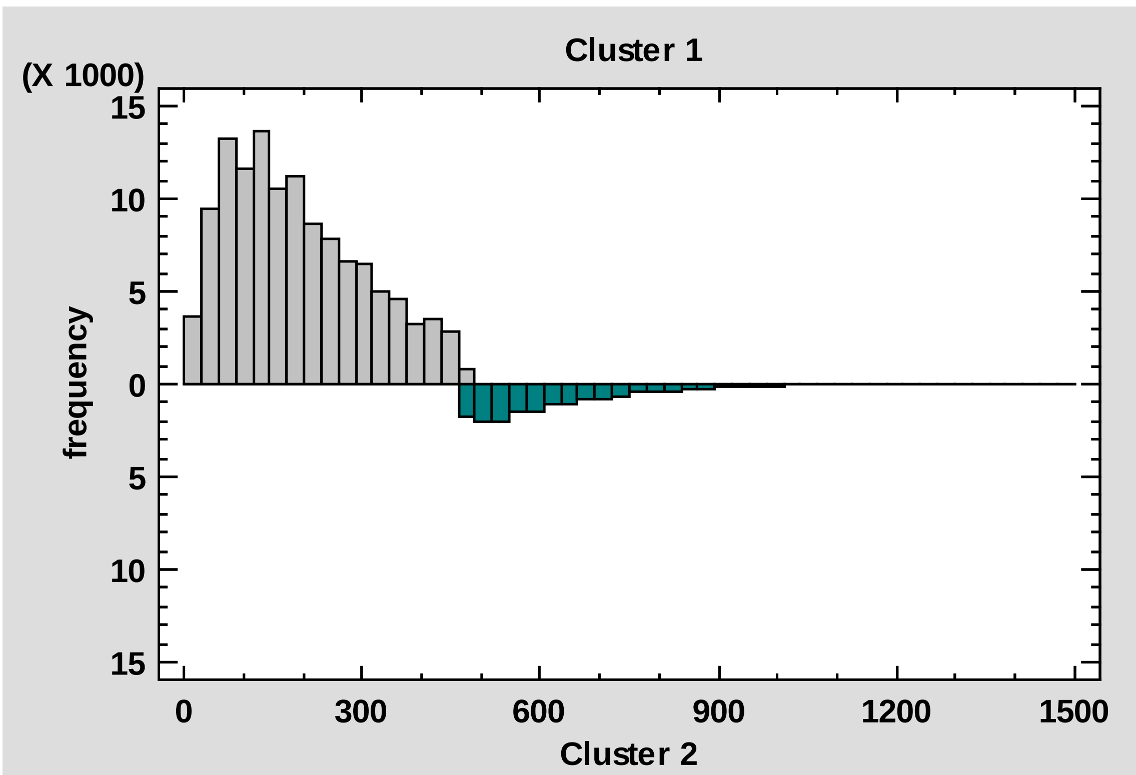


Figure 3.2 Frequency histogram of clusters 1 and 2 (x axis: total time spent in PIPs)

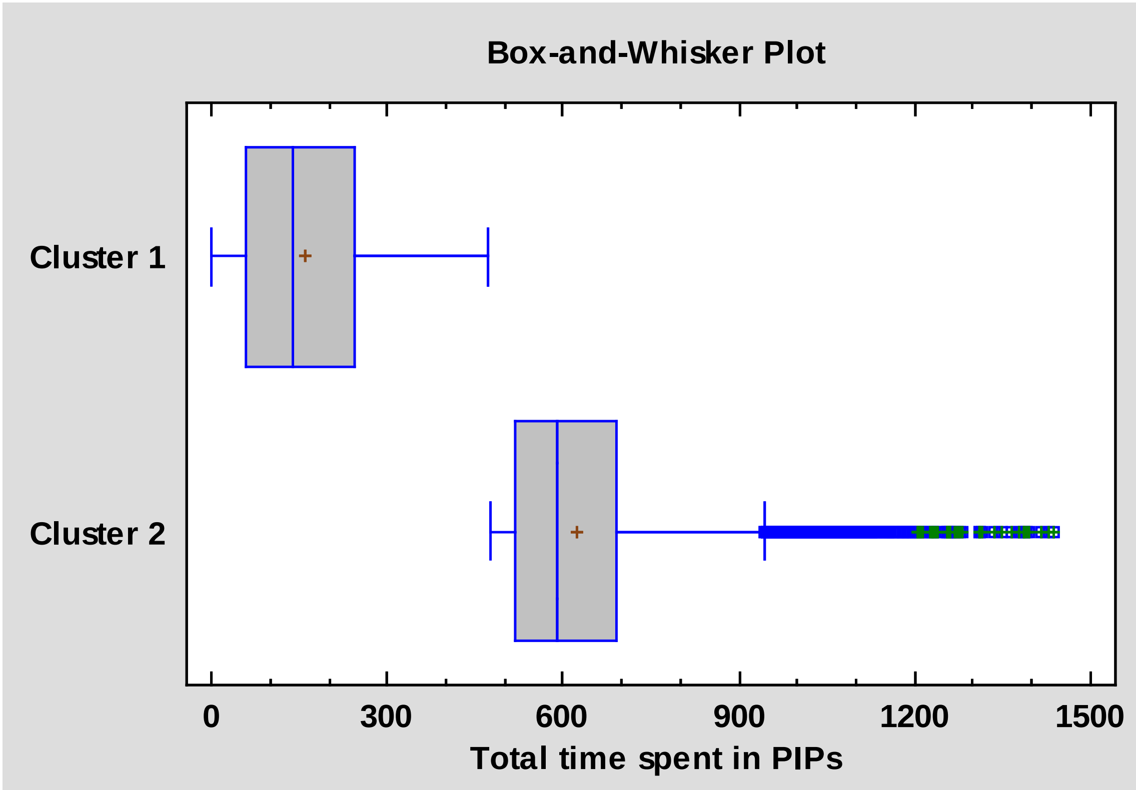


Figure 3.3 Box-and-whisker plot of clusters 1 and 2 (total time spent in PIPs)

The descriptive statistics of each cluster show that cluster 2 is a high information appetite (IA) group. The frequency histogram and the box-and-whisker plot support the label as well.

3.1.1 High information appetite cluster characteristics

The two tables below show that the high IA group tends to be statistically significantly more male and older in age than the regular group cluster. The Pearson chi-square result in table 3.2 shows that males and females are significantly different between clusters ($x^2 = 688.086$, $df = 1$, $N = 159937$, $p < .001$). The effect size (shown in Cramer's V = .066) is quite small, though. The independent samples t-test results in table 3.3 show that the age differences between clusters are also significant with larger than medium effect size ($t(18726) = -75.096$, $p < .001$, $d = .65$).

	Cluster 1	Cluster 2 (High IA)	Total	x^2	*p*	Cramer's V
Gender				688.086	.000	.066
Male	42.6%	53.6%	43.7%			
Female	57.4%	46.4%	56.3%			
Total	100%	100%	100%			

Table 3.2 Chi-square analysis of gender and the clusters

		Mean	Std. Deviation	*t*	*df*	*p*	*d*
Age				-75.096	18726.357	.000	.65
Clusters	1	45.64	17.261				
	2 (High IA)	57.12	18.222				

Table 3.3 T-test result of age differences in the clusters[3]

Findings in the three tables below support HP3. The high IA group has substantially more people not in labor force (x^2 = 7592.531, *df* = 4, *N* = 159937, *p* < .001), and less percentage of people holding more than one job (x^2 = 7074.924, *df* = 2, *N* = 159937, *p* < .001). Also, the high IA group spent significantly less time for work, domestic, and family responsibilities. The effect size (Cohen's *d*) for the differences in time spent working is large at 1.10, indicating a standardized difference between means exceeding one standard deviation. One can assume that the high IA group would have much more free time than the others, which could have led to more time spent in PIPs.

	Cluster 1	Cluster 2 (High IA)	Total	x^2	*p*	Cramer's V
Labor force status				7592.531	.000	.218
Employed – at work	63.2%	28.5%	59.8%			
Employed – absent	2.6%	3.1%	2.7%			
Unemployed – on layoff	0.5%	0.7%	0.5%			

[3] The t and df were adjusted because variances were not equal.

Unemployed – looking	4.4%	5.3%	4.4%			
Not in labor force	29.3%	62.4%	32.5%			
Total	100%	100%	100%			

Table 3.4 Chi-square analysis of the labor force status and the clusters

	Cluster 1	Cluster 2 (High IA)	Total	x^2	*p*	Cramer's V
Has more than one job				7074.924	.000	.210
No	59.4%	29.9%	56.5%			
Yes	6.4%	1.7%	6.0%			
Not applicable	34.2%	68.4%	37.5%			
Total	100%	100%	100%			

Table 3.5 Chi-square analysis of the multiple job status and the clusters

		Mean	Std. Deviation	*t*	*df*	*p*	*d*
Time spent working				212.703	107946.781	.000	1.10
Clusters	1	176.15	247.318				
	2 (High IA)	10.28	53.600				
Time spent on household activities				64.663	25746.141	.000	.45
Clusters	1	125.46	144.525				
	2 (High IA)	73.68	87.972				
Time spent caring for/helping household members				93.893	52518.664	.000	.53
Clusters	1	36.35	84.729				
	2 (High IA)	6.08	29.061				

Table 3.6 T-test results of differences in time spent for responsibilities between clusters[4]

The education level turned out to be lower in the high IA group (x^2 = 1534.714, *df* = 8, *N* = 159937, *p* < .001). However, the effect size (Cramer's V = .098) is on the smaller side.

	Cluster 1	Cluster 2 (High IA)	Total	x^2	*p*	Cramer's V
Highest level of school completed				1534.714	.000	.098
12th grade (no diploma) or below[5]	15.6%	21.4%	16.1%			
High school graduate - diploma or the equivalent (e.g. GED)[6]	25.7%	34.5%	26.6%			
Some college but no degree	17.7%	17.7%	17.7%			
Associate degree in college - occupational/vocational program	4.5%	3.7%	4.4%			
Associate degree in college - academic program	4.8%	3.6%	4.7%			
Bachelor's degree (BA, AB, BS, etc.)	20.0%	12.6%	19.3%			
Master's degree (MA, MS, MEng, MEd, MSW, etc.)	8.6%	4.6%	8.2%			
Professional school degree (MD, DDS, DVM, etc.)	1.6%	1.0%	1.5%			
Doctoral degree (PhD, EdD, etc.)	1.6%	0.9%	1.6%			
Total	100%	100%	100%			

Table 3.7 Chi-square analysis of the education level and the clusters

[4] The t and df were adjusted because variances were not equal in all three cases.
[5] This is recoded from the sum of eight original labels in the variable (Less than 1st grade; 1st, 2nd, 3rd, or 4th grade; and through 12th grade - no diploma).
[6] This is recoded from the sum of two original labels in the variable (GED and diploma).

3.2 Variable Relationships to the Time Spent for PIPs

So far, the analysis showed the differences between clusters by comparing the two groups (e.g. t-test). To find the strength of associations between variables, however, associational inferential statistics such as correlation and regression are needed. This way, one can predict another variable (dependent) with one or more variables (independent). A multiple regression analysis was conducted with the time spent for PIPs as a dependent variable to find out the relationships between personal and situational independent variables and the time spent for PIPs.

The independent variables in the final multiple regression model were the following.

- Age
- Gender (1 = male, 0 = female)
- Time spent working in the reported day (in minutes, ranging from 0 to 1430)
- Has more than one job (1 = yes, 0 = no/NA)

Several independent variables have been removed from the model. First, time spent for household activities and time spent caring for/helping household members were not included because the dependent variable (time spent for PIPs) could potentially affect these two types of time expenditure (e.g. spending too much time reading while sacrificing domestic activity time like cleaning). Allison (1999) pointed out that if the dependent variable affects an independent variable, it might result in a bias of every coefficient in the regression model (pp. 52-54). Therefore, these two independent variables have been removed in the multiple regression model. It is assumed that time spent working is not as affected by the dependent variable because most people are very constrained in terms of the number of hours they work.

However, time spent working (in minutes) turned out to be highly correlated with the “Employed - at work” category in the labor force status for 159,937 respondents ($r = .52$, $p <$,001). Between the labor force status and the time spent working, the model with the time spent working turned out to have a higher explanatory power (higher adjusted $R^2 = .19$ vs $R^2 = .15$), so it was decided to keep the time spent working variable and remove labor force status from the model.

Lastly, highest level of school completed has been removed from the model because it accounted for less than 1% of the variance and the tolerance values in collinearity statistics were low in multiple categories.

According to Allison (1999), independent and dependent variables do not need to be normally distributed in multiple regression (pp. 130-131). What needs to be normally distributed is the disturbance term which cannot be directly observed, but even that is not strictly necessary if the sample size is large (more than 200 cases). Since the total sample size of the ATUS dataset is 159,937, the normality assumption was not considered important in this case.

3.2.1 Multiple regression on time spent for PIPs

Multiple regression was conducted to determine the best linear combination of age, gender, time spent working, and multiple job status for predicting time spent for PIPs. The means, standard deviations, and intercorrelations can be found in the table below.

Variable	M	SD	Age	Gender	Time working	Has > 1 job

Time spent for PIPs	205.20	186.50	.28*	.08*	-.34*	-.10*
Predictor variable						
Age	46.75	17.69	-	-.05*	-.15*	-.07*
Gender (1=male)	.44	.50		-	.14*	.03*
Time spent working	159.99	240.64			-	.18*
Has more than one job (1=yes)	.06	.24				-

*$p < .001$

Table 3.8 Means, standard deviations, and intercorrelations for time spent for PIPs and predictor variables (N = 159,937)

This combination of variables significantly predicted time spent for PIPs, $F(4, 159932) = 9402.04$, $p < .001$, with all four variables significantly contributing to the prediction. The adjusted R^2 value was .19, indicating that 19% of the variance in time spent for PIPs was explained by this model. The beta weights suggest that not spending time working contributes most to predicting time spent for PIPs, and that being older, being male, and not having multiple jobs also contribute to this prediction.

Variable	B	Std. Error of B	beta	Squared part correlation
(Constant)	107.181	1.325		
Age	2.509	.024	.238*	.055
Gender (1=male)	49.932	.854	.133*	.017
Time spent working	-.247	.002	-.319*	.095
Has more than one job (1=yes)	-25.872	1.801	-.033*	.001

Note. Adjusted $R^2 = .19$; $F(4, 159932) = 9402.04$, $p < .001$
*$p < .001$

Table 3.9 Multiple regression analysis summary for age, gender, time spent working, and multiple job status predicting time spent for PIPs (N = 159,937)

For each year increase in age, there was a 2.51 minutes increase in time spent for PIPs. For each minute spent working, there was a .25 minute decrease in time spent for PIPs. Males spent about 50 minutes more in PIPs than females. Those who have more than one job spent about 26 minutes less in PIPs than those without multiple jobs.

The squared part correlations represent the strength of the unique predictive contribution of each particular variable (Warner, 2013). It shows that time spent working uniquely predicted about 9.5% of the variance in time spent for PIPs, when all the other variables were statistically controlled. It also shows that multiple job status uniquely predicted just 0.1% of the variance in time spent for PIPs, which is very low. Age and gender uniquely predicted about 5.5% and 1.7% of the variance in time spent for PIPs, respectively, when the other variables were statistically controlled.

3.2.2 Multiple regression in four groups (not worked/worked, cluster 1/2)

Next, multiple regression was performed on four different groups of respondents. The frequency histogram of the time spent working (see the figure below) showed that there was a large group of people who did not work in the reported day, which would have affected time spent for PIPs. Therefore, the sample was first divided into a group of those who did not work (0 minute spent working) and those who worked (more than 1 minute spent working), and then divided according to the clusters (the high IA group (cluster 2) and the regular group (cluster 1)). Multiple regression was conducted on these four different groups to find out how the independent variables predict the dependent variable (time spent for PIPs).

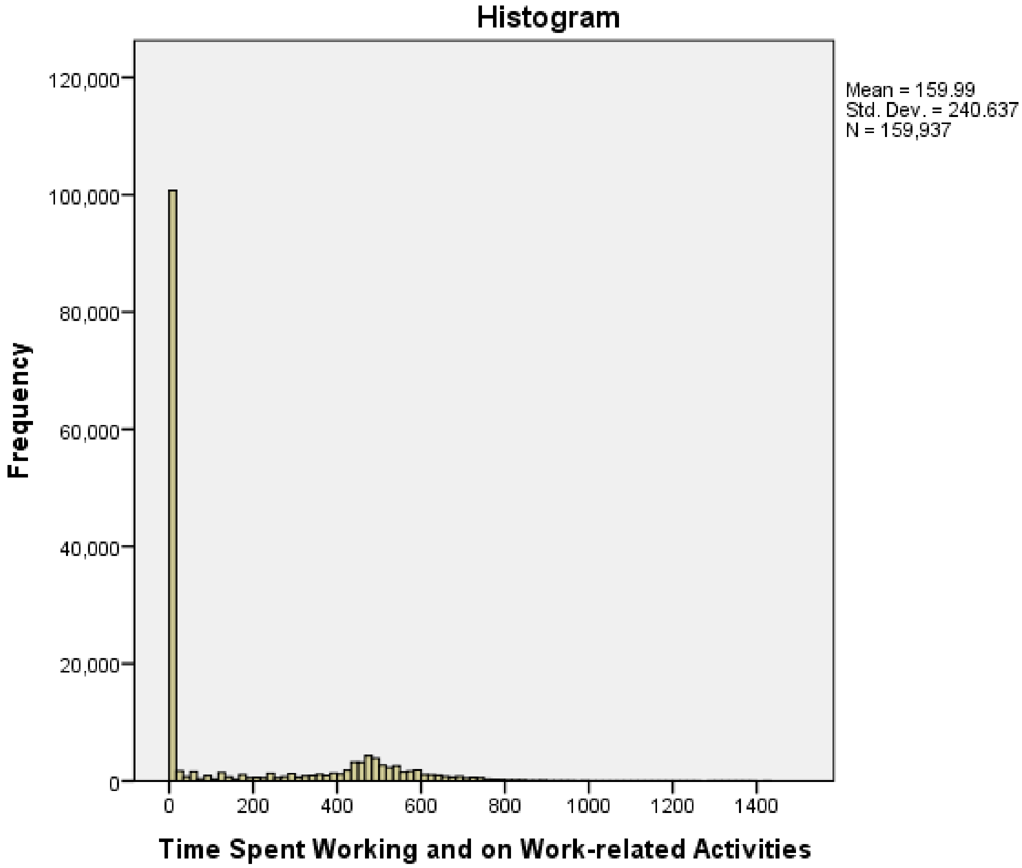


Figure 3.4 Frequency histogram of the time spent working variable

	Not worked		Worked	
	N	Adj. R^2	*N*	Adj. R^2
Cluster 1 (Regular)	84297	.061	60058	.115
Cluster 2 (High IA)	14536	.014	1046	.029

Table 3.10 *N* and adjusted R^2s of the four groups (Not worked/Worked * Cluster 1/2)

3.2.2.1 Not worked, cluster 1

Time spent working was removed from the analysis because it was a constant (0) in this model.

Variable	M	SD	Age	Gender	Time working	Has > 1 job
Time spent for PIPs	188.36	132.759	.226*	.079*	-	-.044*
Predictor variable						
Age	47.79	19.390	-	-.056*	-	-.065*
Gender (1=male)	.37	.482		-	-	.017*

Time spent working	.00	.000			-	-
Has more than one job (1=yes)	.03	.159				-

*$p < .001$

Table 3.11 Means, standard deviations, and intercorrelations for time spent for PIPs and predictor variables in not worked/cluster 1 group (N = 84,297)

This combination of variables significantly predicted time spent for PIPs, $F(3, 84293) = 1811.43$, $p < .001$, with all three variables significantly contributing to the prediction. The adjusted R^2 value was .061, indicating that 6.1% of the variance in time spent for PIPs was explained by this model. The dependent variable in the table below was time spent for PIPs.

Variable	B	Std. Error of B	Beta	Squared part correlation
(Constant)	104.589	1.252		
Age	1.570	.023	.229*	.052
Gender (1=male)	25.607	.921	.093*	.009
Has more than one job (1=yes)	-25.274	2.785	-.030*	.001

Note. Adjusted R^2 = .061; $F(3, 84293) = 1811.43$, $p < .001$
*$p < .001$

Table 3.12 Multiple regression analysis summary for age, gender, and multiple job status predicting time spent for PIPs in not worked/cluster 1 group (N = 84,297)

For each year increase in age, there was a 1.57 minutes increase in time spent for PIPs. Males spent about 26 minutes more in PIPs than females. Those who have more than one job spent about 25 minutes less in PIPs than those without multiple jobs.

The squared part correlations are not much different from the overall sample population multiple regression results shown in the previous section.

