ASSESSING THE PSYCHOMETRIC PROPERTIES OF NEWLY DEVELOPED BEHAVIOR AND ATTITUDE TWITTER SCALES: A VALIDITY AND RELIABILITY STUDY

A dissertation submitted to the
Kent State University College
of Education, Health, and Human Services
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

By

Md Amiruzzaman

December 2019

A dissertation written by

Md Amiruzzaman

B.S., National University, 2005

M.S., Sejong University, 2008

M.S., Kent State University, 2011

M.Tech., Kent University, 2015

Ph.D., Kent State University, 2016

Ph.D., Kent State University, 2019

Approved by

____________________________, Co-Director, Doctoral Dissertation Committee
Tricia Niesz, Ph.D.

____________________________, Co-Director, Doctoral Dissertation Committee
Karl W. Kosko, Ph.D.

____________________________, Member, Doctoral Dissertation Committee
Richard E. Ferdig, Ph.D.

Accepted by

____________________________, Director, School of Foundations, Leadership, and
Kimberly S. Schimmel, Ph.D. Administration

____________________________, Dean, College of Education, Health and Human
James C. Hannon, Ph.D. Services

AMIRUZZAMAN, MD, Ph.D., December 2019 Evaluation and Measurement

ASSESSING THE PSYCHOMETRIC PROPERTIES OF NEWLY DEVELOPED BEHAVIOR AND ATTITUDE TWITTER SCALES: A VALIDITY AND RELIABILITY STUDY (147 pp.)

Co-Directors of Dissertation: Tricia Niesz, Ph.D.
Karl W. Kosko, Ph.D.

The purpose of this study was to explore the psychometric properties of the newly developed Twitter and Scholastic Synchronicity Scale (TSSS) and Twitter and Scholastic Apportionment Assessment (TSAA) items. The study also sought to understand if and how attitude and behavior positively and/or negatively related to undergraduate students' academic performance. The TSSS scale focused on measuring how undergraduate students use Twitter for academics while the TSAA scale focused on their attitudes toward using Twitter for academics.

A comprehensive statistical analysis was conducted to explore both the validity and reliability aspects of these newly developed scales. An online survey collected research data from 327 undergraduate students at one institution. First, Exploratory Factor Analysis (EFA) was used to understand the underlying factor structure. Second, Confirmatory Factor Analysis (CFA) was used to check the proximity of the conceptual model's results to the hypothesized model. Third, reliability and validity aspects of the measure were investigated using Classical Test Theory (CTT) and Rasch Analysis (RA). Fourth, Hierarchical Multiple Regression (HMR) was used to understand the relationship between students' academics and the newly developed scales.

The results here provide evidence for the reliability and validity of the newly developed scales. CFA confirmed that research data support the hypothesized data, and RA indicated that

the items featured in these newly developed scales are based on a single measure. The HMR results indicated that students' academic performance and Twitter scales (TSSS and TSAA) are strongly correlated. Both scales help to explain the variance in undergraduate students' academic performance.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to those people who made the journey a memorable one.

First, I would like to thank Dr. Karl W. Kosko, who served as a co-chair of my dissertations. He helped me to successfully defend two of my dissertations. So many times, I needed his guidance and suggestions, and he was always available to guide me, discuss with me, and read my drafts. I think, today I am a better writer than yesterday—all because of him.

Second, I would like to thank Dr. Tricia Niesz and Dr. Rick Ferdig for helping me to think further about my research findings and limitations. Third, I would like to thank Dr. Joanne Caniglia for serving as a graduate representative, and Dr. Aryn Karpinski for being there whenever I needed some guidance. Also, Dr. Karpinski initially helped me to develop and analyze my survey items.

Finally, most importantly, my heartfelt thanks to my wife Stefanie Amiruzzaman and son Aleo Amiruzzaman who put up with me throughout this educational journey. I dedicate this dissertation in their names. Their love, support and encouragement helped me to complete this dissertation. Aside from my advisors, Stefanie had to read and reread my dissertation numerous times and helped me to explain things clearly. Aleo volunteered to be my audience several times before my proposal and defense presentations.