3.2.2.2 Not worked, cluster 2 (High IA)

Again, time spent working was removed from the analysis because it was a constant (0) in this model.

Variable	M	SD	Age	Gender	Time working	Has > 1 job
Time spent for PIPs	625.17	131.468	.065**	.093**	-	-.023*
Predictor variable						
Age	57.85	18.197	-	-.084**	-	-.084**
Gender (1=male)	.53	.499		-	-	.012
Time spent working	.00	.000			-	-
Has more than one job (1=yes)	.01	.108				-

*p < .01, **p < .001

Table 3.13 Means, standard deviations, and intercorrelations for time spent for PIPs and predictor variables in not worked/cluster 2 group (N = 14,536)

This combination of variables significantly predicted time spent for PIPs, $F(3, 14532) = 70.93$, $p < .001$, with all three variables significantly contributing to the prediction. The adjusted R^2 value was .014, however, indicating that only 1.4% of the variance in time spent for PIPs was explained by this model. The dependent variable in the table below was time spent for PIPs.

Variable	B	Std. Error of B	Beta	Squared part correlation
(Constant)	581.443	3.903		
Age	.522	.060	.072**	.005
Gender (1=male)	26.207	2.176	.100**	.010
Has more than one job (1=yes)	-22.184	10.106	-.018*	.0003

Note. Adjusted R^2 = .014; $F(3, 14532) = 70.93$, $p < .001$
*p < .05, **p < .001

Table 3.14 Multiple regression analysis summary for age, gender, and multiple job status predicting time spent for PIPs in not worked/cluster 2 group (N = 14,536)

For each year increase in age, there was a .52 minutes increase in time spent for PIPs. Males spent about 26 minutes more in PIPs than females, same as not worked, cluster 1 group. Those who have more than one job spent about 22 minutes less in PIPs than those without multiple jobs.

The squared part correlations show very small values, which is not surprising given the low R^2 value.

3.2.2.3 Worked, cluster 1

Unlike in the not worked groups, time spent working was included in the model in the worked groups since it is not a constant anymore (1+).

Variable	M	SD	Age	Gender	Time working	Has > 1 job
Time spent for PIPs	120.75	104.714	.130**	.082**	-.287**	-.053**
Predictor variable						
Age	42.61	13.144	-	.008*	.005	-.008*
Gender (1=male)	.51	.500		-	.118**	.003
Time spent working	423.38	205.773			-	.030**
Has more than one job (1=yes)	.12	.323				-

*p < .05, **p < .001

Table 3.15 Means, standard deviations, and intercorrelations for time spent for PIPs and predictor variables in worked/cluster 1 group (N = 60,058)

This combination of variables significantly predicted time spent for PIPs, $F(4, 60053) = 1949.73$, $p < .001$, with all four variables significantly contributing to the prediction. The adjusted R^2 value was .115, indicating that 11.5% of the variance in time spent for PIPs was explained by this model. The dependent variable in the table below was time spent for PIPs.

Variable	B	Std. Error of B	beta	Squared part correlation
(Constant)	130.352	1.626		
Age	1.039	.031	.130*	.017
Gender (1=male)	24.369	.810	.116*	.013
Time spent working	-.153	.002	-.300*	.089
Has more than one job (1=yes)	-13.910	1.247	-.043*	.002

Note. Adjusted R^2 = .115; $F(4, 60053) = 1949.73$, $p < .001$
*p < .001

Table 3.16 Multiple regression analysis summary for age, gender, time spent working, and multiple job status predicting time spent for PIPs in worked/cluster 1 group (N = 60,058)

In this group, for each year increase in age, there was about 1 minute increase in time spent for PIPs. For each minute spent working, there was a .15 minute decrease in time spent for PIPs. Males spent about 24 minutes more in PIPs than females. Those who have more than one job spent about 14 minutes less in PIPs than those without multiple jobs.

The squared part correlations show that time spent working is the biggest predictor of the time spent for PIPs.

3.2.2.4 Worked, cluster 2 (High IA)

Variable	M	SD	Age	Gender	Time working	Has > 1 job
Time spent for PIPs	574.05	93.266	.013	.099**	-.151***	-.027
Predictor variable						
Age	46.94	15.308	-	-.053*	-.050	-.040
Gender (1=male)	.67	.469		-	-.023	-.045
Time spent working	153.11	144.735			-	.006
Has more than one job (1=yes)	.09	.286				-

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

Table 3.17 Means, standard deviations, and intercorrelations for time spent for PIPs and predictor variables in worked/cluster 2 group (N = 1,046)

This combination of variables significantly predicted time spent for PIPs, $F(4, 1041) = 8.78$, $p < .001$, with two of the four variables (gender, time spent working) significantly contributing to the prediction. The adjusted R^2 value was .029, indicating that only 2.9% of the variance in time spent for PIPs was explained by this model. The dependent variable in the table below was time spent for PIPs.

Variable	B	Std. Error of B	beta	Squared part correlation

(Constant)	573.813	10.933		
Age	.059	.186	.010	.0001
Gender (1=male)	18.966	6.075	.095*	.009
Time spent working	-.096	.020	-.148**	.022
Has more than one job (1=yes)	-7.022	9.956	-.022	.0005

Note. Adjusted R^2 = .029; $F(4, 1041) = 8.78$, $p < .001$
*$p < .01$, **$p < .001$

Table 3.18 Multiple regression analysis summary for age, gender, time spent working, and multiple job status predicting time spent for PIPs in worked/cluster 2 group (N = 1,046)

In this group, for each minute spent working, there was about .1 minute decrease in time spent for PIPs. Males spent about 19 minutes more in PIPs than females.

The squared part correlations show that time spent working is the biggest predictor of the time spent for PIPs, the same with the worked/cluster 1 group.

3.2.3 Factor analysis

Factor analysis was performed to group the eight time use variables into a smaller number of variables. Principal component analysis (PCA) was employed because PCA is useful when "one is simply trying to mathematically derive a relatively small number of variables to use to convey as much of the information in the observed/measured variables as possible" (Leech, Barrett, & Morgan, 2015, p. 68). Therefore, for the purpose of running this analysis here (i.e. reduce the number of time use variables), PCA was appropriate.

PCA with varimax rotation was conducted to assess how the eight time use variables clustered. These variables were time spent for taking class for personal interest, research or homework for class (for personal interest), researching purchases, television and movies (not religious), listening to the radio, computer use for leisure (excluding games), reading for personal interest, and writing for personal interest. Four components were rotated, based on the eigenvalues over 1 criterion. After rotation, the four components combined accounted for 51.8% of the variance; component 1 accounted for 13.4% of the variance, component 2 13%, component 3 12.8%, and component 4 12.5%. The table below displays the items and component loadings for the rotated components.

	Component			
	1	2	3	4
Taking class for personal interest	**0.725**	0.038	-0.008	0.015
Research or homework for class (for personal interest)	**0.718**	0.001	-0.012	-0.008
Computer use for leisure (excluding games)	-0.080	**0.653**	-0.224	-0.098
Television and movies (not religious)	-0.141	**-0.613**	-0.193	-0.044
Listening to the radio	0.028	-0.256	**0.715**	-0.037
Reading for personal interest	-0.061	0.296	**0.638**	-0.023
Researching Purchases	-0.046	0.156	0.064	**0.895**
Writing for personal interest	-0.048	0.251	0.113	**-0.431**

Figure 3.5 Time use variables and component loadings for the rotated components (N = 159,937)

With this result, four new variables have been created. For each component, the sum was computed from the variables (i.e. time spent in minutes) that loaded highly on the component. The new variables are Classes (taking class, research/homework for class), Computer/TV/Movies (computer use for leisure, TV and movies), Radio/Reading (listening to the radio, reading for personal interest), and Researching purchases/Writing (researching purchases, writing for personal interest).

3.2.4 Canonical correlations using the four component variables

Canonical correlation was used to investigate the relationships between the independent variables and the four component variables. Canonical correlation is useful when exploring how the differences in one set of variables relate to the difference in another set of variables.

Since canonical correlation allows scale (numeric) variables only, the independent variables were Age and Time spent working. The variables time spent for household activities and time spent caring for/helping household members were not included because canonical correlation is an extension of multiple regression (Leech et al., 2015): they had been removed in the multiple regression model because the dependent variable (time spent for PIPs) could have inversely affected them, which might result in a bias of coefficients in the regression model (Allison, 1999).

First, canonical correlation was performed with all data in the dataset (*N* = 159,937) to assess the pattern of relationships between age and time spent working, and the four component variables of time spent for PIPs. The first canonical correlation was .437 (19% overlapping variance) and the second was .125 (1.6% overlapping variance). With both canonical correlations included, $x^2(8) = 36445$, $p < .001$, and with the first removed, $x^2(3) = 2506.85$, $p < .001$. The correlations and canonical coefficients are included in the table below. Examination of the loadings suggests that the first canonical correlation seems to involve a relation between age and time spent working, and spending time on computer/TV/movies and on radio/reading. The second canonical correlation loadings capture a similar relationship as well.

	First canonical correlation		Second canonical correlation	
Item content	Loading	Coefficient	Loading	Coefficient
Predictor variables				
Age	-.752	-.649	-.659	-.776
Time spent working	.767	.667	-.641	-.761
Component variables (Time spent)				
Classes	-.018	-.039	.196	.210
Computer/TV/Movies	-.774	-.793	.610	.594

Radio/Reading	-.609	-.631	-.780	-.763
Researching purchases/Writing	-.036	-.033	.032	.045

Table 3.19 Correlation and standardized canonical coefficients between predictor and component variables (*N* = 159,937)

Second, canonical correlation was performed on cluster 1 (*N* = 144,355). The first canonical correlation was .392 (15.4% overlapping variance) and the second was .135 (1.8% overlapping variance). With both canonical correlations included, $x^2(8) = 26706.71$, $p < .001$, and with the first removed, $x^2(3) = 2672.12$, $p < .001$. The correlations and canonical coefficients are included in the table below. Examination of the loadings suggests a similar relationship as in the all-dataset canonical correlation.

	First canonical correlation		Second canonical correlation	
Item content	Loading	Coefficient	Loading	Coefficient
Predictor variables				
Age	-.747	-.664	-.665	-.758
Time spent working	.752	.670	-.659	-.753
Component variables (Time spent)				
Classes	-.014	-.039	.173	.195
Computer/TV/Movies	-.679	-.737	.711	.667
Radio/Reading	-.678	-.734	-.724	-.674
Researching purchases/Writing	-.023	-.025	.054	.070

Table 3.20 Correlation and standardized canonical coefficients between predictor and component variables in cluster 1 (*N* = 144,355)

Third, canonical correlation was performed on cluster 2, the high IA group (*N* = 15,582). The first canonical correlation was .226 (5.1% overlapping variance) and the second was .079 (0.6% overlapping variance). With both canonical correlations included, $x^2(8) = 912.85$, $p < .001$, and with the first removed, $x^2(3) = 96.6$, $p < .001$. The correlations and canonical coefficients are included in the table below. Examination of the loadings and coefficients suggests that the first canonical correlation involves a positive relationship between age and time spent on radio/reading; whereas, the second indicates a negative relationship between time spent working and time spent on computer/TV/movies.

	First canonical correlation		Second canonical correlation	
Item content	Loading	Coefficient	Loading	Coefficient
Predictor variables				
Age	-.994	-.981	.110	.225
Time spent working	.224	.111	.975	1.000
Component variables (Time spent)				
Classes	.198	.094	-.077	-.309
Computer/TV/Movies	.364	-.475	-.868	-1.339
Radio/Reading	-.925	-1.246	.327	-.589
Researching purchases/Writing	-.015	-.066	-.040	-.183

Table 3.21 Correlation and standardized canonical coefficients between predictor and component variables in cluster 2 (high IA group) (N = 15,582)

3.3 Discussion

The first proposition was that individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices (PIPs). The distribution of the total time spent for PIPs was positively skewed in that there were many people who did not spend time for PIPs, but at the same time there were a small number of people that spent a lot of time in PIPs. In addition, cluster analysis of the total time spent for PIPs resulted in identifying the high IA group. This supports that individuals spent varying degrees of time in PIPs, reflecting their information appetites.

As for hypothesis 2 (Individual characteristics such as age, gender, and education level affect the time spent for PIPs), the high IA group tended to be more male and older in age. Both the chi-square analysis and the multiple regression analysis support this finding. The multiple regression findings suggest that males spend more time than female in PIPs, and older people spend more time in PIPs as well. As for the education level (highest level of school completed), it was not included in the multiple regression model because it only explained less than 1% of the variance in time spent for PIPs and there were multiple collinearity issues. When looked at the chi-square statistics of the high IA group, however, although the effect size there was on the smaller side as well, the education level turned out to be lower in the high IA group. This could mean that those with lower education level compensate for their relative lack of knowledge with higher IA levels and more time spent for PIPs.

Hypothesis 3, time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities, was investigated with a t-test (with the high IA group vs the regular group) and multiple regression analysis. The high IA group spent less time working, less time for household activities, and less time caring for/helping household members. Times spent for domestic and family responsibilities were not included in the multiple regression model because these variables can be inversely affected by the dependent variable (e.g. neglecting to clean the house because of too much time spent reading). Time spent working was included in the model, however, and found to be the most powerful factor in predicting time spent for PIPs in a negative direction. That is, the less time spent working, the more time one could spend for PIPs. In addition, chi-square analyses of the high IA group reveal that the high IA group is mostly not in the labor force, and if in the labor force does not tend to have multiple jobs, also indicating that having more free time is an important factor in predicting time spent for PIPs.

The multiple regression model for predicting time spent for PIPs included predictors such as age, gender, time spent working, and multiple job status. However, the four groups (not worked/worked in the reported day, cluster 1/2) multiple regression analyses showed that this is not an effective model for predicting time spent for PIPs in high IA groups, as evidenced by low R^2 scores. In the worked/cluster 2 group, age and multiple job status turned out to be non-significant predictors, leaving only gender and time spent working as viable predictors. Time spent working was the biggest (negative) predictor of time spent for PIPs in worked groups. Older respondents, males, and those without multiple jobs spent more time for PIPs in all four groups.

The canonical correlation result for the high IA group showed that age and time spent on radio/reading have a positive relationship (effect size = 5.1%), and independently, time spent working and time spent on computer/TV/movies have a negative relationship (effect size = 0.6%).

What is interesting here is the overall low effect size of the high IA group analyses. This means that some other factor that is not investigated in this chapter is operating in the high IA group, as the ATUS data analyses failed to capture all the variances that affect the high IA group. What they are might be beyond the scope of the ATUS dataset, however.

Chapter 4 Study 2: Method

4.1 Research Design

4.1.1 Research design of this study

The study design is based on the theoretical model presented in earlier chapters, and encompasses the key concepts and variables to be explored in this research. The design appropriate for this investigation attempts to explain information seeking behavior with individuals' allocation of their time. As such, this exploratory study will employ a survey that may identify correlational relationships among the variables using simple and multivariate models.

The survey was conducted online via Qualtrics, an online survey software system. The survey link was provided in an Amazon Mechanical Turk (MTurk) job. MTurk is an online labor market that has become popular among social scientists for its large pool of constantly available participants and the low cost (Paolacci & Chandler, 2014).

MTurk provides several notable advantages for this study:

- Ease of reaching many participants online
- Being able to reach individuals with wider demographic profiles than college students
- The speed with which data is collected
- Ability to employ pre-tests quickly to apply any changes to the final survey
- Cost-effective and streamlined payment of participants based on work performed

Although Amazon raised their fee from 10% to 40% (for participants more than ten; 20% for less than that) in July 2015, it is still an economical way of gathering a large amount of research data in a relatively short amount of time.

One drawback of using MTurk is that the method is mostly limited to quantitative work. Qualitative techniques such as interviews or focus groups cannot be used, but the possible methods are mostly surveys, experiments, or short essays. Hence, this study consists of a survey questionnaire constructed for quantitative analysis with an optional open-ended qualitative question at the end. It is cross-sectional with all the data being collected at a single point in time.[7]

Two preliminary tests with a small number of subjects (twelve and nine, respectively) were conducted to test the survey procedure. The first preliminary test resulted in a major revision of the research design and data collection procedures, and will be discussed later. The second test was performed to make sure everything runs smoothly before the survey was launched in full scale.

4.1.2 Study procedures

Here are the detailed survey procedures on MTurk:

1. Participant finds the job on MTurk.
2. Participant agrees to the consent form online.
3. Participant fills out the survey online.

[7] Although it is theoretically possible to recruit participants in chunks over a longer period of time (e.g. 100 people every day for six days), I decided not to because it is possible for someone to participate in all six surveys and get paid, resulting in unknown duplicate entries which is problematic. There is a way to prevent a previous survey-taker from taking subsequent surveys, but it requires assigning each previous participant in a group manually which is very labor intensive to do in the current MTurk web interface. Since this is not a critical issue for the purpose of this research, I decided to stick to cross-sectional data gathering.

4. Participant receives a random code at the end of the survey, and enters it in the MTurk job form to receive reward.

4.1.3 Alternative designs

A number of alternative designs are possible to investigate information appetite. Some of these include participant observation, experiment, time diary, and experience sampling method. Furthermore, each approach could be used to investigate participants over time (i.e. longitudinal) or done as a snapshot of information behavior at a particular time (i.e. cross-sectional). Each of these approaches comes with advantages and disadvantages, and such an investigation presents challenges in terms of the extent of effort to achieve a particular goal. Each of these alternatives will now be highlighted as for how fit it is for an initial exploration of the concept of information appetite studied in terms of the differences in the amount of time individuals spend with information.

4.1.3.1 Participant observation

Participants could be observed and records taken in terms of how much time they spend for personal information practices (PIPs). There are three possible ways to investigate this using such a method. One is to invite participants into a lab and observe their behavior; second is to go to their homes and observe what they do at home; and third is using software that monitors activities on their computer and their phone. The first method is more convenient than going to someone's home, but it creates an artificial environment and it might be inappropriate to conclude that such observations represent natural behavior. The second method, observing participants in their home, lets the researcher observe more natural behavior in the most naturalistic setting, but the presence of a researcher at home could create some bias and open up privacy issues. Alternatively, the researcher could ask another member of the household to observe and record the participant. This is much more naturalistic but requires extra training of the other household member as a record-taker. The third method, monitoring with software, requires that participants install a monitoring program on their computer and their phone, and let the software record which websites they visited, which applications they used, and how much time they spent on each activity. The advantage is that the researcher can get a thorough picture of one's online activities. Moreover, the software can be installed on many computers and phones at the same time, so data collection can be done at a much larger scale than the first two methods that require a record-taker's presence. The disadvantage is that the participants might not do what they normally do for fear of invasion of privacy. In addition, this method records online activities only, so offline activities are missed.

4.1.3.2 Experiment

An experiment has the advantage of controlling for extraneous variation by limiting the variables under investigation. However, it may not be suitable for time spent for PIPs because inviting people to a lab creates an artificial environment and it would be harder to observe natural behavior. Nonetheless, for topic-related information appetite it could be a useful method. Participants could be asked to pick two or three topics that they are interested in, asked to indicate for how long they have been interested in them, and then instructed to find the latest news on the topics. It could be hypothesized that the amount of time spent in finding the news would be related to their information appetite level in that those with high information appetite would spend more time finding the news story to share, browsing more sources, and so on. A control group would be assigned a search topic of the researcher's choice. An experiment like this would require an environment where individuals would not be limited by time, which could be problematic to implement, however.

It was considered to run the experiment on MTurk, but the idea was rejected because time spent can be largely impacted by concerns of money. MTurk participants are strongly motivated by monetary compensation (Paolacci & Chandler, 2014), and for a flat amount of compensation it is to their advantage to spend as little time as possible so that they can finish one job quickly and move on to other jobs in order to earn more money. Variable compensation depending on time spent can also skew the result because the participants will be motivated to spend time not because of their information appetite but because of monetary motives. Therefore, MTurk was rejected as a platform for an experiment, but it was still left available for a survey.

4.1.3.3 Time diary

Time diary is a frequently used method for time studies. The biggest reason is that it is hard to lie in time diaries because in order to change time spent on one activity, one has to change the time spent on all the other activities surrounding it. Due to memory and recall problems, past time diaries are usually kept for one day (Harvey, 1999; J. P. Robinson & Godbey, 1999). Future time diaries can be used for more days than one in that the participants are briefed on how to fill out a time diary and then asked to complete it as their days go on. One-day diary has a potential pitfall of capturing a non-typical day, and a week-long diary might solve the issue. However, week-long diaries have a cooperation-rate problem. Other national studies of week-long diaries had a cooperation rate of about 40 percent only (J. P. Robinson & Godbey, 1999, p. 62).

For the purpose of exploring information appetite and how individuals spend time with information, the time diary method can be too complex and resource-consuming. It is more suitable for getting a complete picture of how an individual's daily activities went, but this is not strictly necessary in exploring how much time an individual spends in PIPs. It also does not capture topic interests, which necessitates a survey.

4.1.3.4 Experience sampling method

A beeper study, or Experience Sampling Method (ESM), is a way of sampling human daily life by paging participants at various times of day and asking them to record the nature and quality of their experience (Kubey, Larson, & Csikszentmihalyi, 1996). In traditional beeper studies, participants carry a beeper and a protocol, and when the beeper goes off they fill out the protocol (Harvey & Pentland, 1999, p. 4). With this method, researchers can learn the participants' behavior, feelings, and experience such as product use. It provides a means of understanding human behavior in a contextual and quasi-naturalistic setting, and the researchers can study a larger number of participants using this method than using an ethnographic observation method.

It makes for a viable tool for a number of theoretical and applied questions in social science, including library and information studies. Any research problem that deals with understanding human behavior or emotion in daily life can be answered with this method (Kubey et al., 1996). In information studies, it has been used to study how individuals assess information credibility on the web (Kim, Rieh, Yang, & St. Jean, 2009; Rieh, Kim, Yang, & St. Jean, 2010). Library professionals who would like to learn more about their patrons' experience with the library can also utilize this method to learn their behavior and feelings in situ.

Yet, this method has not been used much in library and information studies in the past. One reason would be its cost and logistical challenges. Beepers are almost obsolete and hard to procure, and they need to be distributed and then collected after research is done. If phone calls are used instead for the signal, it requires human resource, which is costly.

Instead of beepers, many individuals carry smartphones nowadays. According to Pew Research Center's recent reports, 64% of American adults are smartphone owners and 36% of smartphone owners use messaging apps (Duggan, 2015; Pew Research Center, 2015b). Once a message is received on a messaging app, the smartphone notifies the user, just as beepers would. It is the current century's replacement of beepers in ESM.

Moreover, on messaging apps you can send a longer text message than with a beeper, and when combined with mobile web, this means that researchers can send a link to a mobile-compatible online survey that can then be quickly completed by the participants. Instead of carrying a beeper and a protocol, participants will be notified (by a messaging app) and surveyed (on the mobile web) on their smartphones. Many smartphone messaging apps are also compatible with personal computers (PCs), which facilitates a participant's use of a computer to fill out the survey. Pew Research Center (2015a) has employed the ESM with smartphones in nationally representative survey studies, and others are showing an interest in incorporating smartphones in survey research (e.g. Sonck & Fernee, 2013).

This method was tested in the first pretest of this study. Participants were asked to complete an introductory survey, and then were signaled in a smartphone messenger app (programmed using Telegram messenger app's bot API system) on three randomly-selected days to complete a follow-up survey where time use variables were asked. All four surveys asked about time spent in the past three hours. As a result of this pretest, two problems emerged. The first problem was with the participation rate. Of the twelve who completed the introductory survey[8], only eight people participated in the follow-ups, and just five of them completed all three of the follow-up surveys, resulting in fragmented data. The second problem was with the user interface of MTurk's bonus payment system. The introductory surveys were paid out as a job completed, but the follow-ups which happened outside the MTurk eco-system (via signaling on a smartphone messenger app) were paid for as bonuses on MTurk. The bonus payment system on the web was very complicated to use, however, and required too much time and complex record-keeping to deploy for hundreds of participants. Therefore, a decision was made not to use this method but survey the participants for a single three-hour period of time use.

4.1.4 Working with time estimates

Robinson and Godbey (1999) listed multiple risks in asking participants "how many hours do you work?" or "how many hours do you watch TV?" They argued that this assumes that each participant (pp.58-59):

- Interprets "work" or "TV" the same way.
- Separates the most important activity (the primary activity) from other activities that are taking place simultaneously but are ancillary or less important (secondary activities).
- Undertakes the work of searching memory for all episodes of work or television yesterday or the last week.

[8] The pretest was released for nine participants on MTurk, but twelve people ended up completing the introductory survey, resulting in three people not being compensated for their work. I posted a question on Reddit's MTurk community, and the most plausible response was that it might be due to the survey link being available in the job preview before the job is accepted by the participant. Therefore, in the second pretest and the final survey I employed Javascript codes that display the survey link only after a participant accepted to do the job. It seemed to have worked because the second pretest was released for nine and completed by nine, and the final survey was released for 591 and was completed by 593 with only two overages without compensation. I still do not have an answer why two more people did the work in the final survey without reporting their completion for compensation, though.

- Is able to properly add up all the episode lengths across the day yesterday or across days in the last week.
- Feels comfortable describing this duration to an interviewer when it may not be a typical day or week.
- Avoid reverting to social norms, stereotypes, or images of themselves about how much a "normal" person ought to work, like the normal 40-hour workweek. (p. 59)

The above list of problems for time-estimate questions is applicable to many typical surveys. However, the survey method was selected instead of a time diary for the second reason they list above in that a survey is more conducive to capturing ancillary activities as discussed in chapter 2, section 2.3. Moreover, time use activities are not to be treated as mutually exclusive in the survey for this study. This means that if someone spent 30 minutes to do personal research on computer, the 30 minutes count towards both Computer use and Reading/doing research, which is harder to accomplish with time diary. The survey method facilitates asking about topic interests, temporal history of the interest and time spent on topics.

In addition, for the third and fourth problems (memory and episode lengths), instead of asking about a typical day or week, the survey asks specifically about the last three hours, a very small time period that is the freshest in the participants' minds. This is done to minimize the impact of memory bias as an issue. Robinson and Godbey (1999) call this a random-hour technique (p. 62). Calculating episode lengths would also be easier with the small time period. Moreover, for time data entry method, a slider is employed instead of text entry. This way, participants can visually add up episode lengths using a slider instead of having to do the math in the head.

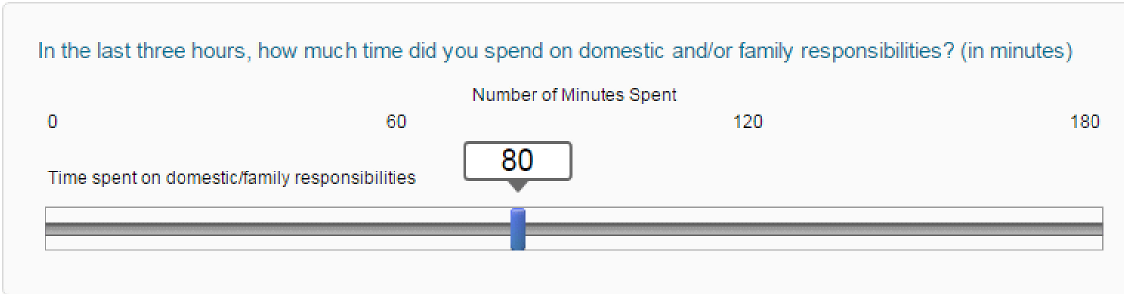


Figure 4.1 Setting the value of a slider question on Qualtrics

4.2 Participant

4.2.1 Amazon Mechanical Turk

Data was collected online using Qualtrics that can be used with an online sample group such as MTurk. To establish similar parameters in MTurk demographics (i.e. those living in the same country) and to limit the scope of time zones to a manageable size for analysis, the sample was limited to the residents of continental U.S. (US/Eastern, US/Central, US/Mountain, and US/Pacific time zones only). Age restriction was 18 years of age or older.

One thing to note about MTurk is the population characteristics. It is unlikely to be representative of the general population: previous studies using MTurk show that Internet users are different from non-Internet users, and MTurk workers tend to be "younger (about 30 years old), overeducated, underemployed, less religious, and more liberal than the general population" (Paolacci & Chandler, 2014, p. 185). However, because this study is not about generalizing the result over the US population but more about understanding a human behavioral characteristic, as long as the sample population can be well identified and described, it is not a deterrent enough to reject MTurk for such research.

Amazon allows requesters (the ones settings up the jobs on MTurk) to set participant parameters such as total number of jobs approved and job approval rate. In the first pretest no parameters were set, and some participants were found to be obvious beginners in using the system and did not follow instructions well. Therefore, from then on, in order to avoid complete beginners, the parameters were set to 98% approval rating and 1,000 jobs[9] approved.

4.2.2 Alternatives

An alternative to MTurk is using college students. College students are convenient to access and can be recruited in a large mass. However, college students have a limited generalizability and can overly emphasize a restricted range of individuals with similar characteristics that afforded them an admission to a university. In addition, there was a study where college student samples have been found unreliable and the findings suspect (Peterson & Merunka, 2014).

Moreover, one recent study compared MTurk and college student populations in terms of how attentive they are to instructions, and found that MTurk participants were more attentive and responded more to small text manipulations than the undergraduates (Hauser & Schwarz, 2015).

Another alternative is to use the general public. However, random sampling of the general public population is particularly challenging and requires excessive time and money for this exploratory study. Even if there were a way to identify and contact a random sample of individuals from a general population, there is still the question of

[9] MTurk workers can accept 100 jobs in their first 10 days, and then can do up to 3,800 jobs per day afterwards.

what types of incentives would be ideal to get the most participation in this study. Hence, for practical reasons, it was decided to use the MTurk instead.

4.2.3 Sample size

The sample size can be determined by use of a formula. The larger the sample size, the more likely it is representative of the population characteristics (Salkind, 2012). Qualtrics offers the following formula for an unknown or a very large population size (Smith, n.d., p. 3):

Necessary Sample Size = $(\text{Z-score})^2 * \text{StdDev} * (1\text{-StdDev}) / (\text{margin of error})^2$

Z-score corresponds to the confidence level, which means how confident one wants to be that the actual mean falls within the confidence interval (also called margin of error in this formula.) With 98% confidence level (z-score: 2.326), 0.5 standard deviation and 95% confidence interval (+/- 5% margin of error), the sample size comes out to $2.326^2 * 0.5 * 0.5 / 0.5^2 = 5.410276 * 0.25 / 0.0025 = 1.352569 / 0.0025 =$ 541.0276. Considering there will be some missing and/or unusable entries, 600 is appropriate given these parameters.

4.3 Measurement/Variables

The survey collects data on the following.

- Demographic: Gender, Age, Local time zone
- Socio-economic status: Education level, Labor force status
- If employed, time spent actually working in the past three hours (in minutes)
- Time spent on domestic/family responsibilities in the past three hours (in minutes)

- Time spent on personal information practice (PIP) activities in the past three hours (in minutes)
- Topic interests (Name two topics that the participant was interested in in the last three hours)
- For each topic, for how long the participant has been interested in it
- Time spent on each of the two topics in the past three hours (in minutes)
- Epistemic curiosity scale (J. Litman, n.d.; J. A. Litman, 2008; J. A. Litman & Spielberger, 2003)
- Acquisition subscale from the Saving Inventory-Revised scale (Frost et al., 2004)
- [Optional] An open-ended question on the topic of interest and how the participant spent time with information regarding the topic(s) in the last 24 hours

The optional open-ended question was designed to capture topic interests and time use activities that may not have occurred in the last three hours since three hours is such a small window of time.

The unit of analysis is the person. Although it is possible that epistemic curiosity may occur at a topic level, Litman's scale has been tested at a person level only, so it will not be used for individual topics.

The complete questionnaire is in Appendix B.

Hypothesis/Research Question	Item on the Survey (Variables)
HP1. Information appetite in one type of activity does not predict information appetite in other types of activities.	Question 9: Time spent for PIPs

HP2. Individual characteristics such as age, gender, and education level affect the time spent for PIPs.	• Questions 1, 2, 4: Demographic variables • Question 9: Time spent for PIPs
HP3. Time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities.	• Questions 5, 6: Labor force status • Questions 7, 8: Time spent for work, time spent for domestic and/or family responsibilities • Question 9: Time spent for PIPs
RQ1. What are the relationships between the activity-based information appetite and the topical information appetite along with the topic's temporal context?	• Question 9: Time spent for PIPs • Questions 10, 11, 12: Interested topics, topics' temporal context, time spent for topics • Question 21: Topics, temporal context and time spent in the last 24 hours
HP4. Those with a higher epistemic curiosity score will spend more time for PIPs.	• Question 9: Time spent for PIPs • Question 13: Epistemic curiosity scale
RQ2. How do the participants' reported compulsive acquisition symptoms relate to the time they spent for PIPs?	• Question 9: Time spent for PIPs • Questions 14 – 20: Acquisition (SI-R) scale

Table 4.1 Hypothesis/research questions, and items on the survey (variables)

The number of variables used in this study has been kept purposefully small, primarily in order to reduce the complexity of the information appetite model. Also, an increase in the number of variables means potentially an increase in the sample size to maintain the same level of statistical power as the power is a function of sample size and is directly related to the number of variables investigated (Krathwohl, 1997). Finally, if the number of variables is large, for example several dozen variables, it implies that, independent of each other, they respectively explain only a very small percentage of the overall variance or effect size.

The following explains each of the variables used in this study.

1. Demographic and socio-economic status (Gender, Age, Education level, Labor force status): These are included as personal characteristic variables. They are routinely investigated in time use studies (e.g. J. P. Robinson & Godbey, 1999), as well as in ATUS.
2. Personality markers
   2.1. Epistemic curiosity scale (J. Litman, n.d.; J. A. Litman, 2008; J. A. Litman & Spielberger, 2003): This scale measures two types of intellectual curiosity. One is I-type that is about the pleasure in learning new ideas, which is similar to browsing in information behavior. The other is D-type that is about spending time and effort to find a solution to a problem, which is analogous to information seeking in response to a task or a problem. The hypothesis is that those with a higher score on this scale will spend more time for PIPs. Validity of the scale has been tested by Litman (2008).
   2.2. Acquisition subscale from the Saving Inventory-Revised scale (Frost et al., 2004): This is a subscale about excessive acquisition in a measure of compulsive hoarding. The assumption is that when an opportunity to pursue information is present as a form of new desire, those with an excessive acquisition characteristic would resist less. Hence, the relationship with this individual characteristic and information appetite

level is a research question. Validity of the SI-R scale was tested by Frost, Steketee, and Grisham (2004).

3. Time spent on responsibilities (work/domestic/family): Time spent working has been termed “actually working” in the questionnaire in order to differentiate it from time spent at work yet idly or doing things that are not related to work as it is possible to do PIP activities at work. Time spent on domestic/family responsibilities means time spent taking care of household affairs and/or family members. All three of these represent time spent for responsibilities that take away from one’s free time. Therefore, the hypothesis is that time spent for these responsibilities affect time spent for PIPs negatively.
4. Time spent on PIP activities: The previous chapter explained the PIPs compiled for the purpose of this study. These activities cover all gamut of PIP: consumption, dissemination, and creation of information for personal reasons. They are the following.
    4.1. Taking/studying for a class on a subject of personal interest: This is different from taking/studying for a class for academic reasons. One may be taking an online or offline class on a variety of topics of personal interest, ranging from hobby to personal development to self-defense. Learning new information and potentially sharing it with others are good examples of PIPs.
    4.2. Researching purchases: This involves mostly comparison shopping such as browsing through circulars, comparing prices at different stores (online

and/or offline), reading product reviews, and so on. This also involves learning new information and sharing it with others.

4.3. Computer use for personal interest (excluding games): All the other six activities can be done on a computer, or not on a computer, therefore this activity is asked separately. Because participants are asked not to treat these PIP activities mutually exclusively in terms of time use, if someone spent 30 minutes reading product reviews on the computer, the 30 minutes count for both Researching purchases and Computer use. Hence, this covers all varieties of PIPs performed on the computer.

4.4. Reading/doing research for personal interest: This includes reading print materials such as books, newspapers, and magazines, or reading materials online, as well as doing research for personal interest. It is a classic example of consuming information.

4.5. Writing/blogging for personal interest: This involves creating content which is one of the major groups of information behaviors (St. Jean et al., 2012), and is one of the three major activities in PIP (create, consume, and disseminate). Blogging is also creating and sharing information combined.

4.6. Listening/watching radio, podcast, TV, or videos for personal interest: This is consuming multimedia contents on TV, radio, computer, or mobile devices. Multimedia use is found to be one of the five major groups of information behaviors (St. Jean et al., 2012). For example, watching YouTube videos for instructions is consuming information and falls under this category.

4.7. Creating multimedia contents for personal interest: As mentioned before, content creation is one of the major groups of information behaviors (St. Jean et al., 2012) and a major part of PIP, and here is multimedia content creation separate from writing.

5. Topic interests (temporal context, time spent, optional open-ended question): Temporal context of topic interest means for how long one has been interested in a topic of personal interest. Temporal context and time spent on a topic of interest are collected in two sets. The optional open-ended question asks about topic interests and time use in the last 24 hours because three hours is a small window of time and may not capture topic interests activity well.

4.4 Plan for Data Analysis

The actual steps for data analysis will depend on the results obtained from each statistical analysis; the results would indicate the next appropriate test to employ. However, this method still assumes an overall view of what could be applicable after certain tests are run. The following represents a basic framework of key tests in the data analysis plan. It has been informed by the analysis performed on the ATUS dataset in the previous chapter. Statistical procedures are to be performed in SPSS and Statgraphics.

- Report information about the population characteristic such as the number of people in different time zones, gender, education level, and labor force status.

- Check descriptive statistics and frequency distributions of variables to assess the measures of central tendency, variability, and so on.
- Activity variables
    - Sum up the activity-based time use variables to create a PIP time use variable.
    - Run cluster analysis on the PIP time use variable to identify the high information appetite group.
    - Perform factor analysis to reduce the activity-based time use variables to a smaller number of dimensions.
    - Conduct hypothesis driven analyses using assessments of basic differences (e.g. t-test and chi-square) and multiple regression.
- Topic variables
    - Run tests that compare paired means (Wilcoxon Signed-Ranks test) to compare the means of time spent on topics.
    - Assess whether there are information appetite level and topic differences in time spent on topics (mixed ANOVA).
    - Conduct hypothesis driven analyses using multiple regression.
    - Analyze the open-ended question on how the participant spent time with information regarding their topics of interest in the last 24 hours.

Chapter 5 Study 2: Results and Discussion

This chapter reports the results and findings of the study presented in the previous chapter on Method. Discussed here are the results of the survey data gathering process, the demographic profile of the participants, and review of the findings using the propositions proposed in chapter 1 as well as the research questions and the hypotheses proposed in chapter 2. The propositions, research questions, and hypotheses are the following.

Proposition 1. Individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices (PIPs). PIPs encompass consumption, dissemination, and creation of information for personal reasons.

Proposition 2. An individual has different information appetites for different topical areas.

HP1. Information appetite in one type of activity does not predict information appetite in other types of activities.

HP2. Individual characteristics such as age, gender, and education level affect the time spent for PIPs.

HP3. Time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities.

RQ1. What are the relationships between the activity-based information appetite and the topical information appetite along with the topic's temporal context?

HP4. Those with a higher epistemic curiosity score will spend more time for PIPs.

RQ2. How do the participants' reported compulsive acquisition symptoms relate to the time they spent for PIPs?

5.1 Survey Data Gathering Process

The MTurk job with the Qualtrics survey link was posted on Amazon Mechanical Turk on Friday, January 29, 2016 at 6:35pm EST, and was live until 9:20pm EST for 591 participants.[10] (Nine participants were pretested to make sure the survey process runs smoothly, earlier in the same day from 5:28pm EST to 6:16pm EST.) The survey was filled out by 593 participants[11], resulting in 602 total participants in the dataset when the nine from the pretest were included. According to the MTurk system, the participants took 15 minutes 36 seconds on average to finish the job, resulting in $15.385 average hourly rate after their $4.00 compensation.

The pretest of nine participants had taken 48 minutes to complete. In the survey of 591 participants, as shown in the figure below, 99% of the survey was taken during the first hour and a half, and then the rate of completion slowed down dramatically. How exactly Amazon displays a job to a participant is not disclosed, but from this experience it can be surmised that Amazon releases a larger job to a larger number of people, and as the number of participants left for a job decreases, it is exposed to a smaller and smaller number of people.

[10] Given that the pretest of nine participants took 48 minutes to complete, at first it was assumed that the full study of 600 participants would take several days to complete. Hence Friday was chosen to include some weekdays and weekend days in the sample. The data collection on MTurk was completed much faster than anticipated, however. Moreover, as mentioned in the previous chapter, it was decided not to run the survey in chunks of smaller number of people due to technical difficulties. Hence the survey was conducted in the one-time cross-sectional method.

[11] This issue was discussed in the previous chapter.