TABLE OF CONTENTS

Page
ACKNOWLEDGEMENTS iv

LIST OF FIGURES ix

LIST OF TABLES x

CHAPTER
I. INTRODUCTION 1
Introduction and Background 1
Need for This Study 2
Purpose of the Study 6
Summary 7
Organization of the Dissertation 8

II. LITERATURE REVIEW 9
Introduction 9
Literature Review 10
Theme/Scale Development 16
Multitasking 18
Responsibility 20
Perception 21
Schoolwork 22
Procrastination and Leisure 23
Measure Development 25
Measures/Items 25
Twitter and Scholastic Synchronicity Scale (TSSS) 27
Twitter and Scholastic Apportionment Scale (TSAA) 28
Item Characteristics 29
Construct map 30
Construct Map for Twitter and Scholastic Synchronicity Scale (TSSS) 31
Construct map for multitasking. 31

Construct map for responsibility 31
Construct map for perception 32
Construct Map for Twitter and Scholastic Apportionment Assessment (TSAA) 33
Construct map for schoolwork. 33
Construct map for procrastination and leisure 33
Validity Argument 34
Summary 39

III. METHODOLOGY 40
Introduction 40
Data Collection Procedure 42
Instrument Development 42
Pilot Study 43
Participants 44
Population 44
Sample 45
Research Data Collection and Administration 45
Administration 46
Data Analysis Procedure 47
Factor Analysis 47
Explanatory factor analysis 49
Confirmatory factor analysis 50
Validity and Reliability Analysis 51
Evaluation of validity indices 51
Evaluation of reliability indices 53
Validity of the Study 53
Rasch Analysis 54
Unidimensionality 55
Wright map and construct key map 56
Logit 56
Item fit 57
Linear Regression Analysis 58
Method of entry 58

Assumptions 61
Summary 62

IV. RESULTS AND DISCUSSION 63
Introduction 63
Factor Analysis 64
Explanatory Factor Analysis 64
EFA and Item Analysis of the TSSS 65
EFA and Item Analysis of the TSAA 68
Confirmatory Factor Analysis 70
CFA and Item Analysis of the TSSS 70
CFA and Item Analysis of the TSAA 72
Rasch Analysis 74
Unidimensionality 75
Item Fit Statistics 75
Item difficulty and rating scales of individual subscales 76
Principal Component Analysis (PCA) of Residuals 80
Wright Map, Construct Key Map, and Logit 81
Wright map 81
Construct keymap 83
Multiple Regression 84
TSAA and GPA 84
TSSS and GPA 89
Summary 93

V. DISCUSSION, IMPLICATIONS, AND CONCLUSION 95
Introduction 95
Summary of the Study 95
Findings of the Study 96
Research Question 1: Psychometric Properties of TSSS Scale 96
Research Question 2: Psychometric Properties of the TSAA Scale 98
Research Question 3: Academic Performance and Twitter Scales 99
Implications and Future Research Directions 100

Implications for undergraduate student users 100
Implications for Collegiate Faculty and Administrators 102
Implications for Research 103
Limitations 106
Summary and Conclusion 109

APPENDICES 111
APPENDIX A. SAMPLE DATA FROM PILOT STUDY 112
APPENDIX B. RATING SCALE AND WRIGHT MAP 122

REFERENCES 125

LIST OF FIGURES

Figure Page

1. Constructs and scales of the study 26

2. Block diagram of the research method 41

3. Block-diagram of different blocks for the Hierarchical Regression 60

4. CFA model for TSSS. 72

5. CFA model for TSAA. 74

6. Standard Residuals 80

7. Wright map of the newly developed Twitter scales. 82

8. Construct Keymap of the newly developed Twitter scales 84

9. Rating scale of Twitter Scale Measures 123

10. Wright map of the Twitter scales 124

LIST OF TABLES

Table Page

1. Scales, Sub-scales, and Items Descriptions ... 29
2. Construct map for TSSS subscales ... 32
3. Constract map for TSAA subscales ... 34
4. Summary of Validity Evidences, by Form, Provided in Current Study ... 36
5. Sample items for the instrument ... 43
6. Problematic items in associated scales and subscales ... 44
7. Summary of Validity Evidences, by Form, Provided in Current Study ... 54
8. Pattern Matrix for TSSS ... 65
9. Descriptive Statistics TSSS ... 66
10. Pattern Matrix II for TSSS ... 67
11. Pattern Matrix of TSAA ... 68
12. Descriptive Statistics TSAA ... 69
13. Pattern Matrix II of TSAA ... 69
14. Model Identification TSSS ... 70
15. Model Identification TSAA ... 73
16. Item Analysis Statistics for TSSS and TSAA Scales ... 76
17. Level of difficulty of individual subscales of TSAA subscales ... 78
18. Level of difficulty of individual subscales of TSSS subscales ... 79
19. ANOVA Table TSAA and GPA ... 86
20. Hierarchical Multiple Regression Analysis Summary Predicting College GPA ... 87
21. Coefficient Table GPA and TSAA ... 88

22. ANOVA Table TSSS and GPA 90

23. Hierarchical Multiple Regression Analysis Summary Predicting College GPA 91

24. Coefficient Table GPA and TSSS 92

25. Sample data from the pilot study demographic information (gender, age, ethnicity) 113

26. Sample data from the pilot study demographic information (parents' education) 114

27. Sample data from the pilot study Academic Information 115

28. Sample Data Pilot Study Twitter Use 116

29. Sample Data Pilot Study TSSS and Academic Performance 118

30. Sample Data Pilot Study TSAA and Academic Performance 120

CHAPTER I

INTRODUCTION

Introduction and Background

Twitter was founded in 2006 (Alhabash & Ma, 2017; Arceneaux & Schmitz Weiss, 2010), and, at present, the total number of monthly active Twitter users is approximately 326 million. It is a type of social networking site (SNS). *Social networking sites* can be defined as an Internet-based platform that allows users to communicate and share information with other users. In this study, the terms social networking site, social networking application, and social media are used interchangeably. Because Twitter is a real-time networking platform, it allows users to communicate and share information via small bursts of information (Alhabash & Ma, 2017). Among Twitter users, the *small burst* of information is known as a *tweet*. Each tweet is a short message that might contain the user's comments or an update of previous information (Alhabash & Ma, 2017; Jansen, Zhang, Sobel & Chowdury, 2009; Tierman, 2014).

Several studies have claimed that Twitter is one of the most commonly used social networking applications among college students (Aharony, 2010; Alhabash & Ma, 2017; Junco, Elavsky & Heiberger, 2013). According to Chaffey (2019), the world population is approximately 7.6 billion, and among them, 45% people actively use some sort of social media (e.g., Facebook, YouTube, Twitter, etc.). To describe the popularity of Twitter among young adults, Alhabash and Ma (2017) suggested that more than one-third of online adult users use Twitter. Further, several studies indicate that most Twitter users are aged 18 to 29, and the popularity of Twitter continues to grow among young adults (Basak, Sural, Ganguly & Ghosh, 2019; Gao, Cao, Li, Yao, Chen & Tang, 2019). Additionally, the use of Twitter among students is increasing quickly (Tierman, 2014).

Various authors have found that the use of social media helps students to improve academic performance (Junco, Elavsky & Heiberger, 2013; Junco, Heiberger & Loken, 2011; Lin, Hoffman & Borengasser, 2013; Mao, 2014). Other researchers, however, found that the use of social media among students could negatively effect on their academic performance (Flanigan & Babchuk, 2015; Kirschner & Karpinski, 2010). For example, Kirschner and Karpinski (2010) reported that the average user of Facebook (one of the more popular SNSs) has a lower GPA than a non-user. When students use SNS, it often takes time away from their studies; hence, SNSs can be characterized as an academic distraction (Flanigan & Babchuk, 2015; Gonzalez, Gasco & Llopis, 2019; Kirschner & Karpinski, 2010).

Previous studies have focused on either students' attitudes or behaviors towards the use of SNSs. In one instance, Kirschner and Karpinski (2010) focused on students' attitude and Twitter use, while Junco, Heiberger, and Loken (2011) focused on students' behavior and Twitter use. Attitude and behaviors relate to students' academic performance; however, each of these factors could relate differently with students' academic performance (Ajzen & Fishbein, 1977). One might have a positive relationship with academic performance whereas the other could have a negative relationship with academic performance (Lassen, Steele & Sailor, 2006; Ostroff, 1992). It thus becomes important that a study is conducted focusing on both aspects (i.e., attitude and behavior), thereby discovering their relationship with students' academic performance in terms of Twitter use.