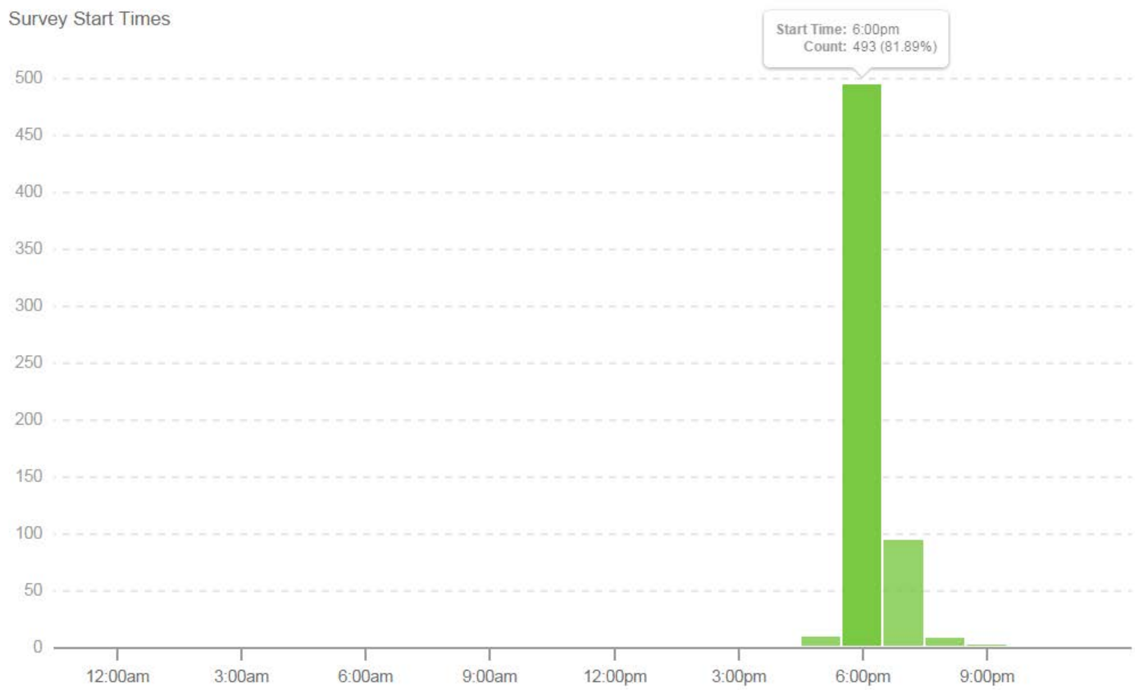


Figure 5.1 Qualtrics survey start times (y axis: number of participants)

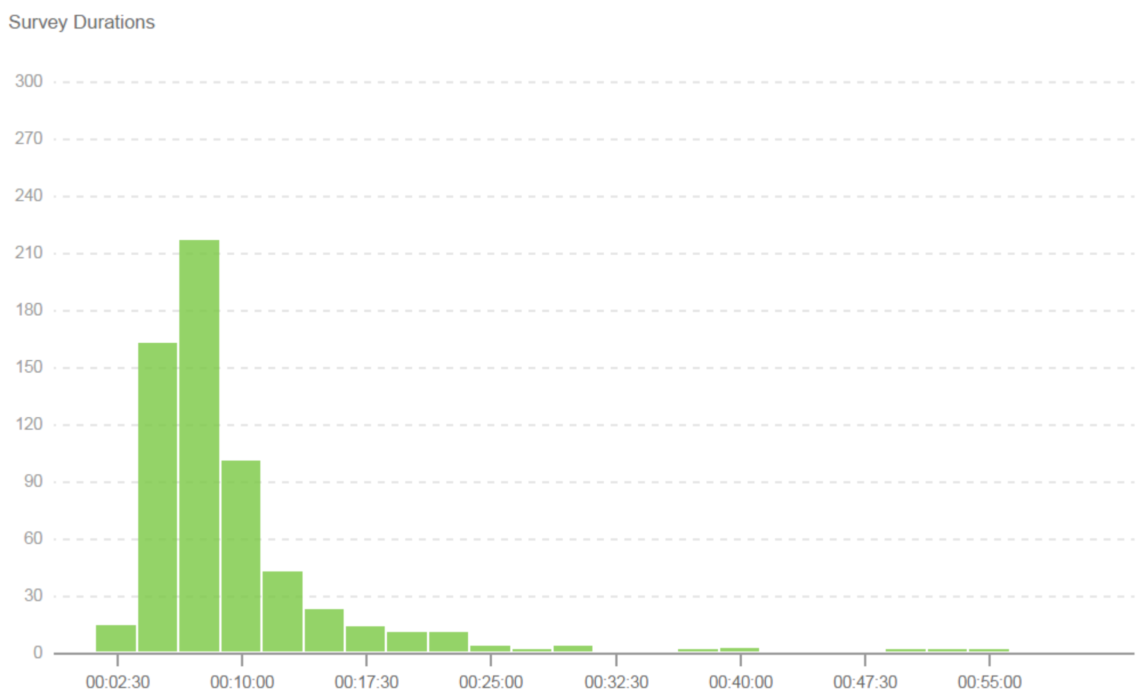


Figure 5.2 Qualtrics survey durations (y axis: number of participants)

According to the survey durations report from Qualtrics, the participants took from around 2:30 to 55 minutes to complete the survey, with 86% of them in the 5 to 12:30 minutes range.

Although some answers turned out to be invalid (most notably, not indicating a unit in the question “for how long you have been interested in the topic”), all respondents who completed the survey received payment for their participation.

5.2 Participant Profile

Below are the participant profiles based on their gender, age, time zone (geographic location), education level, and labor force status.

5.2.1 Gender

Gender	Frequency	Percent
Male	317	52.66%
Female	285	47.34%

Table 5.1 Frequency table of gender

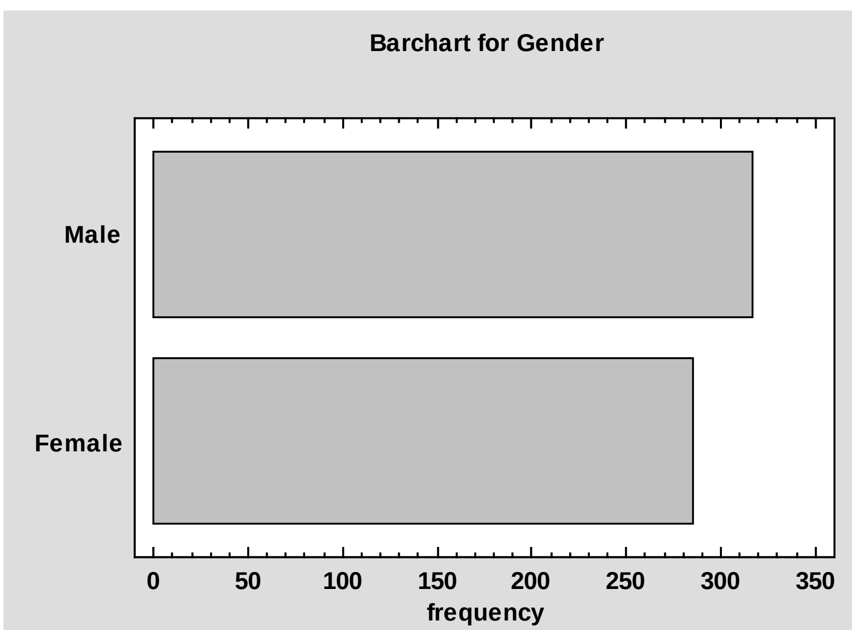


Figure 5.3 Bar chart for gender

5.2.2 Age

	Min	Max	Mean	Std. Dev.
Age	19	83	35.09	11.15

Table 5.2 Summary statistics for age

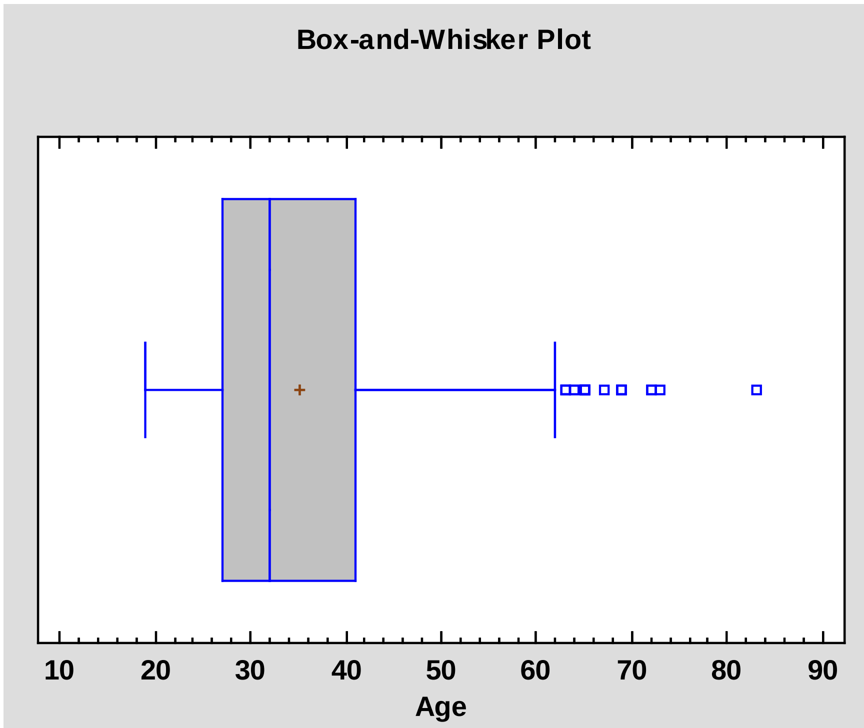


Figure 5.4 Box-and-whisker plot of age

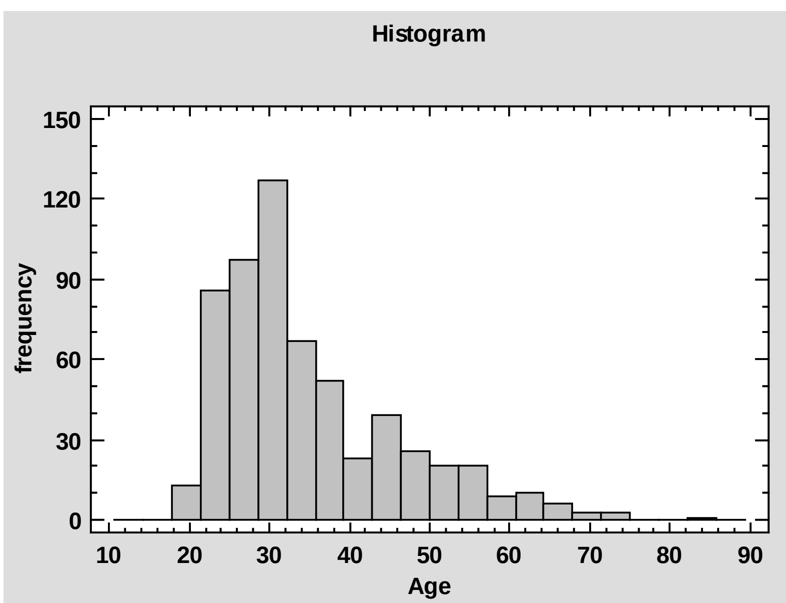


Figure 5.5 Frequency histogram of age

5.2.3 Time zone

Local time zone	Frequency	Percent
US/Eastern	302	50.17%
US/Central	133	22.09%
US/Mountain	29	4.82%
US/Pacific	138	22.92%

Table 5.3 Frequency table of time zone

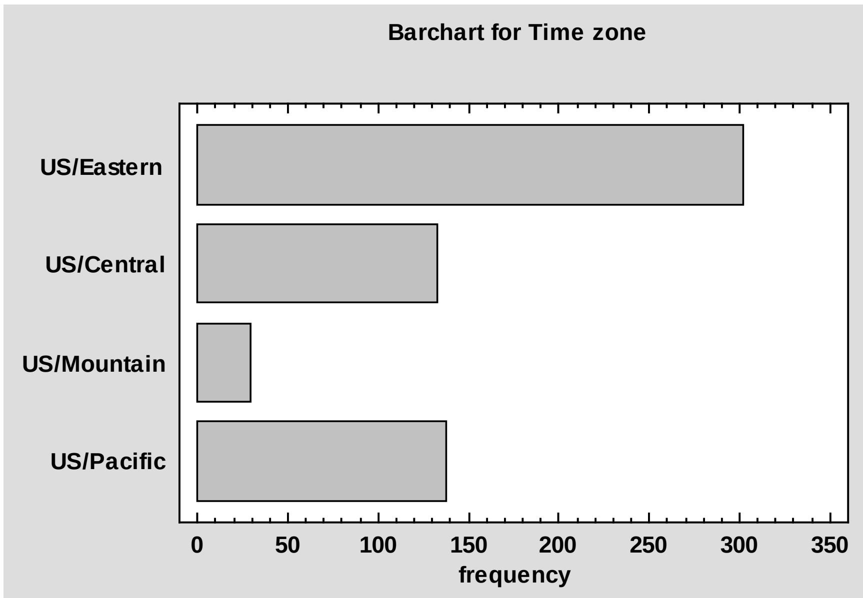


Figure 5.6 Bar chart for time zone

5.2.4 Education level

Highest level of school completed	Frequency	Percent
12th grade (no diploma) or below	2	0.33%
High school graduate - diploma or the equivalent (e.g. GED)	83	13.79%
Some college but no degree	154	25.58%
Associate degree in college - occupational/vocational program	34	5.65%
Associate degree in college - academic program	39	6.48%
Bachelor's degree (BA, AB, BS, etc.)	224	37.21%
Master's degree (MA, MS, MEng, MEd, MSW, etc.)	51	8.47%
Professional school degree (MD, DDS, DVM, etc.)	9	1.50%
Doctoral degree (PhD, EdD, etc.)	6	1.00%

Table 5.4 Frequency table of the highest level of school completed

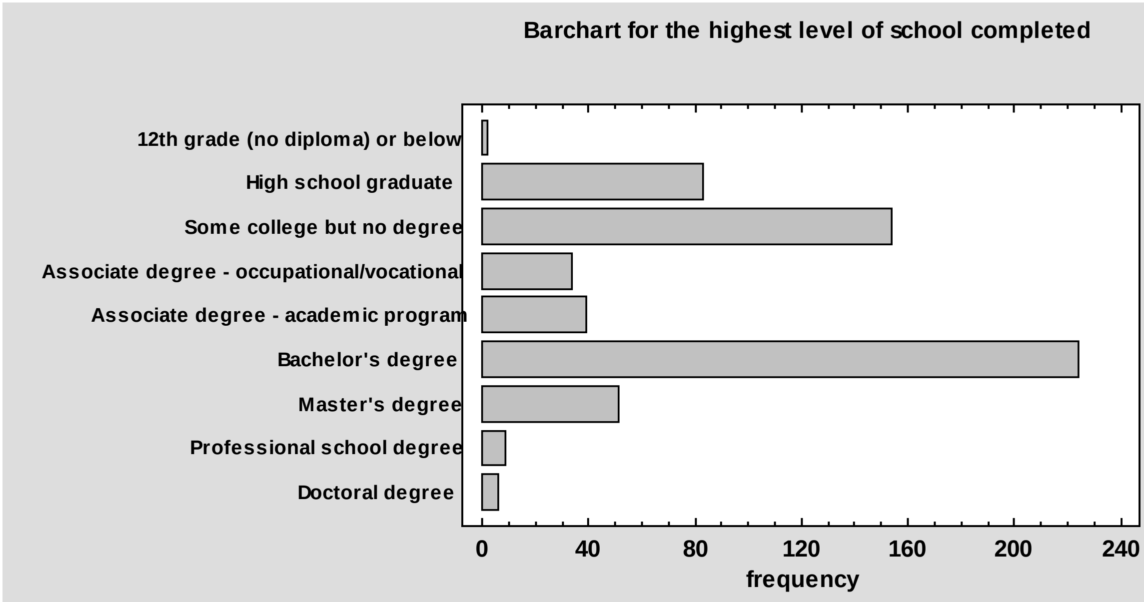


Figure 5.7 Bar chart for the highest level of school completed

5.2.5 Labor force status

Labor force status	Frequency	Percent
Employed - at work	440	73.09%
Employed - absent	35	5.81%
Unemployed - on layoff	4	0.66%
Unemployed - looking	59	9.80%
Not in labor force	64	10.63%

Table 5.5 Frequency table of labor force status

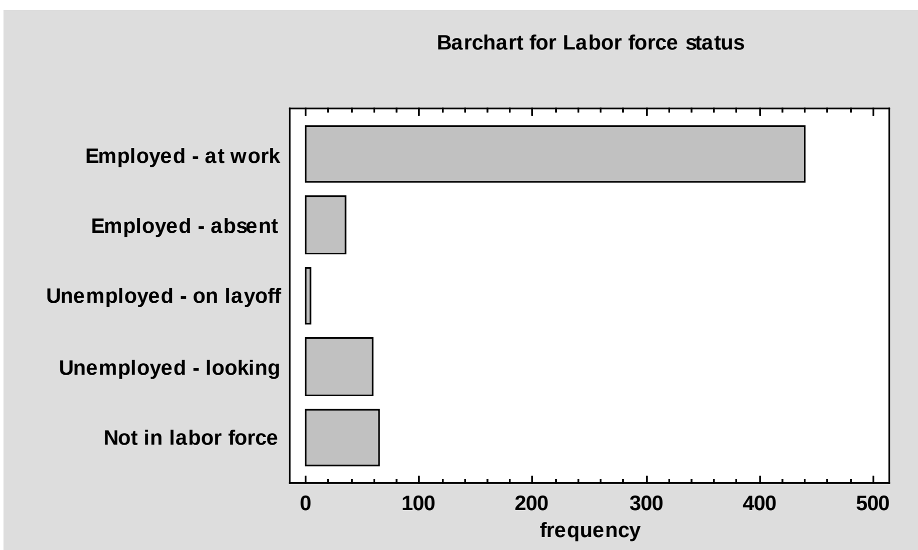


Figure 5.8 Bar chart for labor force status

5.2.6 Multiple job status

Here, NA (Not Applicable) refers to those who were unemployed or not in labor force because this question was asked to only those who were employed. The question was if the respondent has more than one job.

Multiple job status	Frequency	Percent
Yes	102	16.94%
No	373	61.96%
NA (Not Applicable)	127	21.10%

Table 5.6 Frequency table of multiple job status

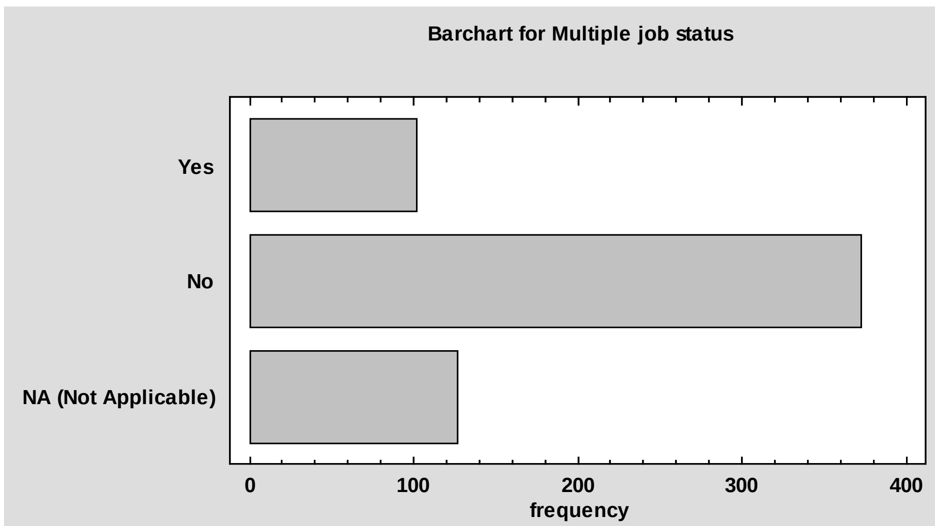


Figure 5.9 Bar chart for multiple job status

5.2.7 Summary

In summary, the participant demographics show the following:

- Gender: Male 52.7%, Female 47.3%
- Average age: 35 years old
- Local time zone: US/Eastern 50.2%, US/Central 22.1%, US/Mountain 4.8%, US/Pacific 22.9%
- About 75% have some college education or hold a college degree.
- About 79% are employed, and 17% hold more than one job.

5.3 Data Analyses

5.3.1 Activity-based information appetite

First, a PIP variable was created from the sum of time spent in all seven PIP activity categories: Taking/studying for a class on a subject of personal interest, Researching purchases, Computer use for personal interest (excluding games), Reading/doing research for personal interest, Writing/blogging for personal interest, Listening/watching radio, podcast, TV, or videos for personal interest, and Creating multimedia contents for personal interest. Then, another variable (PIP_count) was created for hypothesis 1 (information appetite in one type of activity does not predict information appetite in other types of activities), by counting the number of PIP activities that a participant engaged in (i.e. spent more than 1 minute on the activity).

Unlike in the ATUS dataset, however, analysis of the PIP_count variable as shown in the frequency table below suggests that hypothesis 1 is not supported. The mean is 3.49 with the standard deviation of 1.84. Although in the ATUS dataset 83% of the individuals engaged in one to two activities only, here the number of activities engaged in is much larger and more dispersed. Also, it is notable that 11.6% participated in all seven activities.

	Frequency	Percent	Cumulative Percent
0	14	2.3	2.3
1	54	9.0	11.3
2	131	21.8	33.1
3	149	24.8	57.8
4	107	17.8	75.6
5	43	7.1	82.7
6	34	5.6	88.4
7	70	11.6	100.0
Total	602	100.0	

Table 5.7 Frequency of the number of PIP activities engaged

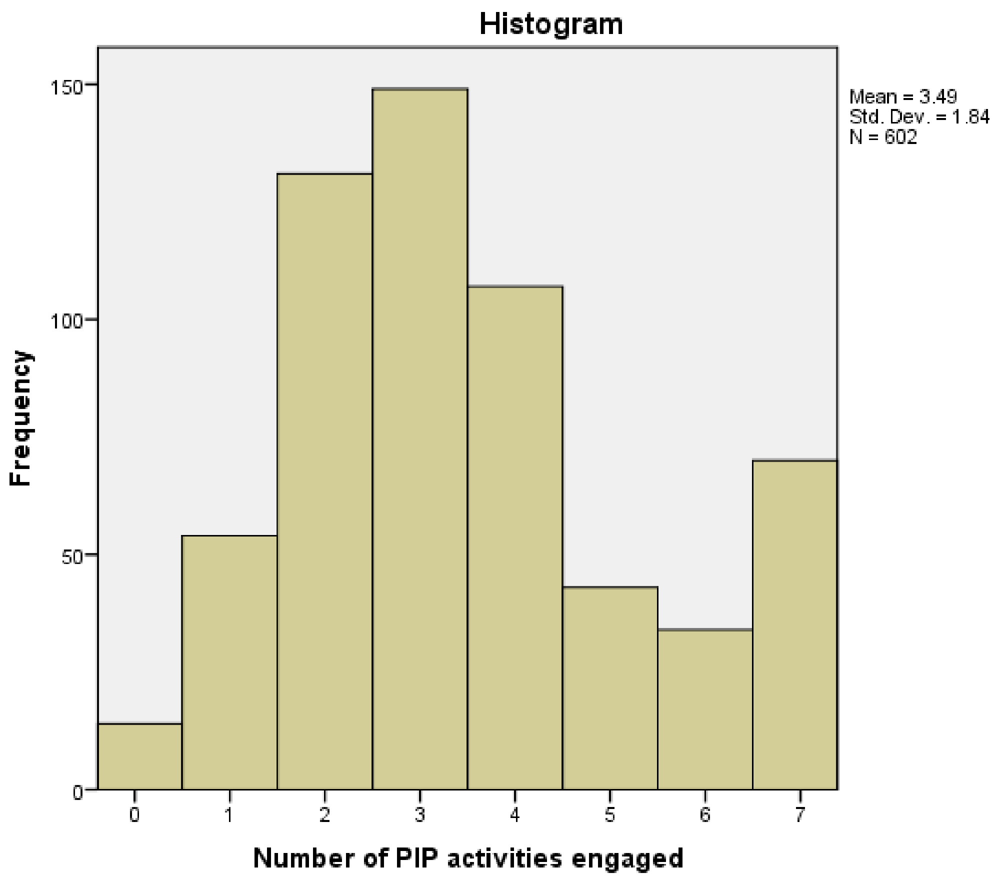


Figure 5.10 Frequency histogram of the number of PIP activities engaged

The table below shows participation rate of each of the PIP activities and descriptive statistics of the time spent. Overall, the participation rates are much higher than in the ATUS dataset (see table 2.2 for comparison). The highest participation rate is in the computer use for personal interest category, followed by reading/doing research and listening/watching multimedia contents. The highest participation rate category is perhaps not surprising because the participants were working on MTurk when they took the survey, which means they had already spent some time on the computer.

PIP Activity	Participation rate (%)[12]	Time spent by person (minutes)[13]				
		n	Min.	Max.	Average	Std. Dev.
Taking/studying for a class on a subject of personal interest	25.7%	155	1	180	25.24	35.485
Researching purchases	48.2%	290	1	162	18.92	21.548
Computer use for personal interest (excluding games)	90.9%	547	1	180	46.95	40.844
Reading/doing research for personal interest	70.1%	422	1	180	28.14	27.325
Writing/blogging for personal interest	23.3%	140	1	180	15.84	30.990
Listening/watching radio, podcast, TV, or videos for personal interest	69.9%	421	1	180	52.53	46.271
Creating multimedia contents for personal interest	20.8%	125	1	162	15.79	31.554
All seven activities combined	97.7%	588	1	842	124.60	114.694

Table 5.8 Descriptive statistics of the seven PIP activities

5.3.2 MTurk and ATUS datasets

As shown above, the MTurk and the ATUS datasets paint very different pictures. In this study, the statistical findings from the data collected on MTurk will not be compared to the findings from the ATUS dataset for the following reasons. First, in this study the ATUS dataset was not used to generalize findings, but instead as a way of exploring hypotheses and the first proposition. The ATUS dataset was examined to explain the presence of subgroup clusters (the high information appetite group vs. the regular group), to generate a hypothesis, and to explore some research hypotheses and proposition. It was not intended to provide benchmarks for comparable results between the MTurk and ATUS samples.

[12] The percentage of population that spent time on the activity. Participation rate percentage is calculated as Doers / All persons * 100 (Harvey, 1999, pp. 32–36).
[13] *n* = Those who participated in the activity (Total sample size: 602)

Second, the MTurk participants were already using a computer when they took the survey, which can explain the reason behind the large participation rate in the computer use activity category. This is a crucially different context than in the ATUS dataset. Third, the ATUS dataset collected data on one full day, but the MTurk study here collected data on three hours in the middle of a weekday. This is another contextual discrepancy. These contextual differences make it hard to compare the two datasets.

5.3.3 High information appetite group

Cluster analysis was performed to segregate the high information appetite (IA) group from the PIP variable (total time spent in PIPs). A Two-Step cluster analysis with the Euclidean distance measure resulted in two clusters as shown below.

Cluster	*N*	% Distribution	Min	Max	Mean	Std. Deviation	Label
1	575	95.5%	0	330	103.73	76.290	Regular
2	27	4.5%	355	842	504.30	132.352	High IA
Combined	602	100%	0	842	121.70	114.898	

Table 5.9 Cluster analysis on total time spent in PIPs

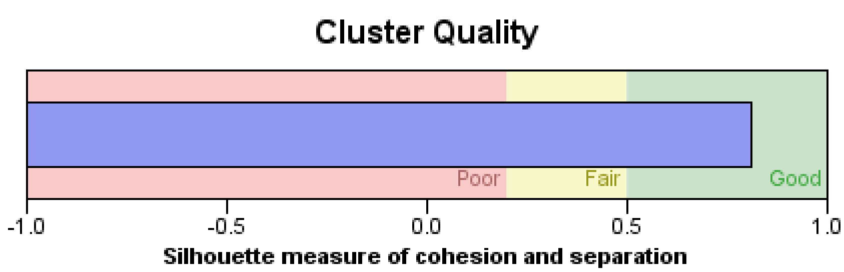


Figure 5.11 Cluster quality measure (Average silhouette = 0.8)

Since the size of cluster 2 was too small to conduct further statistical analyses, it was decided to find an alternative approach to define the high IA group. This time, the high IA group was defined as more than one standard deviation away from the mean (total time spent in PIPs' value greater than 121.7 + 114.898 = 236.598).

IA Group	*N*	% Distribution	Min	Max	Mean	Std. Deviation	Label
1	536	89%	0	236	91.08	61.849	Regular
2	66	11%	237	842	370.33	141.702	High IA
Combined	602	100%	0	842	121.70	114.898	

Table 5.10 IA groups redefined, based on total time spent in PIPs

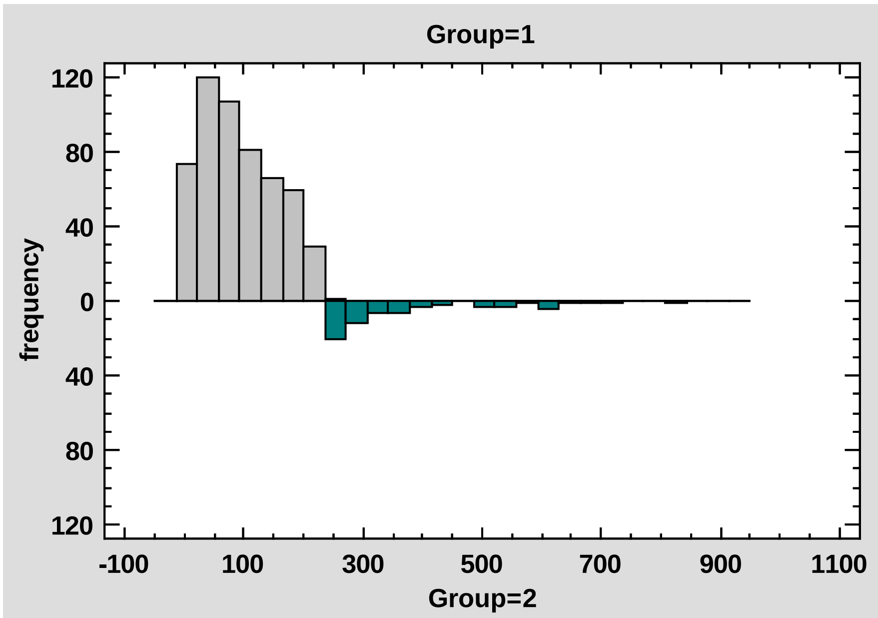


Figure 5.12 Frequency histogram of the IA groups (x axis: total time spent in PIPs)

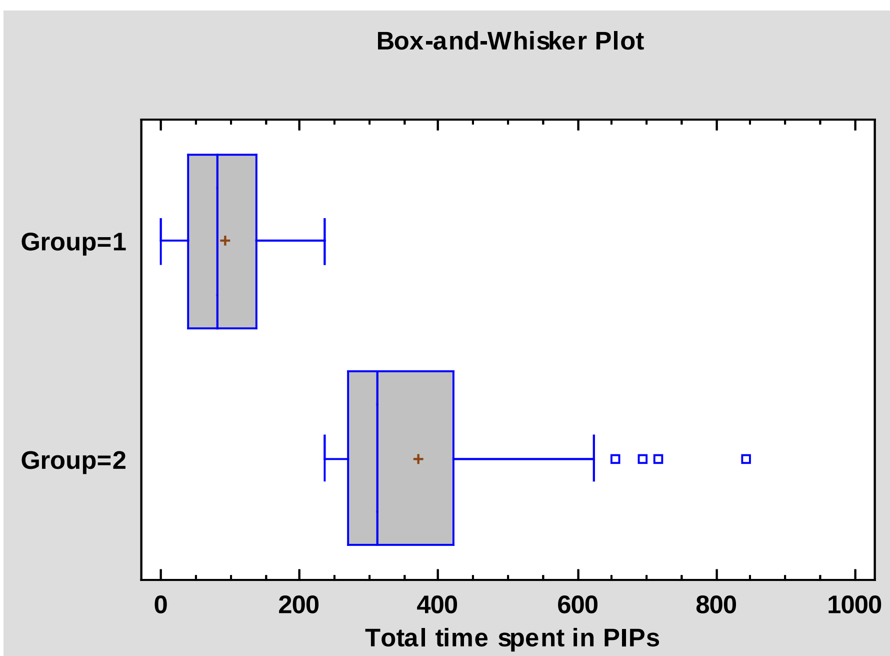


Figure 5.13 Box-and-whisker plot of the IA groups (total time spent in PIPs)

The Pearson chi-square result shows that there was not a significant gender difference between the two groups (*p* = .949).

	Group 1 (Regular)	Group 2 (High IA)	Total	x^2	*p*	Cramer's V
Gender				.004	.949	.003
Male	52.6%	53.0%	52.7%			
Female	47.4%	47.0%	47.3%			
Total	100%	100%	100%			

Table 5.11 Chi-square analysis of gender between the IA groups

The table below shows that the high IA group is different from the regular group in terms of age (*t*(600) = 2.339, *p* = .020, *d* = .31), which is statistically significant. Inspection of the two group means indicates that the average age (*M* = 32.08) of the high IA group is younger than the age (*M* = 35.47) of the regular group. The effect size (*d* = .31) is on the low side, however.

	Mean	Std. Deviation	*t*	*df*	*p*	*d*
Age			2.339	600	.020	.31
Group 1 (Regular)	35.47	11.201				
Group 2 (High IA)	32.08	10.347				

Table 5.12 T-test result of age differences in the IA groups

The education level was not significantly different between the groups (*p* = .410).

	Group 1 (Regular)	Group 2 (High IA)	Total	x^2	*p*	Cramer's V
Highest level of school completed				8.249	.410	.117
12th grade (no diploma) or below[14]	0.2%	1.5%	0.3%			
High school graduate - diploma or the equivalent (e.g. GED)[15]	14.6%	7.6%	13.8%			

[14] This is recoded from the sum of eight original labels in the variable (Less than 1st grade; 1st, 2nd, 3rd, or 4th grade; and through 12th grade - no diploma).

[15] This is recoded from the sum of two original labels in the variable (GED and diploma).

Some college but no degree	25.4%	27.3%	25.6%			
Associate degree in college - occupational/vocational program	5.6%	6.1%	5.6%			
Associate degree in college - academic program	6.7%	4.5%	6.5%			
Bachelor's degree (BA, AB, BS, etc.)	36.6%	42.4%	37.2%			
Master's degree (MA, MS, MEng, MEd, MSW, etc.)	8.6%	7.6%	8.5%			
Professional school degree (MD, DDS, DVM, etc.)	1.3%	3.0%	1.5%			
Doctoral degree (PhD, EdD, etc.)	1.1%	-	1.0%			
Total	100%	100%	100%			

Table 5.13 Chi-square analysis of the education level between the IA groups

There was a statistically significant difference in the labor force status between the two IA groups (x^2 = 10.538, *df* = 4, *N* = 602, *p* = .032). The high IA group had fewer people employed and more people unemployed. There was not a significant difference in multiple job status, however (*p* = .212).

	Group 1 (Regular)	Group 2 (High IA)	Total	x^2	*p*	Cramer's V
Labor force status				10.538	.032	.132
Employed – at work	73.5%	69.7%	73.1%			
Employed – absent	6.3%	1.5%	5.8%			
Unemployed – on layoff	0.7%	-	0.7%			
Unemployed – looking	8.6%	19.7%	9.8%			
Not in labor force	10.8%	9.1%	10.6%			
Total	100%	100%	100%			

Table 5.14 Chi-square analysis of the labor force status between the IA groups

	Group 1 (Regular)	Group 2 (High IA)	Total	x^2	*p*	Cramer's V
Has more than one job				3.100	.212	.072
Yes	16.8%	18.2%	16.9%			
No	63.1%	53.0%	62.0%			
Not applicable	20.1%	28.8%	21.1%			
Total	100%	100%	100%			

Table 5.15 Chi-square analysis of the multiple job status between the IA groups

As for the time spent actually working and the time spent for domestic/family responsibilities, there was a marginally significant difference in the former (p = .077) with small effect size (d = .23) but no significant difference in the latter (p = .198). The mean difference in the time spent actually working indicates that the high IA group spent less time working (M = 61.21) than the regular group did (M = 75.13).

	Mean	Std. Deviation	*t*	*df*	*p*	*d*
Time spent actually working			1.771	600	.077	.23
Group 1 (Regular)	75.13	60.477				
Group 2 (High IA)	61.21	58.432				
Time spent on domestic/family responsibilities			-1.290	600	.198	.16
Group 1 (Regular)	41.08	42.042				
Group 2 (High IA)	48.20	44.213				

Table 5.16 T-test result of time spent on responsibilities between the IA groups

Also investigated were the epistemic curiosity (EC) scale scores and the acquisition scale scores. Both Interest-type and Deprivation-type EC scores were found significantly different between the IA groups (p = 0.41 and p = .001 respectively), but there was no significant difference in the acquisition scores (p = .173). Both EC scores were higher in the high IA group.

	Mean	Std. Deviation	t	df	p	d
Interest-type EC			-2.047	600	.041	.27
Group 1 (Regular)	15.09	3.189				
Group 2 (High IA)	15.94	3.127				
Deprivation-type EC			-3.434	600	.001	.46
Group 1 (Regular)	11.76	3.397				
Group 2 (High IA)	13.27	3.218				
Acquisition scale			-1.363	600	.173	.18
Group 1 (Regular)	8.13	3.999				
Group 2 (High IA)	8.85	4.159				

Table 5.17 T-test result of the epistemic curiosity and the acquisition scale scores between the IA groups

5.3.4 Variable relationships to the time spent for PIPs

Additional inferential statistics procedures were performed to find the strength of associations between variables. A multiple regression analysis was conducted with the time spent for PIPs as a dependent variable to find out the relationships between personal and situational independent variables and the time spent for PIPs.

The independent variables in the final multiple regression model were the following.

- Gender (1 = male, 0 = female)
- Age
- Labor force status (Reference category: Not in labor force)
- Has more than one job (1 = yes, 0 = no/NA)
- Interest-type epistemic curiosity (EC) scale score
- Deprivation-type EC scale score
- Acquisition scale score

Time spent actually working and time spent for domestic/family responsibilities have been removed from the model because the dependent variable (time spent for PIPs) could potentially affect them (e.g. spending time on PIPs at work instead of actually working, or spending too much time reading instead of cleaning the house.) According to Allison (1999), if the dependent variable affects the independent variable it might result in a bias of every coefficient in the regression model.

In addition, highest level of school completed was removed from the model because it accounted for less than 1% of the variance (.001 in adjusted R^2) and the tolerance values in collinearity statistics were very low in multiple categories.

The combination of the independent variables listed above significantly predicted time spent for PIPs, $F(10, 591) = 4.061$, $p < .001$, with three of the variables significantly contributing to the prediction. The adjusted R^2 value was .048, indicating that 4.8% of the variance in time spent for PIPs was explained by this model. The beta weights suggest that being employed at work contributes the most to predicting time spent for PIPs in the negative direction, and that being younger and having the tendency to acquire things also contribute to this prediction.

Variable	B	Std. Error of B	beta	Squared part correlation
(Constant)	132.357	33.370		
Gender (1=male)	14.476	9.555	.063	.004
Age	-1.084	.433	-.105**	.010
Labor force status: Employed – at work	-28.969	15.671	-.112*	.005
Labor force status: Employed – absent	-16.876	23.894	-.034	.001

Labor force status: Unemployed – on layoff	-8.449	57.894	-.006	.00003
Labor force status: Unemployed – looking	29.794	20.561	.077	.003
Has more than one job (1=yes)	-2.647	12.652	-.009	.0001
Interest-Type EC scale score	-.861	1.742	-.024	.0004
Deprivation-Type EC scale score	2.479	1.665	.074	.004
Acquisition scale score	2.801	1.191	.098**	.009

Note. Adjusted $R^2 = .048$; $F(10, 591) = 4.061$, $p < .001$
*$p < .10$, **$p < .05$

Table 5.18 Multiple regression analysis summary predicting time spent for PIPs ($N = 602$)

The squared part correlations represent the strength of the unique predictive contribution of each particular variable (Warner, 2013). It shows that age uniquely predicted about 1% of the variance in time spent for PIPs, when all the other variables were statistically controlled. The rest of the variables predicted less than 1% of the variance in time spent for PIPs.

Although a normal distribution is not strictly required in multiple regression (Allison, 1999), to see if transforming the y variable (PIP) would yield a different result, the PIP variable was transformed into a square root of it (SQRT(PIP) in SPSS). The square root was chosen because it is a good transformation method for positively skewed data. Another good method, a log transformation, was not selected because a log of zero results in an undefined value, and yet zero is a meaningful value in total time spent for PIPs. A square root of zero is zero.