Need for This Study

The literature review indicates that several research studies have focused on Facebook use and students' academic performance (Flanigan & Babchuk, 2015; Gonzalez, Gasco & Llopis, 2019; Junco, Elavsky & Heiberger, 2013; Junco, Heiberger & Loken, 2011; Kirschner &

Karpinski, 2010; Lin, Hoffman & Borengasser, 2013; Mao, 2014). For example, Kirschner and Karpinski (2010) and Junco (2012) focused on multitasking with Facebook and students' academic performance. Kirschner and Karpinski (2010) found that multitasking and academic performance is negatively correlated, although Junco (2012) noticed that students' academic performance improved after using Facebook. Studies found that social networking sites help students to stay in touch with family and peers and that Facebook has a positive influence on academic outcomes (Ainin, Naqshbandi, Moghavvemi & Jaafar, 2015; Junco, 2012). In a separate study, Jacobsen and Forste (2011) mentioned that multitasking while studying leads students to incomplete or inadequately done tasks. Seeing this, the use of social networking sites can be harmful for schoolwork, and it often decreases students' academic performance.

A few studies indicated that the use of social networking sites (e.g., Facebook, Pinterest, etc.) is an academic distraction (Fewkes & McCabe, 2012; Michikyan, Subramanyam & Dennis, 2015; Ozer, Karpinski & Kirschner, 2014). Those studies focused on finding relationships between students' academic performance and their academic responsibility as well as perception about the use of social networking sites (Fewkes & McCabe, 2012; Michikyan, Subramanyam & Dennis, 2015; Gonzalez, Gasco & Llopis, 2019; Michikyan, Subramanyam & Dennis, 2015; Ozer, Karpinski & Kirschner, 2014). Nonetheless, other studies reported that students have mixed beliefs about the use of social networking sites (Gonzalez, Gasco & Llopis, 2019; Michikyan, Subramanyam & Dennis, 2015). For instance, Flanigan and Babchuk (2015) posited that some students view social networking sites as an academic distraction. On the contrary, Ozer, Karpinski and Kirschner (2014) found that some students think that social networking sites helped them to improve their academic performance (Ozer, Karpinski & Kirschner, 2014).

Mixed findings were reported by those researchers who focused on students' behavior in terms of the use of social networking sites and their academic performance (Lian, et. al., 2018; Meier, Reinecke & Meltzer, 2016; Kirschner & Karpinski, 2010). Kirschner and Karpinski (2010) reported that the use of Facebook during study break brought fun to students. Happiness serves as a positive trait for students in terms of their academics; it helps to relieve stress and improve their academic performance (Villavicencio & Bernardo, 2013). Conversely, Lian et al. (2018) mentioned that the use of social networking sites during study breaks could lead students to subjective feelings and tiredness; ultimately, such feelings could be bad for their academic achievement. *Subjective feelings* can be defined as discomfort, decreased motivation, and increased physical lassitude (Lee, Son & Kim, 2016).

One of the reasons for such mixed findings in various studies is that researchers focused on different aspects of social networking sites use among students while they focused on different social traits. *Social traits* can be defined as an attitude and behavior towards someone or something (Ajzen, 1987; Singh & Teoh, 2000). Students' attitudes and behavior often relate to academic performance (Credé & Kuncel, 2008; Fuligni, 1997). Previous studies either focused on attitude-related traits or on behavior-related traits. Few to no research studies have focused on both attitude- and behavior- related traits.

Most studies have attempted to measure the constructs of social-networking sites, use, and academic achievement using different scales without considering the combination of both traits (Ahn, 2011; Amiruzzaman, Amiruzzaman, Ozer-Guclu & Karpinski, 2019; Mehmood & Taswir, 2013; Paul, Baker & Cochran, 2012). In order to focus on students' academic performance in relation to the use of social networking sites, a study should focus on blog-based (i.e., text-based) information sharing social networking sites, and both attitude and behavior traits should be considered. This is because *blog-based* social networking sites are topic oriented. On blog-based social networking sites, a user can focus and follow a specific topic, discussion, and user(s) based on their interest (Minocha, 2009). For students, blog-based social networking sites, such as Twitter and Wiki, can be quite helpful as they can learn and share information about a topic that is important to them (Grosseck & Holotescu, 2008; Kassens-Noor, 2012).

More specifically, several studies mentioned that the use Twitter in academia could help increase students' learning opportunities (Junco, Heiberger, & Loken, 2011; Rankin, 2009; Tang & Hew, 2017; Tiernan, 2014; Young, 2009). In one specific study, Rankin (2009) posited that students could use Twitter and communicate with others during class lecture. Other researchers suggested that Twitter could be an outlet for students to express their ideas and opinions (Denker, Manning, Heuett & Summers, 2018; Young, 2009). In addition, Twitter can help students to engage in class lectures and activities to aid in enhancing their knowledge (Thompson 2007; Beldarrain 2007). As such, undergraduate students' use of Twitter among could prove beneficial (Junco, Heiberger, & Loken, 2011; Tang & Hew, 2017). Even so, a specific research study is necessary to understand the use of Twitter for academic purpose and its relationship with as well as the effect on students' academic performance.

Studies that focused on students' academic performance found that students' behavior and attitude traits are the most predictive of their academic performance (Ajzen, 1987; Chun Chu & Choi, 2005; Credé & Kuncel, 2008). Fuligni (1997), for instance, mentioned that 70% variance in students' academic performance can be explained by their behavior and attitude traits after controlling for their socioeconomic backgrounds. As a result, it is important to measure the relationship between undergraduate students' academic performance and their attitude and behavior towards the use of Twitter (i.e., a blog-based or information sharing social networking site).

Based on previous studies, this study developed two scales: the first scale measures undergraduate students' attitude towards the use of Twitter for academics; and the second scale measures undergraduate students' behavior towards the use of Twitter for academics. The attitude scale is called the Twitter and Scholastic Synchronicity Scale (TSSS), and the behavioral scale is called the Twitter and Scholastic Apportionment Scale (TSAA). The TSSS scale consists of multitasking, responsibility, and perception subscales, while the TSAA scale consists of schoolwork and procrastination and leisure subscales.

Purpose of the Study

The purpose of this study was to develop scales to measure the psychometric properties of developed measure. Further, it sought to understand relationships between newly developed scales and undergraduate students' academic performance. The study endeavored to answer the following research questions:

(1) What are the psychometric properties of the newly developed TSSS items?

   i. What is the construct validity of TSSS?

   ii. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSSS?

   iii. What are item properties of the subscales of TSSS (i.e., analyze individual subscales: level of difficulty and rating scale)?

(2) What are the psychometric properties of the newly developed TSAA items?

   i. What is the construct validity of TSAA?

ii. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSAA?

iii. What are item properties of the subscales of TSAA (i.e., analyze individual subscales: level of difficulty and rating scale)?

(3) What is the relationship between Twitter attitude and behavior measures, and students' academic performance?

i. What is the relationship between TSAA (i.e., a behavioral frequency measure) and academic performance controlling other variables?

ii. What is the relationship between TSSS (i.e., an attitudinal measure) and academic performance controlling other variables?

Summary

This study developed two different scales to measure undergraduate students' attitude and behavior in terms of the use of Twitter for academic purpose. The first scale, which is known as an attitude scale, has three subscales: multitasking, responsibility, and perception. The second scale, which is known as a behavioral scale, has two subscales: (a) schoolwork and (b) procrastination and leisure.

The study used a quantitative design to collect research data and analyze the collected data. Different types of statistical analysis were used to validate and explore the psychometric properties of the newly developed scales (i.e., attitude and behavioral scales). Factor Analysis, Rasch Analysis, and Classical Test Theory (CTT) were each used to analyze the validity and reliability aspects of these scales. Finally, this study used Hierarchical Multiple Regression to discern the relationship between students' academic performance and these scales.

Organization of the Dissertation

Chapter 2 describes the literature review of this study. That chapter presents how the theme for the scales have been extracted from existing literature. Moreover, it describes how the measures were developed and presents the validity argument to analyze the validity and psychometric properties of newly developed measures.

Chapter 3 describes the design of this study. That chapter includes the purpose of the study along with research questions, data collection, and the data analysis procedures employed in the study. All necessary statistical tools and models used in the study are likewise explained in chapter 3.

Chapter 4 presents collected data and results of the data analysis. The data analysis is using different statistical tools, such as, to explore factor structure Exploratory Factor Analysis, to confirm factor structure Confirmatory Factor Analysis is used. Further, to find the evidence of unidimensionality Rasch Analysis, and to understand relationship between newly developed scales and students' academic performance Hierarchical Multiple Regression is used. Results are interpreted in light of the research questions and discussed in conjunction with other literature.