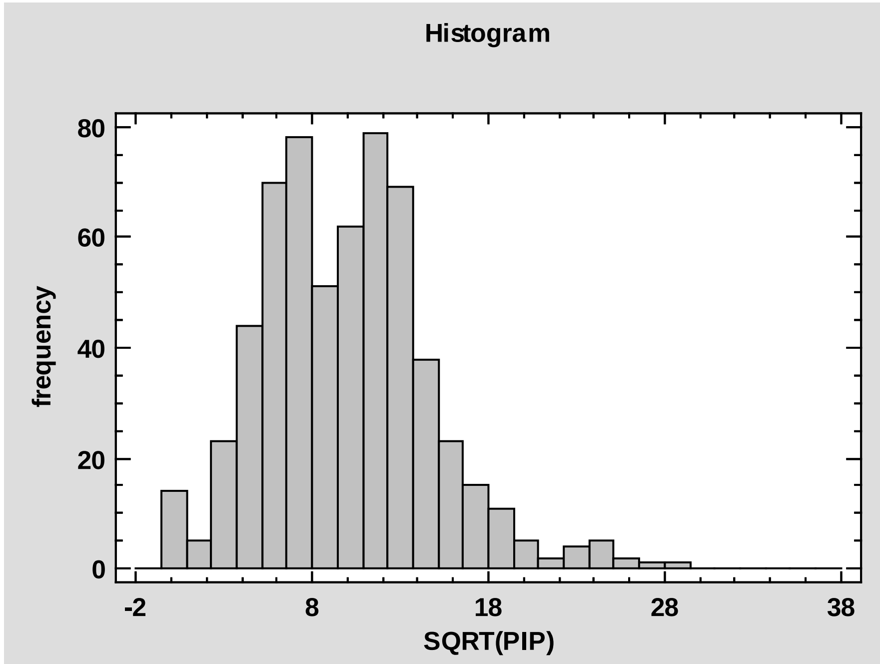


Figure 5.14 Frequency histogram of the square root of total time spent in PIPs

Then, multiple regression was conducted with the same set of independent variables to determine the best linear combination of the variables for predicting time spent for PIPs. The combination of independent variables significantly predicted time spent for PIPs, $F(10, 591) = 4.889$, $p < .001$, with four of the variables significantly contributing to the prediction. The adjusted R^2 value was .061, indicating that 6.1% of the variance in time spent for PIPs was explained by this model. The beta weights suggest that being employed at work contributes the most to predicting time spent for PIPs in the negative direction, and that being male, being younger, and having the tendency to acquire things also contribute to this prediction.

Variable	B	Std. Error of B	beta	Squared part correlation
(Constant)	10.828	1.380		
Gender (1=male)	.925	.395	.097*	.009
Age	-.037	.018	-.087*	.007

Labor force status: Employed – at work	-1.812	.648	-.168*	.012
Labor force status: Employed – absent	-.623	.988	-.031	.001
Labor force status: Unemployed – on layoff	-.462	2.394	-.008	.0001
Labor force status: Unemployed – looking	.895	.850	.056	.002
Has more than one job (1=yes)	-.281	.523	-.022	.0005
Interest-Type EC scale score	-.028	.072	-.019	.0002
Deprivation-Type EC scale score	.052	.069	.037	.001
Acquisition scale score	.130	.049	.109*	.011

Note. Adjusted R^2 = .061; $F(10, 591) = 4.889$, $p < .001$
*$p < .05$

Table 5.19 Multiple regression analysis summary predicting time spent for PIPs (square root) (N = 602)

The squared part correlations show that being employed at work uniquely predicted about 1.2% of the variance in time spent for PIPs, when all the other variables were statistically controlled. Having the tendency to acquire things predicted about 1.1% of the variance. The rest of the variables predicted less than 1% of the variance in time spent for PIPs.

5.3.5 Factor analysis

Factor analysis was performed to find out which of the seven PIP time use variables exhibit similar patterns. The seven PIP time use variables are Taking/studying for a class on a subject of personal interest, Researching purchases, Computer use for personal interest (excluding games), Reading/doing research for personal interest, Writing/blogging for personal interest, Listening/watching radio, podcast, TV, or videos for personal interest, and Creating multimedia contents for personal interest.

Here, principal axis factor analysis (PA) method was used to extract relationships because it is good for understanding the covariation among variables (not reproducing all information including variance and covariance as principal component analysis does) (Leech et al., 2015). PA with varimax rotation was conducted, and two components were rotated based on the eigenvalues over one criterion. After rotation, the two components combined accounted for 43% of the variance. Component 1 accounted for 25% of the variance, and component 2 accounted for 18%. The table below displays the items and component loadings for the rotated components.

	Factor	
	1	2
Writing/blogging for personal interest	**.825**	.066
Creating multimedia contents for personal interest	**.693**	.247
Taking/studying for a class on a subject of personal interest	**.507**	.371
Researching purchases	**.474**	.262
Reading/doing research for personal interest	.242	**.793**
Computer use for personal interest (excluding games)	.166	**.489**
Listening/watching radio, podcast, TV, or videos for personal interest	.080	**.347**

Table 5.20 Factor loadings from principal axis factor analysis with varimax rotation for a two factor solution for PIP time use categories (N = 602)

In factor 1, it is notable that creating information (writing/blogging and creating multimedia contents) load together. Taking/studying for a class and researching purchases also load in factor 1. In factor 2, consuming information (reading/research and listening/watching multimedia contents) and computer use load together. This indicates that most of the consumption might have been done on the computer.

5.3.6 Topic-based information appetite

To test proposition 2 that an individual has different information appetites for different topical areas, a nonparametric Wilcoxon Signed-Ranks test was conducted. It was chosen in order to compare paired means for continuous data that are not normally distributed. In this case, the pairs were time spent for topic A and time spent for topic B. As with time spent in PIPs, time spent for topics also exhibit positively skewed, long tail distributions, so Wilcoxon Signed-Ranks test was appropriate.

The Wilcoxon Signed-Ranks test showed that 329 people spent more time on topic A, 175 spent more time on topic B, and there were 98 ties (N = 602). The difference indicating more time spent for topic A was statistically significant (Z = -7.032, p < .001, r = -0.20).

	N	Mean	Std. Deviation	Minimum	Maximum	Median
Time spent on topic A	602	37.47	36.311	0	180	25.00
Time spent on topic B	602	28.10	28.796	0	180	20.00

Table 5.21 Descriptive statistics of the time spent on topics

Then, the SPSS file was split according to the IA group (high IA vs. regular), and the same statistical test was run again.

In the regular IA group, the Wilcoxon Signed-Ranks test showed that 287 people spent more time on topic A, 155 spent more time on topic B, and there were 94 ties (N = 536). The difference indicating more time spent for topic A was statistically significant (Z = -6.481, $p < .001$, r = -0.20).

	N	Mean	Std. Deviation	Minimum	Maximum	Median
Time spent on topic A	536	33.25	32.729	0	180	22.00
Time spent on topic B	536	24.96	25.660	0	180	16.00

Table 5.22 Descriptive statistics of the time spent on topics in the regular IA group

In the high IA group, the Wilcoxon Signed-Ranks test indicated that 42 people spent more time on topic A, 20 spent more time on topic B, and there were 4 ties (N = 66). The statistics show that difference indicating more time spent for topic A was statistically significant (Z = -2.483, p = .013, r = -0.22).

	N	Mean	Std. Deviation	Minimum	Maximum	Median
Time spent on topic A	66	71.77	45.083	5	180	60.00
Time spent on topic B	66	53.64	38.777	5	180	46.50

Table 5.23 Descriptive statistics of the time spent on topics in the high IA group

Then, a mixed ANOVA was conducted to assess whether there were IA level and topic differences in time spent on topics. Results indicated a statistically significant main effect of topic, $F(1, 600) = 36.173$, $p < .001$, partial eta^2 = .057. In other words, the difference on time spent on topics between topic A and B was statistically significant. IA group effect was statistically significant also at $F(1, 600) = 97.712$, $p < .001$, partial eta^2 = .140. This means that the test of whether people in different IA groups spent time on topics differently is significant. However, the topic main effect was qualified by a statistically significant interaction between topic and IA group, $F(1, 600) = 5.023$, $p = .025$, partial eta^2 = .008. Examining the means and standard deviations of the time spent on topics between regular and high IA groups shows that both regular and high IA groups spent more time on topic A, and high IA group spent more time on both topics A and B than the regular group did.

Next, multiple regressions were conducted as the time spent on topic A and topic B as dependent variables, respectively. The independent variables were the same as in the multiple regression model for the time spent for PIPs, except the addition of time interested in the topic. They were the following.

- Gender (1 = male, 0 = female)
- Age
- Labor force status (Reference category: Not in labor force)
- Has more than one job (1 = yes, 0 = no/NA)
- Interest-type epistemic curiosity (EC) scale score
- Deprivation-type EC scale score
- Acquisition scale score
- Time interested in the topic (in days)

N was 553 for topic A, reduced from the total *N* of 602, because there were some missing values in time interested in topic A. The model's prediction of time spent on topic A was marginally significant, $F(11, 541) = 1.671$, $p = .077$, with four of the variables significantly contributing to the prediction. The adjusted R^2 value was .013, indicating that 1.3% of the variance in time spent on topic A was explained by this model.

The statistics show that the time interested in the topic A did not statistically significantly predict time spent on topic A.

Variable	B	Std. Error of B	beta	Squared part correlation
(Constant)	59.116	10.929		
Gender (1=male)	3.663	3.107	.052	.002
Age	-.097	.144	-.031	.001
Labor force status: Employed – at work	-14.266	5.117	-.182***	.014
Labor force status: Employed – absent	-13.189	7.638	-.091*	.005
Labor force status: Unemployed – on layoff	-18.746	18.050	-.045	.002
Labor force status: Unemployed – looking	-6.897	6.680	-.059	.002
Has more than one job (1=yes)	9.849	4.164	.104**	.010
Interest-Type EC scale score	-1.133	.564	-.102**	.007
Deprivation-Type EC scale score	.568	.536	.055	.002
Acquisition scale score	-.024	.391	-.003	.00001
Time interested in topic A	4.685E-05	.000	.005	.00002

Note. Adjusted R^2 = .013; $F(11, 541) = 1.671$, $p = .077$
*$p < .10$, **$p < .05$, ***$p < .01$

Table 5.24 Multiple regression analysis summary predicting time spent on topic A (N = 553)

For topic B, N was 551, reduced from the total N of 602, because of the missing values in time interested in the topic B. The independent variables did not predict the time spent on topic B, however, $F(11, 539) = 1.460$, $p = .143$. The statistics showed that the time interested in the topic B statistically significantly predicted time spent on topic B (beta = .097), but since the overall model is not significant, this has little meaning.

5.3.7 Topics and time with information in the last 24 hours

Lastly, the open-ended question on how the participant spent time with information regarding their topics of interest in the last 24 hours was analyzed. This question was optional and was completed by 301 out of 602 participants.

Largely four themes emerged from the data. First was the availability of time, or the lack there of. Second was the temporal context, or the length of time they have been interested in the topic. Third was why they were interested in a certain topic, and the fourth was doing multiple PIP activities at the same time.

First, the participants noted that they did not spend much time for PIPs in the past three hours because they were at work or attending to domestic/family responsibilities. Several mentioned that they spend more time for PIPs in the evenings (e.g. after work, or after kids have gone to bed) or weekends. Several also wished they had more free time for PIPs. Some mentioned that they spent time on PIPs while at work, often during breaks or down time.

Second, the temporal context turned out to be a complex one. Many times it was nested at multiple levels – such as being interested in politics for many years, then the upcoming presidential election, and caucuses. Interests in caucuses were short-termed, but it was nested within a medium-term interest in the presidential election, and long-term interests in politics and news. Some had a long-term interest in favorite TV shows, and then a short-term burst of interest in a new cast member. Then there were long-term interests such as cooking, and a short-term interest of a recipe for the night.

The data showed that there were two modes of attention for long-term, ongoing interests: One is that it may stay dormant in the back of the mind until a related short-term interest pops up, which then takes the center stage over other interests. For instance, one participant had a 30 year interest in gardening, but warmer weather on the day of the survey made her want to take action and research plants. Another participant was a lifelong gamer, and finding out about a new game prompted her to research it to see if she wanted to purchase it. A long-term movie goer read reviews of new movies because of a plan to go to the movie with friends on the weekend. The other is that one continuously and regularly (almost every day) checks for new information on the long-term interested topic. Being up to date on the topic is important to these participants.

Third, the participants described why they were interested in a certain topic. The most prominent reason was research for practical purposes such as income tax, travel planning, redecorating, fixing a computer, and so on. Some of the other major reasons included researching purchases (e.g. car, cell phone, small appliances, etc.) and the upcoming presidential election. The next was a major life event for themselves or close ones such as moving, pregnancy, or having a serious illness such as cancer. Personal goals such as losing weight, writing a book, or a career goal were also mentioned.

Fourth, some participants mentioned that they did multiple PIP activities together such as researching information on a TV show on a tablet while watching the TV show, listening to a radio episode and looking up a print copy of the story at the same time, or having YouTube running in the background while being online.

5.4 Discussion

The first proposition was that individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices (PIPs). The distribution of the total time spent for PIPs was positively skewed with a long tail in that there were many people who did not spend much time for PIPs, but at the same time there were a small number of people who spent a lot of time in PIPs. The distributions of times spent on topics also exhibited a similar shape. In addition, a cluster analysis identified a small group of people who spent much time for PIPs, namely a high information appetite (IA) group. The number was very small (4.5% of the sample), though, so throughout the data analyses mean plus one standard deviation was used to distinguish the high IA group (11% of the sample). The proposition was supported in that individuals spent varying degrees of time in PIPs, reflecting their IA level.

The second proposition was that an individual has different information appetites for different topical areas. This was supported by statistical analyses of the time spent on topics. Respondents were asked to provide two different topics that they were interested in the last three hours, topic A and topic B, and then asked how much time they spent on each topic. According to the Wilcoxon Signed-Ranks test, 504 out of 602 respondents (84%) spent either more time on topic A or more time on topic B (i.e. not a tie between the two). The difference on time spent on topics between topic A and B was statistically significant in mixed ANOVA. Based on these results, proposition two was also supported. Additionally, more people spent more time on topic A, which happened even though the participants were instructed to name the topics in no particular order.

Hypothesis 1, information appetite in one type of activity does not predict information appetite in other types of activities, was not supported by the MTurk dataset. Although in the ATUS dataset 83% of the individuals engaged in one to two activities only, in the MTurk dataset the number of activities engaged in was much larger and more dispersed (Mean = 3.49, Std. Dev. = 1.84). 64.4% were involved in two to four activities. Moreover, 11.6% of the respondents participated in all seven activities. One thing to note is that this difference between the ATUS and the MTurk results might be due to the fact that in the MTurk survey PIP activities were not treated as mutually exclusive, so the participants could report more activities than one for a given time period. Moreover, the MTurk sample was using a computer when they answered the survey, and 90.9% of the participants said they used a computer for PIPs. So the PIP activities they performed likely happened on the computer, and the participants must have counted computer use in addition to their other PIP activities in the survey answer, which might be the reason of 64.4% being involved in two to four activities.

Hypothesis 2, individual characteristics such as age, gender, and education level affect the time spent for PIPs, was partially supported. According to comparisons of regular and high IA groups, age difference was found to be statistically significant, but gender and education level differences were not. As for the age, the high IA group was younger (Mean = 32.08, Std. Dev. = 10.35) than the regular group (Mean = 35.47, Std. Dev. = 11.20). Multiple regression analysis also showed that being younger contributes to the prediction of time spent for PIPs.

Hypothesis 3, time spent for PIPs is negatively affected by time spent for work, domestic, and family responsibilities, was also partially supported. When the regular and high IA groups were compared, the difference in time spent actually working was found marginally significant in that the high IA group had spent less time actually working. However, the difference in time spent on domestic/family responsibilities was not found to be statistically significant.

These two variables were not included in the multiple regression analysis because time spent actually working and time spent for domestic/family responsibilities could be affected by the time spent for PIPs (e.g. spending time on PIPs at work instead of actually working, or spending too much time reading instead of cleaning the house). If the dependent variable affects the independent variable, it might create a bias in the regression model (Allison, 1999).

Research question 1 was: what are the relationships between the activity-based information appetite and the topical information appetite along with the topic's temporal context? This was investigated with mixed ANOVA analysis to assess whether there were IA level (determined from the activity-based IA) and topic differences in time spent on topics. The results showed that the difference on time spent on topics between topic A and B was statistically significant, and the test of whether people in different IA groups spent time on topics differently was also significant. Both regular and high IA groups spent more time on topic A[16], and the high IA group spent more time on both topics A and B than the regular group did.

As for the topic's temporal context, statistical analyses did not reveal any useful results and it could not be determined whether a topic's temporal context affects time spent on topics. However, analysis of the open-ended question on time spent on

[16] This must be an order effect, although the participants were instructed to list the topics "in no particular order." The full survey questionnaire is in Appendix B.

interested topics in the last 24 hours indicated a complex relationship between the topics and the temporal context. A nested structure was found where a short-term interest stems from a medium-term interest, which in turn is nested inside a long-term interest. As such, when answering the survey some participants responded on time spent on their current short-term interest along with how long they have been interested in their long-term interest. This could have thrown off the survey results.

Hypothesis 4, those with a higher EC score will spend more time for PIPs, was supported by the t-test between regular and high IA groups, but was not supported by the multiple regression analysis. According to the t-test, both EC scores were higher in the high IA group (the group that spent more time in PIPs). However, in the multiple regression model, the relationship between the EC variables and the time spent in PIPs was not found to be statistically significant.

RQ2 was how the participants' reported compulsive acquisition symptoms related to the time they spent for PIPs. The relationship between the acquisition scale scores and the time spent for PIPs was found to be statistically significant in the multiple regression analysis, and according to the beta weight a higher scale score (i.e. having the tendency to acquire things) related to more time spent in PIPs.

Overall, the multiple regression model of time spent for PIPs had a very small effect size of 4.8%. There was a similar result with the ATUS dataset in its high IA group. This suggests that some other factors that were not investigated in this study might be affecting results. The next chapter proposes plans for future research to further study these undiscovered aspects of information appetite.

Chapter 6 Conclusion

In this chapter, implications of the findings and limitations in the studies reported in this dissertation as well as plans for future research are discussed. It ends with concluding remarks.

6.1 Variables and the Revised Model

What approaches or perspectives might explain information appetite (IA)? The answer can be considered in terms of person and situation variables. First would be personal characteristics embedded within an individual such as personality, which acts to produce a particular information behavior that is consistent across different situations and is stable over time. However, personality variables such as epistemic curiosity were found to have minimal effect in the MTurk study. The results did not support that an individual's IA level would be universal across all situations and be persistent over time. Rather, it is driven by opportunity such as free time available for personal information practices (PIPs), by need-driven interests, and so on.

Second, IA might be primarily explained by situation variables alone. A situation variable was found important in that free time available for PIPs was a powerful predictor of time spent for PIPs. Yet, there is no reason to assume that the situational circumstances can be generalized over individuals and then over time. For example, some situations generate short-lived interests: needing a car may invoke a high IA in a car-related topic but for only a period of time while the necessity exists. Needing to fix something creates a high IA in the topic until the deed is done. Moreover, relying on situation variables alone betrays the common observation that some people exhibit excessive information seeking tendencies even when there is no pressing need to do so.

Therefore, it can be concluded that a combination of both person and situation variables is at work. Future research will have to find variables that are reasonable at the collection level, however.

In the studies reported in this dissertation, the person variables were not found to make meaningful contributions to the model. Rather, it was a situation variable such as having free time that turned out to be the most important factor in predicting the time spent for PIPs. A person variable that was found to be a significant factor, the compulsive acquisition symptoms personality variable, explains that when the desire to spend time is presented, those with a higher acquisition impulse will more readily act on the desire and spend more time in PIPs.

The preliminary model presented in Chapter 1 has been revised as follows. Situation variables have been added to the model, an example of which is short-term needs. Topic interests have been revised to long-term topic interests to reflect the nested structure of temporal contexts in topic interests. The nested structure means that individuals were shown to have long term interests, and then medium term and short term interests that are nested within the long term interests. What is learned from the MTurk study is that a short-term interest born of a short-term need (e.g. fixing something broken) and a short-term interest born of a long-term interest (e.g. interest in some new TV actors as potential casts while being interested in a TV show for a long time, or interest in a caucus while being interested in politics for a long time) should be differentiated. As such, a short-term topic interest born of a short-term need would be classified as the situation variable, whereas a short-term topic interest born of a long-term interest would be classified as the long-term topic interest.

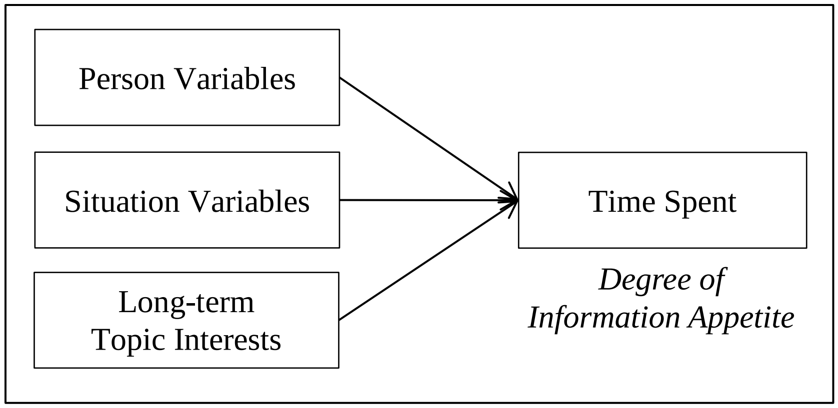


Figure 6.1 Revised model

6.2 Implications

The major implication of the research in this dissertation is that time is used as the measure of information appetite, and as both propositions show, individuals were found to be spending various amount of time in PIP activities and in different topics. Hence the studies validate the use of time as the measure.