Chapter 5 justifies the research study and concludes the dissertation. The chapter presents answers to the research questions. This chapter also provides implications for researchers, practitioners, and educators; it also includes directions for future research studies and limitations of the study.

CHAPTER II

LITERATURE REVIEW

Introduction

Twitter is one of the most commonly used social networking applications (Aharony, 2010; Junco, Elavsky & Heiberger, 2013). The authors of earlier studies explained that Twitter helps to broadcast and share information via the Internet (Aharony, 2010; Junco, Elavsky & Heiberger, 2013; Junco, Heiberger & Loken, 2011). In fact, millions of people use Twitter daily to communicate and share information (Aslam, 2017; Coad, 2017; Mao, 2014; Parra, Trattner, Gómez, Hurtado, Wen & Lin, 2016).

Some researchers have found that social media use helps students improve academic performance (Junco, Elavsky & Heiberger, 2013; Junco, Heiberger & Loken, 2011; Lin, Hoffman & Borengasser, 2013; Mao, 2014). Nevertheless, other researchers suggested that students spend considerable time using social media, which affects academic performance (Flanigan & Babchuk, 2015; Kirschner & Karpinski, 2010). For instance, Kirschner and Karpinski (2010) conducted a study to examine how many college students use Facebook and how its use influences academic performance. Their findings indicated that Facebook users have a lower mean GPA than non-users due to fewer hours of studying per week (Kirschner & Karpinski, 2010).

Similarly, Flanigan and Babchuk (2015) viewed social networking technology as an academic distraction for students. These researchers had conducted an in-depth exploration of students' use of Facebook and Twitter. Students' learning and achievement served as the basis of Flanigan and Babchuk (2015)'s research study.

It is important that students communicate with each other, which help them to learn from each other (Coad, 2017; Lin, Hoffman & Borengasser, 2013; Tu, 2000). The fundamental idea of Coad's (2017) study was that those social networking sites, such as Twitter and Facebook, positively affected students' learning. Interestingly, however, a recent study on the use of social networking sites among college students revealed that when students spend too much time on social networking sites, academic performance decreases (Gonzalez, Gasco & Llopis, 2019).

Students' academic performance also relates to their attitude and behavior (Credé & Kuncel, 2008; Fuligni, 1997). What students think about Twitter use related to their academics can help to understand their attitude towards it, and how they use Twitter for academics can help to understand their behavior towards it. Because of this, it is important to investigate how this social networking application impacts students' learning. In other words, what are the different ways to measure the relationship between students' academic performance and their attitude and behavior towards a social networking site, such as Twitter?

Literature Review

As technology is rapidly evolving, Twitter, as a social networking site, is gaining more attention from students (Kim & Kim, 2019; Roopchund, Ramesh & Jaunky, 2019). This increase in attention occurs because it allows students to communicate and share information in real-time (Chang & Hsiao, 2014; Wise, Alhabash & Park, 2010). This is one of the advantages of technology-based environment. In technology-based environments, students are referred to as the "Net Generation" (Flanigan & Babchuk, 2015) or "Homo Zappiens" (Kirschner & Karpinksi, 2010). Regardless of generational references or identities, technology is a very present part of their world. With this, a discussion has emerged about the advantages and disadvantages in terms of this technological environment for students (Liu & Tsai, 2012). In addition, the continued work analyzing the consequences of social networking technology now exists (Danciu & Grosseck, 2011; Hargittai & Hsieh, 2010). The use of social networking technology can be faster and convenient, but questions arise whether it actually helps students in their academics (Ellison, Steinfield & Lampe, 2007; Pempek, Yermolayeva, & Calvert, 2009).

Kirschner and Karpinski (2010), for example, wanted to see how many college students use Facebook and how their use of Facebook impacts their academic performance, if at all. The results indicated that Facebook users have a lower mean GPA than non-users; this is most likely due to users putting in fewer hours of studying per week than non-users. The evidence of their study suggests that there is a relationship between lower GPAs and the use of social media or social networking sites among students.