The second implication is the situation variable, having free time. In information overload, the insufficiency of time was a major factor in determining the state of information overload, and some scholars argued that individuals exercise various strategies to avoid being overloaded (Jacoby, 1984; Savolainen, 2007). The strategies such as filtering and withdrawal might be at work in satisfying information appetite as well. The filtering strategy is filtering out useless information from the sources while focusing on the content to avoid; the withdrawal strategy is minimizing the daily information sources, focusing on the information supply to avoid in order not to be overwhelmed with information (Savolainen, 2007). This means that when individuals know that they have limited free time for information use, they might employ similar strategies to economically spend their time, and such time use may not accurately represent their actual desire to spend time with information. This presents a challenge in measuring information appetite as it is difficult to measure an individual's desire directly. Hence, extra questions would be needed to find out if the participants economized their time spending or not.

The situation variable, having free time, also relates to the question of information appetite and satiety. The studies here assumed that people stop information seeking when they are satiated. However, for those with high information appetite, it is questionable whether they have been truly satiated: they may not have been satiated, but stopped because they ran out of available time. Moreover, several people mentioned in the open-ended question that they spend more time with information during evenings and weekends when their time use is less constricted. In a future study, asking individuals why they stopped information seeking would be one way to learn if they stopped because their information appetite was satiated or because of other reasons such as running out of free time.

And yet, for long-term interests, perhaps satiety is a state that is never reached. Case (2012) pointed out that with an interest that is from a lifetime curiosity, there is no identifiable time pressure (p. 41). Any such information seeking could be picked up later where it had been left off and continued indefinitely. So, for long-term interests, one might never be fully satiated, yearning for more information endlessly.

One area that is left for future study is the three dimensions of PIPs – consumption, dissemination, and creation of information for personal reasons. In the factor analysis of the time spent on PIP activities in the MTurk study, creation activities loaded together as well as the consumption activities. However, since the PIP activity categories were largely informed by the ATUS study, several activities covered multiple dimensions, which made it impossible to analyze the effect of each dimension separately. A future study on PIPs and their three dimensions, as well as their influence on information appetite, might uncover the separate contributions of PIP dimensions and yield valuable insights.

Another area for future study is the kinds of information behaviors that represent high information appetite. Examples would be reading newspapers regularly or keeping up with the YouTube subscription feed daily. This dissertation focused on measuring information appetite with time, but did not investigate the kinds of behaviors that are exhibited by those with high information appetite. Such data could be gathered from focus groups, interviews, or a panel study where the researcher follows a group of participants over time.

The third area left for future study is if information appetite runs in families. Does parents' information appetite affect their children? Once the above-mentioned data on information behaviors that represent high information appetite is gathered, the behaviors could be used as markers to analyze both parents and children in a family and see if there is any pattern of similarities. There are some studies that argue that children inherit their parents' habits via genetic disposition and model learning (e.g. driving habits (Bianchi & Summala, 2004; Ferguson, Williams, Chapline, Reinfurt, & De Leonardis, 2001; G. Miller & Taubman - Ben-Ari, 2010; Taubman - Ben-Ari, Mikulincer, & Gillath, 2005)), but there is little such research in information behavior except for early research on reading habits (e.g. Stone & Wetherington Jr., 1979). If high information appetite is found to involve promotable behavior, studying intergenerational influence would be useful.

6.3 Limitations

To minimize the effect of memory, the MTurk survey asked about time expenditure in the last three hours only. In the open-ended question that asked about their time spending in interested topics in the last 24 hours, some participants noted that the last three hours did not represent their typical time expenditure. A time diary or a survey of a full day's record would have been more appropriate; or multiple samplings of three hours over multiple days would have been good, too. In fact in the MTurk study it was planned to execute the latter plan, but the plan changed to a single survey after the first pretest, as discussed in Chapter 4 for reasons that were particular to MTurk. It would be appropriate to execute the plan again on samples that are more controllable than MTurk, and see if this makes a difference in the results.

Moreover, the survey was conducted in the one-time cross-sectional way. This also happened due to the way MTurk worked, as discussed earlier in Chapters 4 and 5. In future studies, it would be interesting to include various days (both weekdays and weekends) and time periods in the data.

The next limitation was that the MTurk sample was using a computer when the individuals responded to the survey. This created a context where they have been using a computer (for potentially PIPs) for a while before answering the survey. As such, the MTurk sample's participation rate in the computer use for PIPs was the highest among PIP activities at almost 91%. This is an important contextual distinction to note.

In addition, in the MTurk study of topic interests, it did not differentiate topic interests born of a short-term need or those born of a long-term interest. After the study was done, in the data there was a confusion in the reporting of the length of time interested in the topic in that those who had a short-term interest born of a long-term interest did not seem sure whether to report the short time or the long time that they were interested in the topic. Therefore, these distinctions would have to be clearly made in future studies.

Next, there could be an alternative way of defining the high IA group. Instead of running cluster analysis or using mean plus standard deviation on the total time spent for PIPs, specific outliers in the total time spent for PIPs could be picked and of those outliers, other variables could be investigated to identify those who share similar characteristics. This way, those who have potentially high IA but did not spend much time during the surveyed time period might be identified in the high IA group. This method has not been explored in this dissertation, however. Also, such grouping would need to be confirmed with an additional, follow-up method such as an interview.

Lastly, the ATUS dataset and the study on MTurk were conducted on samples drawn from the United Sates population. There are other time use datasets available from other countries such as Europe; it would be interesting to analyze them and see if the propositions that were supported in this dissertation are also supported in samples from other countries.

6.4 Future Studies

Three future studies will be presented here. First is an in-depth interview/focus group, second is a survey with follow-up interview, and third is an experiment.

First, the purpose of the in-depth interviews (individual interviews and focus groups) is to explore what affects IA, to discover the variables that will explain more variance in the IA model, and to learn the kinds of information behaviors that represent high information appetite. The focus group will be broken by discipline/profession and age groups. The proposition is that people of high IA are drawn to certain disciplines/professions. The interviews and focus groups will be conducted around the following themes: what spurs them to begin and end information seeking, how regularly and repeatedly a single topic is searched for, time spent on the topic vs. time spent normally, topics that are searched for that are from a long-term interest vs. a short-term need, and so on.

Second, a survey will be conducted on two student groups with different profiles: one is college freshmen, and the other is graduate students in library studies. The hypothesis is that graduate students in library studies will have higher IA than college freshmen and have been attracted to information professions such as librarianship for that reason, following the new proposition that people of high IA are drawn to certain disciplines/professions; because college freshmen have not chosen their field yet, they make a good comparison group. In the survey, they will be asked about time spent on PIPs. It would take the form of time diary with PIPs as optional secondary activities. After the survey is finished, groups of high IA participants and regular IA participants will be identified; then they will be invited for a follow-up interview. They will be interviewed on their time spending in PIPs, their PIP-related habits, and so on.

Third, a group of participants will be recruited for an observational experiment that will take a set amount of time. The purpose of this experiment is to investigate the influence of topic interests on the time expenditure and information seeking patterns. The proposition is that IA is influenced by long-term interests. They will be asked which topics they have long-term interests in, and one of them will be randomly picked for their task. The control group will be given a topic that is not one of the participants' long-term interests. Using computers that are on the Internet, they will be given a task to find one or more interesting, latest news on the topic. They are free to determine how much time they want to spend on the task, and if they finish early, they can spend the rest of the time however they want but they may not leave the room early. How much time they spend on the task and their online activities will be recorded and analyzed. After the experiment, they will be asked about the experience and the reasons behind their actions.

The experiment could also be a 2x2 design in that the sample is divided into long-term interest/no interest and information field/non-information field (e.g. library program students vs. college freshmen) groups.

6.5 Concluding Remarks

Despite the limitations noted earlier, this study made an important contribution in the study of information behavior by uncovering a human characteristic that has not been studied before. As an explorative study of information appetite, and the first of its kind, it served its role well.

Two propositions about IA were made, and both were supported by the datasets. The first proposition was that individuals have varying degrees of information appetite, as measured by the amount of time spent engaging in personal information practices, which was supported by the ATUS dataset and the study on MTurk. The second proposition was that an individual has different information appetites for different topical areas, which was supported by the MTurk data.

A model of information appetite was proposed and explored in the MTurk study. Although the statistical representation of the model was poorly supported in general with very small explanatory power, this suggests that the variables in the model did not fully explain the variance in information appetite, which is still a meaningful finding.

As for what affects spending time for PIPs, having free time (e.g. not working/not having a job, or evenings and weekends) was found to be an important factor in both statistical analyses and the analysis of the open-ended question.

Three future studies were suggested to explore what affects IA and to study two more propositions. The first new proposition is that people of high information appetite are drawn to certain disciplines/professions, and the second new proposition is that information appetite is influenced by long-term interests.

This dissertation stemmed from the question about when individuals are satisfied with information. It attempted to answer the question from the perspective of information appetite, expressed in the amount of time spent engaging in PIPs. Even though the finding is meager, the topic still has academic and practical importance. It presents a behavioral model of when individuals stop information seeking. It is commonly observed that some people spend much time even when the same task is assigned to them. IA is not a part of relevance judgement, but instead it may represent a habit and how people spend their time. In addition, being able to predict how much time people spend has a very useful implication in online advertisement marketing, for instance. As the first study of information appetite, this study served its role and paves the road for next studies.

Appendix A

American Time Use Survey – PIP Activity Examples

Sources:

Hofferth, S. L., Flood, S. M., & Sobek, M. (2013). *American Time Use Survey Data Extract System: Version 2.4 [Machine-readable database]*. College Park, MD, and Minneapolis, MN: Maryland Population Research Center, University of Maryland, College Park and Minnesota Population Center, University of Minnesota. Retrieved from http://www.atusdata.org

U.S. Bureau of Labor Statistics. (2015). *American Time Use Survey activity lexicon: 2014*. Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from http://www.bls.gov/tus/lexicons.htm

Taking Class for Personal Interest (ATUS code 060102)

- Attending Bible study
- Attending Lamaze class
- Attending Sunday school
- Dance class (personal interest)
- Prenatal/child care classes (personal interest)
- Taking a car maintenance/repair class (personal interest)
- Taking a cooking class (personal interest)
- Taking a financial planning class (personal interest)
- Taking a massage class (personal interest)
- Taking a pottery class (personal interest)
- Taking a retirement planning seminar
- Taking a sewing class (personal interest)
- Taking a wine appreciation class (personal interest)
- Taking academic class (personal interest)

- Taking an art, craft, hobby, recreational course (personal interest)
- Taking CPR, first aid (personal interest)
- Taking driver's education
- Taking driving lessons
- Taking music/voice lessons (personal interest)
- Taking on-line course (personal interest)
- Taking parenting class
- Taking personal development classes (personal interest)
- Taking photography class (personal interest)
- Taking self-defense (personal interest)
- Talking to classmates (class for personal interest)
- Talking to teacher (class for personal interest)

Research or Homework for Class (For Personal Interest) (ATUS code 060302)

- Attending study group (class for personal interest)
- Listening to language CD (class for personal interest)
- Organizing notes (class for personal interest)
- Reading (class for personal interest)
- Reading/sending e-mail (class for personal interest)
- Studying (class for personal interest)
- Writing paper/essay (class for personal interest)

Researching Purchases (ATUS code 070200)

- Comparison Shopping (ATUS code 070201)
    - Browsing through circulars
    - Comparing prices at different stores
    - Comparison shopping on the internet
    - Reading product reviews
    - Researching items/prices/availability
- Researching Purchases, N.E.C. (ATUS code 070299)

Television and Movies (Not Religious) (ATUS code 120303)

- Borrowing movies from the library
- Returning movies to library
- Rewinding a tape or DVD (2003-2004)
- Setting TiVo/DVR (2011+)
- Setting the VCR or DVD player
- Watching a DVD/video/instructional video
- Watching home movies/home videos
- Watching TV
- Watching TV/DVDs on computer (personal interest) (2011+)
- Watching videos on YouTube (2011+)

Listening to the Radio (ATUS code 120305)

- Listening to a radio talk show
- Listening to music on the radio

- Listening to public radio
- Listening to the top ten on the radio

Computer Use for Leisure (Excluding Games) (ATUS code 120308)

- Browsing on the internet (personal interest)
- Burning CDs (personal interest) (2006+)
- Checking Facebook, MySpace, or other social networking Web sites (personal interest) (2011)
- Checking Facebook (personal interest) (2012+)
- Checking MySpace (personal interest) (2012+)
- Computer programming (personal interest) (2012+)
- Computer use, leisure (personal interest)
- Computer use, unspecified
- Designing/updating website (personal interest)
- Downloading files, music, pictures (personal interest)
- Participating in a chat room (personal interest)
- Surfing the internet (personal interest)
- Surfing the web (personal interest)
- Using Twitter or tweeting (personal interest) (2011+)
- Writing computer software (personal interest) (2012+)

Reading for Personal Interest (ATUS code 120312)

- Being read to (personal interest)

- Borrowing books from the library (personal interest)
- Browsing at the library (personal interest)
- Checking out library books (personal interest)
- Doing research (personal interest)
- Flipping/leafing through magazine (personal interest)
- Listening to books on tape (personal interest)
- Reading a book on a Kindle or other electronic book reader (personal interest) (2011+)
- Reading a magazine/book (personal interest)
- Reading scripture (personal interest)
- Reading the Bible (personal interest)
- Reading the newspaper (personal interest)
- Reading, unspecified
- Returning library books

Writing for Personal Interest (ATUS code 120313)

- Blogging (personal interest) (2011+)
- Editing (personal interest) (2004+)
- Writing in diary (personal interest)
- Writing in journal (personal interest)
- Writing lyrics (2005+)
- Writing stories (personal interest)

Appendix B

Survey Questionnaire

I. Demographic

Q1. What is your gender?

- Male
- Female

Q2. Please enter your age. [__________]{→ End of survey if younger than 18}

Q3. In which time zone are you currently in?

- US/Eastern
- US/Central
- US/Mountain
- US/Pacific
- Other {→ End of survey}

II. Socio-economic status

Q4. What is the highest level of school you have completed or the highest degree you have received?

- 12th grade (no diploma) or below
- High school graduate - diploma or the equivalent (e.g. GED)
- Some college but no degree
- Associate degree in college - occupational/vocational program
- Associate degree in college - academic program
- Bachelor's degree (BA, AB, BS, etc.)
- Master's degree (MA, MS, MEng, MEd, MSW, etc.)

- Professional school degree (MD, DDS, DVM, etc.)
- Doctoral degree (PhD, EdD, etc.)

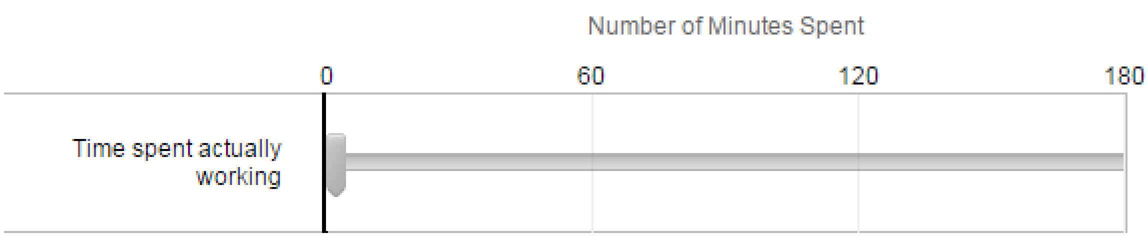
Q5. Do you currently have a job?

- Employed - at work
- Employed - absent
- Unemployed - on layoff
- Unemployed - looking
- Not in labor force

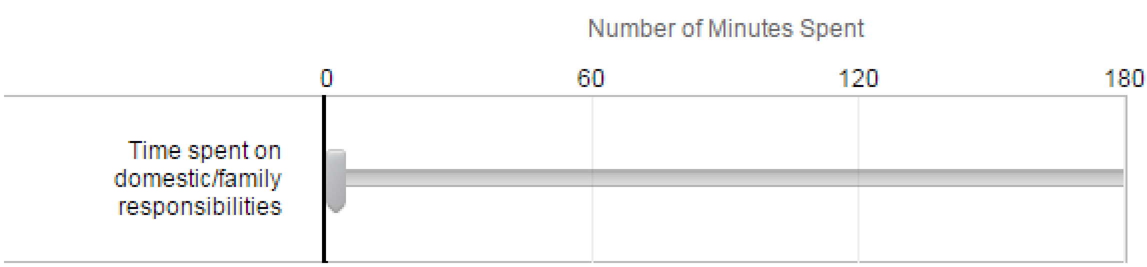
Q6. {If Employed} Do you have more than one job?

- Yes
- No

Q7. {If Employed} In the last three hours, how much time did you spend actually working? (in minutes)



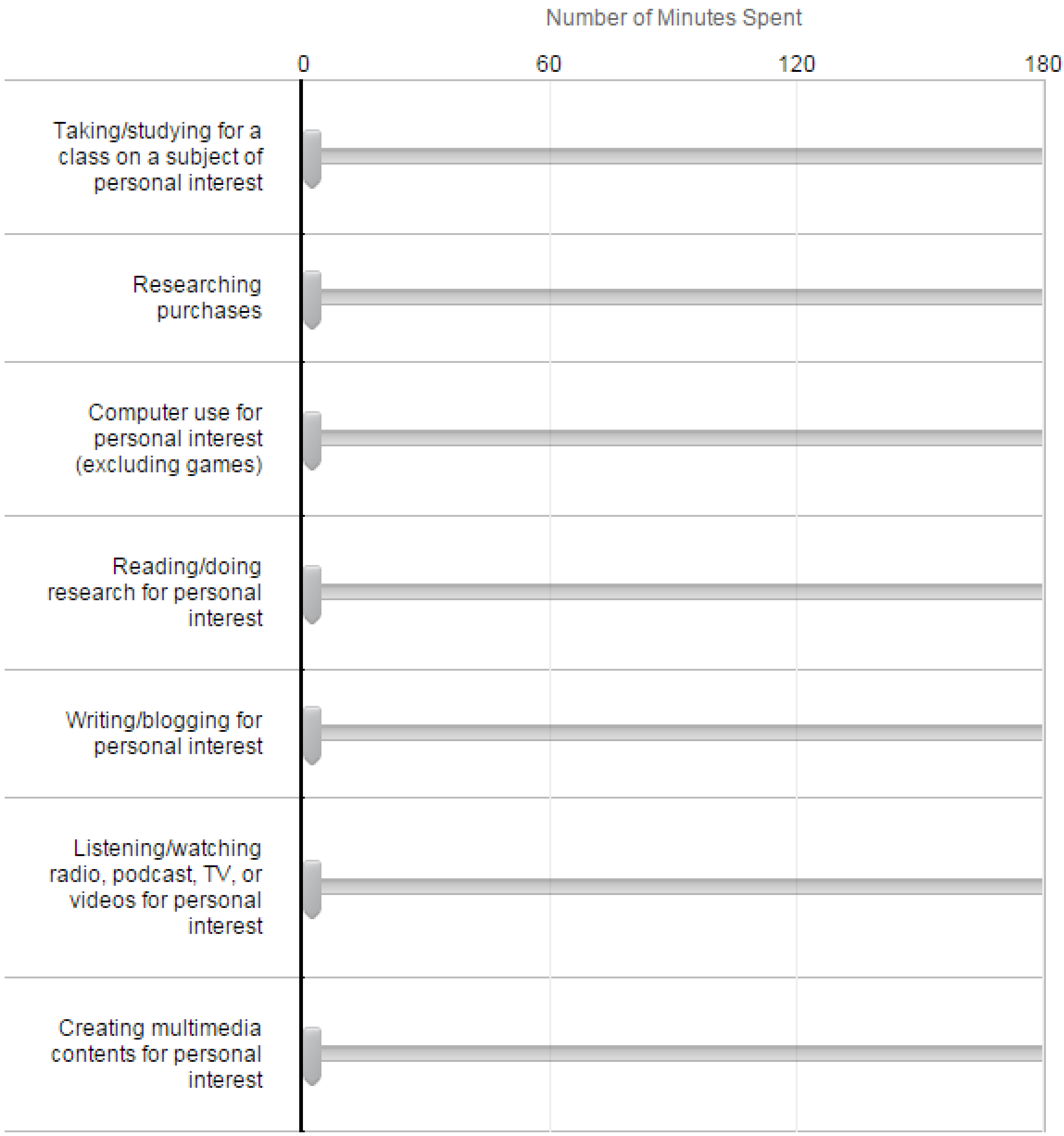
Q8. In the last three hours, how much time did you spend on domestic and/or family responsibilities? (in minutes)



III. Time use: Personal information practices

Q9. You will now be asked to indicate the amount of minutes you spent in the past three hours for six personal information activities. The times you spent for each activity are NOT mutually exclusive: that is, if you spent 30 minutes to do personal research on computer, the 30 minutes count towards both Computer use and Reading/Doing research.

In the last three hours, for how long did you engage in each of the following activities for personal purposes (NOT for work)? Please use the slider to select the amount of minutes you spent for each activity.



IV. Topic interests and time use

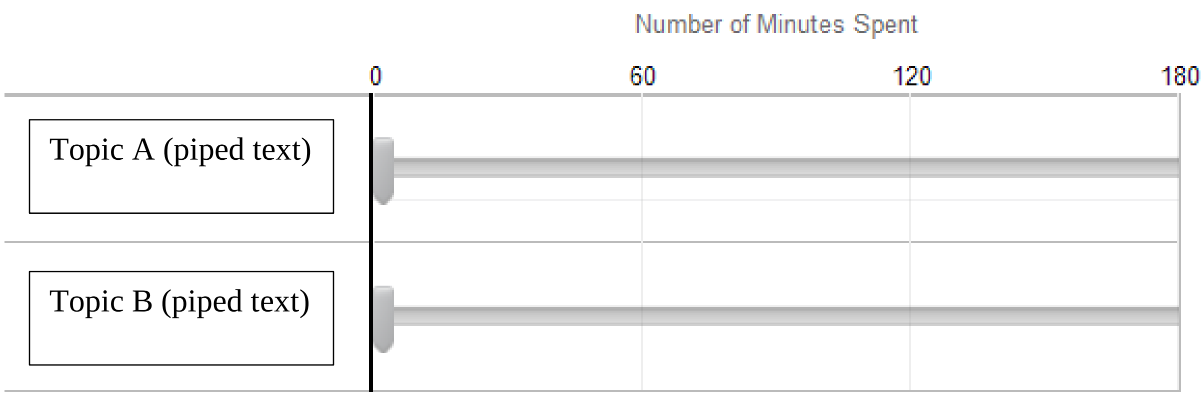
Q10. Please name two topics on which you spent time with information for personal purposes in the last three hours. (in no particular order)

- Topic A: __________
- Topic B: __________

Q11. For each topic, please tell us for how long you have been interested in it. (e.g. 3 days, 5 years, etc.)

- {Topic A (piped text)}: __________
- {Topic B (piped text)}: __________

Q12. In the last three hours, how much time did you spend with information about your interested topics? Please use the slider to select the amount of minutes you spent per each topic.



V. Epistemic curiosity scale

Q13. A number of statements that people use to describe themselves are given below. Read each statement and then select the appropriate response using the scale below to indicate how you generally feel. There are no right or wrong answers. Do not spend too much time on any one statement but give the answer that seems to describe how you generally feel.

1 = Almost Never 2 = Sometimes 3 = Often 4 = Almost Always

1) I enjoy exploring new ideas.
2) Difficult conceptual problems can keep me awake all night thinking about solutions.
3) I enjoy learning about subjects that are unfamiliar to me.

4) I can spend hours on a single problem because I just can’t rest without knowing the answer.
5) I find it fascinating to learn new information.
6) I feel frustrated if I can’t figure out the solution to a problem, so I work even harder to solve it.
7) When I learn something new, I would like to find out more about it.
8) I brood for a long time in an attempt to solve some fundamental problem.
9) I enjoy discussing abstract concepts.
10) I work like a fiend at problems that I feel must be solved.

VI. Acquisition (SI-R) scale

Please select the response that is most appropriate.

Q14. How distressed or uncomfortable would you feel if you could not acquire something you wanted?

0 = Not at all

1 = Mild, only slightly anxious

2 = Moderate, distress would mount but remain manageable

3 = Severe, prominent and very disturbing increase in distress

4 = Extreme, incapacitating discomfort from any such effort

Q15. How often do you feel compelled to acquire something you see (e.g., when shopping or offered free things)?

0 = Never feel compelled

1 = Rarely feel compelled

2 = Sometimes feel compelled

3 = Frequently feel compelled

4 = Almost always feel compelled

Q16. How strong is your urge to buy or acquire free things for which you have no immediate use?

0 = Urge is not at all strong

1 = Mild urge

2 = Moderate urge

3 = Strong urge

4 = Very strong urge

Q17. How much control do you have over your urges to acquire possessions?

0 = Complete control

1 = Much control, usually able to control urges to acquire

2 = Some control, can control urges to acquire only with difficulty

3 = Little control, can only delay urges to acquire only with great difficulty

4 = No control, unable to stop urges to acquire possessions

Q18. How often do you actually buy (or acquire for free) things for which you have no immediate use or need?

0 = Never

1 = Rarely

2 = Sometimes

3 = Frequently

4 = Almost always

Q19. How upset or distressed do you feel about your acquiring habits?

0 = Not at all upset

1 = Mildly upset

2 = Moderately upset

3 = Severely upset

4 = Extreme embarrassment

Q20. To what extent has your saving or compulsive buying resulted in financial difficulties for you?

0 = Not at all

1 = A little financial difficulty

2 = Some financial difficulty

3 = Quite a lot of financial difficulty

4 = An extreme amount of financial difficulty

VII. Open-ended question

Q21. Earlier, you answered a question on your personal topic of interest in the past three hours. When you look back the last 24 hours, is there anything else you would like to say about your personal topic of interest and how you spent time with information regarding the topic(s)? [Optional]

VIII. End of survey

Here is a randomly regenerated code for you: {CODE}

Please copy and paste the code on Amazon Mechanical Turk to prove that you completed the survey. Researcher will verify the code for your reward.

References

Allen, D. (2011). Information behavior and decision making in time-constrained practice: A dual-processing perspective. *Journal of the American Society for Information Science and Technology*, *62*(11), 2165–2181. http://doi.org/10.1002/asi.21601

Allison, P. D. (1999). *Multiple regression: A primer*. Thousand Oaks, CA: Pine Forge Press.

Appetite (diet). (2015, January 9). *Encyclopedia Britannica*. Retrieved from http://www.britannica.com/EBchecked/topic/30576/appetite

Bawden, D., & Robinson, L. (2009). The dark side of information: Overload, anxiety and other paradoxes and pathologies. *Journal of Information Science*, *35*(2), 180–191. http://doi.org/10.1177/0165551508095781

Belkin, N. J. (2005). Anomalous state of knowledge. In K. E. Fisher, S. Erdelez, & L. (E. F. . McKechnie (Eds.), *Theories of information behavior* (pp. 44–48). Medford, NJ: Information Today.

Berlyne, D. E. (1954). A theory of human curiosity. *British Journal of Psychology*, *45*(3), 180–191.

Bianchi, A., & Summala, H. (2004). The “genetics” of driving behavior: parents’ driving style predicts their children’s driving style. *Accident Analysis & Prevention*, *36*(4), 655–659. http://doi.org/10.1016/S0001-4575(03)00087-3

Case, D. O. (Ed.). (2012). *Looking for information: A survey of research on information seeking, needs, and behavior* (3rd ed.). United Kingdom: Emerald Group Publishing.

Casserly, M. (2012, October 4). Top Five Personality Traits Employers Hire Most. *Forbes*. Retrieved from http://www.forbes.com/sites/meghancasserly/2012/10/04/top-five-personality-traits-employers-hire-most/

Chen, S.-Y. (Max), & Rieh, S. Y. (2009). Take your time first, time your search later: How college students perceive time in Web searching. *Proceedings of the American Society for Information Science and Technology*, *46*(1), 1–19. http://doi.org/10.1002/meet.2009.1450460253

Crescenzi, A., Capra, R., & Arguello, J. (2013). Time pressure, user satisfaction and task difficulty. *Proceedings of the American Society for Information Science and Technology*, *50*(1), 1–4. http://doi.org/10.1002/meet.14505001121

Dholakia, U. M. (2014). *Three senses of desire in consumer research* (SSRN Scholarly Paper No. ID 2420250). Rochester, NY: Social Science Research Network. Retrieved from http://papers.ssrn.com/abstract=2420250

Diderot, D. (1964). The encyclopedia. In J. Barzun & R. H. Bowen (Trans.), *Rameau's nephew and other works* (pp. 277–307). Indianapolis, IN: Bobbs-Merrill Company.

Duggan, M. (2015). *Mobile messaging and social media - 2015*. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/2015/08/19/mobile-messaging-and-social-media-2015/

Edmunds, A., & Morris, A. (2000). The problem of information overload in business organisations: A review of the literature. *International Journal of Information Management*, *20*(1), 17–28. http://doi.org/10.1016/S0268-4012(99)00051-1

Eppler, M. J., & Mengis, J. (2004). The concept of information overload: A review of literature from organization science, accounting, marketing, MIS, and related disciplines. *The Information Society*, *20*(5), 325–344. http://doi.org/10.1080/01972240490507974

Ferguson, S. A., Williams, A. F., Chapline, J. F., Reinfurt, D. W., & De Leonardis, D. M. (2001). Relationship of parent driving records to the driving records of their children. *Accident Analysis & Prevention*, *33*(2), 229–234. http://doi.org/10.1016/S0001-4575(00)00036-1

Frost, R. O., Steketee, G., & Grisham, J. (2004). Measurement of compulsive hoarding: saving inventory-revised. *Behaviour Research and Therapy*, *42*(10), 1163–1182. http://doi.org/10.1016/j.brat.2003.07.006

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. *American Psychologist*, *54*(7), 493–503.

Hartel, J. (2010). Time as a framework for information science: Insights from the hobby of gourmet cooking. *Information Research*, *15*(4), colis715.

Harvey, A. S. (1999). Guidelines for time use data collection and analysis. In W. E. Pentland, A. S. Harvey, M. P. Lawton, & M. A. McColl (Eds.), *Time use research in the social sciences* (pp. 19–45). New York: Kluwer Academic/Plenum Publishers.

Harvey, A. S., & Pentland, W. E. (1999). Time use research. In W. E. Pentland, A. S. Harvey, M. P. Lawton, & M. A. McColl (Eds.), *Time use research in the social sciences* (pp. 3–18). New York: Kluwer Academic/Plenum Publishers.

Hauser, D. J., & Schwarz, N. (2015). Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants. *Behavior Research Methods*, 1–8. http://doi.org/10.3758/s13428-015-0578-z

Hofferth, S. L., Flood, S. M., & Sobek, M. (2013). *American Time Use Survey Data Extract System: Version 2.4 [Machine-readable database]*. College Park, MD, and Minneapolis, MN: Maryland Population Research Center, University of Maryland, College Park and Minnesota Population Center, University of Minnesota. Retrieved from http://www.atusdata.org

Hofmann, W., Baumeister, R. F., Förster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology*, *102*(6), 1318–1335. http://doi.org/10.1037/a0026545

Jacoby, J. (1984). Perspectives on information overload. *Journal of Consumer Research*, *10*(4), 432–435.

Kimmel, J. (2008). Introduction. In J. Kimmel (Ed.), *How do we spend our time? Evidence from the American Time Use Survey* (pp. 1–9). Kalamazoo, Michigan: W.E. Upjohn Institute for Employment Research.

Kim, Y.-M., Rieh, S. Y., Yang, J. Y., & St. Jean, B. (2009). An online activity diary method for studying credibility assessment on the Web. *Proceedings of the American Society for Information Science and Technology*, *46*(1), 1–5. http://doi.org/10.1002/meet.2009.1450460388

Krathwohl, D. R. (1997). *Methods of educational & social science research: An integrated approach* (2nd ed.). New York: Longman.

Kubey, R., Larson, R., & Csikszentmihalyi, M. (1996). Experience Sampling Method applications to communication research questions. *Journal of Communication*, *46*(2), 99–120. http://doi.org/10.1111/j.1460-2466.1996.tb01476.x

Leech, N. L., Barrett, K. C., & Morgan, G. A. (2015). *IBM SPSS for intermediate statistics: Use and interpretation* (5th ed.). New York: Routledge.

Litman, J. (n.d.). The Interest (I) Type and Deprivation (D) Type Scales (English). Retrieved from http://drjlitman.net/?page_id=171

Litman, J. A. (2008). Interest and deprivation factors of epistemic curiosity. *Personality and Individual Differences*, *44*(7), 1585–1595. http://doi.org/10.1016/j.paid.2008.01.014

Litman, J. A. (2010). Relationships between measures of I- and D-type curiosity, ambiguity tolerance, and need for closure: An initial test of the wanting-liking model of information-seeking. *Personality and Individual Differences*, *48*(4), 397–402. http://doi.org/10.1016/j.paid.2009.11.005

Litman, J. A., & Spielberger, C. D. (2003). Measuring epistemic curiosity and its diversive and specific components. *Journal of Personality Assessment*, *80*(1), 75–86. http://doi.org/10.1207/S15327752JPA8001_16

Liu, C., Yang, F., Zhao, Y., Jiang, Q., & Zhang, L. (2014). What does time constraint mean to information searchers? In *Proceedings of the 5th Information Interaction in Context Symposium* (pp. 227–230). New York, NY, USA: ACM. http://doi.org/10.1145/2637002.2637029

Miller, G., & Taubman - Ben-Ari, O. (2010). Driving styles among young novice drivers—The contribution of parental driving styles and personal characteristics. *Accident Analysis & Prevention*, *42*(2), 558–570. http://doi.org/10.1016/j.aap.2009.09.024

Miller, J. G. (1960). Information input overload and psychopathology. *American Journal of Psychiatry*, *116*(8), 695–704.

Motivation: Research starters topic. (2015). *Research Starter Topic Pages*. EBSCOhost. Retrieved from http://eds.a.ebscohost.com.proxy.libraries.rutgers.edu/eds/detail/detail?sid=56cb9f3f-a33e-4f0e-97f2-d3c299e1b353%40sessionmgr4004&vid=3&hid=4111&bdata=JnNpdGU9ZWRzLWxpdmU%3d#AN=91919385&db=ers

Naqvi, N., Shiv, B., & Bechara, A. (2006). The role of emotion in decision making: A cognitive neuroscience perspective. *Current Directions in Psychological Science*, *15*(5), 260–264. http://doi.org/10.1111/j.1467-8721.2006.00448.x

Paisley, W. (1993). Knowledge utilization: The role of new communication technologies. *Journal of the American Society for Information Science*, *44*(4), 222–234. http://doi.org/10.1002/(SICI)1097-4571(199305)44:4<222::AID-ASI10>3.0.CO;2-0

Paolacci, G., & Chandler, J. (2014). Inside the Turk: Understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23*(3), 184–188. http://doi.org/10.1177/0963721414531598

Pertusa, A., Frost, R. O., Fullana, M. A., Samuels, J., Steketee, G., Tolin, D., … Mataix-Cols, D. (2010). Refining the diagnostic boundaries of compulsive hoarding: A critical review. *Clinical Psychology Review*, *30*(4), 371–386. http://doi.org/10.1016/j.cpr.2010.01.007

Perugini, M., & Bagozzi, R. P. (2001). The role of desires and anticipated emotions in goal-directed behaviours: Broadening and deepening the theory of planned behaviour. *The British Journal of Social Psychology*, *40*, 79–98.

Perugini, M., & Bagozzi, R. P. (2004). The distinction between desires and intentions. *European Journal of Social Psychology*, *34*(1), 69–84. http://doi.org/10.1002/ejsp.186

Peterson, R. A., & Merunka, D. R. (2014). Convenience samples of college students and research reproducibility. *Journal of Business Research*, *67*(5), 1035–1041. http://doi.org/10.1016/j.jbusres.2013.08.010

Pew Research Center. (2015a). *App vs. web for surveys of smartphone users: Experimenting with mobile apps for signal-contingent experience sampling method surveys*. Washington, DC: Pew Research Center. Retrieved from http://www.pewresearch.org/2015/04/01/app-vs-web-for-surveys-of-smartphone-users/

Pew Research Center. (2015b). *The smartphone difference: U.S. smartphone use in 2015*. Washington, DC: Pew Research Center. Retrieved from http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/

Rieh, S. Y., Kim, Y.-M., Yang, J. Y., & St. Jean, B. (2010). A diary study of credibility assessment in everyday life information activities on the web: Preliminary findings. *Proceedings of the American Society for Information Science and Technology*, *47*(1), 1–10. http://doi.org/10.1002/meet.14504701182

Robinson, J. (2014, September). Pay attention! *Entrepreneur*, *42*(9), 60–65.

Robinson, J. P., & Godbey, G. (1999). *Time for life: The surprising ways Americans use their time* (2nd ed.). University Park, PA: Pennsylvania State University Press.

Rogers, E. M. (1986). *Communication technology: The new media in society*. New York: Free Press; London: Collier Macmillan.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. http://doi.org/10.1006/ceps.1999.1020

Salkind, N. J. (2012). *100 questions (and answers) about research methods*. Thousand Oaks, CA: Sage. Retrieved from http://www.amazon.com/Questions-Answers-About-Research-Methods/dp/1412992036/ref=tmm_pap_title_0?_encoding=UTF8&qid=&sr=

Sasaki, Y., Kawai, D., & Kitamura, S. (2015). The anatomy of tweet overload: How number of tweets received, number of friends, and egocentric network density affect perceived information overload. *Telematics and Informatics, 32*(4), 853–861. http://doi.org/10.1016/j.tele.2015.04.008

Savolainen, R. (2006). Time as a context of information seeking. *Library & Information Science Research*, *28*(1), 110–127. http://doi.org/10.1016/j.lisr.2005.11.001

Savolainen, R. (2007). Filtering and withdrawing: Strategies for coping with information overload in everyday contexts. *Journal of Information Science*, *33*(5), 611–621. http://doi.org/10.1177/0165551506077418

Shanker, D. (2015, March 8). 15 secrets from recruiters on how to get hired. *Fortune*. Retrieved from http://fortune.com/2015/03/08/getting-hired-by-best-companies/

Slone, D. J. (2007). The impact of time constraints on Internet and Web use. *Journal of the American Society for Information Science and Technology*, *58*(4), 508–517. http://doi.org/10.1002/asi.20525

Smith, S. M. (n.d.). Determining sample size: How to ensure you get the correct sample size. Qualtrics. Retrieved from http://success.qualtrics.com/rs/qualtrics/images/Determining-Sample-Size.pdf

Sonck, N., & Fernee, H. (2013). *Using smartphones in survey research: A multifunctional tool*. The Hague: The Netherlands Institute for Social Research.

Sonnenwald, D. H., & Iivonen, M. (1999). An integrated human information behavior research framework for information studies. *Library & Information Science Research*, *21*(4), 429–457. http://doi.org/10.1016/S0740-8188(99)00023-7

St. Jean, B., Rieh, S. Y., Kim, Y.-M., & Yang, J. Y. (2012). An analysis of the information behaviors, goals, and intentions of frequent Internet users: Findings from online activity diaries. *First Monday*, *17*(2). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/3870

Stone, G. C., & Wetherington Jr., R. V. (1979). Confirming the newspaper reading habit. *Journalism Quarterly*, *56*(3), 554–566.

Taubman - Ben-Ari, O., Mikulincer, M., & Gillath, O. (2005). From parents to children—similarity in parents and offspring driving styles. *Transportation Research Part F: Traffic Psychology and Behaviour*, *8*(1), 19–29. http://doi.org/10.1016/j.trf.2004.11.001

U.S. Bureau of Labor Statistics. (2015a). *American Time Use Survey activity lexicon: 2014*. Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from http://www.bls.gov/tus/lexicons.htm

U.S. Bureau of Labor Statistics. (2015b). *American Time Use Survey user's guide: Understanding ATUS 2003 to 2014*. Washington, DC: U.S. Bureau of Labor Statistics. Retrieved from http://www.bls.gov/tus/atususersguide.pdf

von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science*, *6*(6), 574–588. http://doi.org/10.1177/1745691611421204

Warner, R. M. (2013). *Applied Statistics: From bivariate through multivariate techniques* (2nd ed.). Thousand Oaks, CA: Sage.

Zach, L. (2005). When is "enough" enough? Modeling the information-seeking and stopping behavior of senior arts administrators. *Journal of the American Society* *for Information Science and Technology*, *56*(1), 23–35. http://doi.org/10.1002/asi.20092

Zuckerman, M. (2009). Sensation seeking. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (pp. 455–465). New York: Guilford Press.