Some researchers have hypothesized about students' multitasking skills (Jacobsen and Forste, 2011; Lau 2017; Rouis, 2012). Karpinski, Kirschner, Ozer, Mellott, and Ochwo (2013) defined multitasking as "the simultaneous/concurrent execution of two or more cognitive or information processing activities" (p. 1183). A few research studies have indicated that multitasking can be helpful for schoolwork (Ellison, Steinfield & Lampe, 2007; Rouis, 2017). In one instance, Rouis (2012) found that students who possessed multitasking skills were better able to manage their time and exerted less effort when it came to schoolwork. In a separate study, Kirschner and Karpinski (2010) focused on multitasking with the assumption that students can work effectively and efficiently while multitasking. Still, their findings indicated that multitasking does not lead to effectiveness and/or efficiency.

Furthermore, Kirschner and Karpinski (2010) found that people were making assumptions that multitasking is a good skill for students to have, but they discovered that multitasking leads to lack of attention. In fact, when students tried multitasking during their study or other tasks, those tasks were not being completed or were attacked in broken periods of time, leaving the work done inadequately (Jacobsen & Forste, 2011; Lau 2017). This provides opportunities for further research into multitasking and students' attention in academics.

Little proof exists that technology, specifically social media access, is advantageous for students (Cain, 2008; Pempek, Yermolayeva & Calvert, 2009). Rouis (2012) discovered that social media can be helpful for students' emotional well-being. Her study revealed that social media provides students more connections with people in a positive manner, which in turn causes students to care more about academic performance. The premise was that if students felt connected to friends and family, then they would have the motivation and desire to do well in school (Ellison, Steinfield & Lampe, 2007).

Naqshbandi, Ainin, Jaafar, and Shuib (2017) looked at the personality traits of the social media (i.e., Facebook) users. Personality factors were considered as an important component to molding students' academic performance. Students who were shy or who were considered extroverts had different types of responses or purposes of social media use. People could have been using social media (i.e., Facebook, WhatsApp, etc.) to disseminate information and others might have used it as a manner to socialize (DeAndrea, Ellison, LaRose, Steinfield & Fiore, 2012). These factors can be helpful to understand how social media affects students' academic performance.

Researchers have mixed opinions about the use of social media among students. Whereas some have viewed the use of social media as emotionally helpful for students' academics (Ellison, Steinfield & Lampe, 2007; Rouis, 2017), others have viewed the use of social media as a distraction for their academics (Flanigan & Babchuk, 2015; Moreno et al., 2011). In one study, Flanigan and Babchuk (2015) noticed that Facebook and Twitter negatively affected students' learning and achievement. Other studies correspondingly indicated that the distraction of social networking technology can adversely affect students' academic performance (Dietz & Henrich, 2014; Kirschner & Karpinski, 2010). This occurs because students feel a constant pull to stay connected socially via social media (Rouis, 2012). The temptation to stay connected might in actuality be the areas where students cannot control, and it could lead them into being less productive in their academics (Flanigan & Babchuk, 2015; Lin, Hoffman & Borengasser, 2013). Findings regarding these aspects have determined that even though most college students have continued access to social networking technology, they might not have the skills or knowledge to best use it for academics or professional purposes (Kennedy, Judd, Churchward, Gray & Krause, 2008). This includes the self-discipline to act upon the best time of using social media (Dietz & Henrich, 2014). As mentioned earlier, students often remark that they cannot help themselves, so they view their devices often to check social media updates. This lack of discipline most often occurs during instruction or homework time (Lin, Hoffman & Borengasser, 2013; Rouis, 2012).

Hargittai (2010) interviewed students who use social media often and found that those students did not feel confident about their professional computer skills. According to Lohnes and Kinzer (2007), digital-age students have enough knowledge to use technology to get around socially or leisurely, but do not know how to use computer programs sufficiently nor do they have strategies to stay off a device when appropriate (Waycott, Bennett, Kennedy, Dalgarno & Gray, 2010). A study conducted by Kennedy, Judd, Churchward, Gray and Krause (2008) indicated that digital-age students lack understanding on how to gain a technology savvy skill set. Flanigan and Kiewra (2018) found that most technology-based educational programs are being used within classrooms, but those are not being used to teach the basics of computer use.

The basics could consist of keyboarding, shortcuts, correct grammar or sentence structure, and/or appropriate communication etiquette. These basics are some of the computer skills that students need to have in their professional lives. This scenario indicates that students focus more on the use of social media skills and less on developing professional computer skills.

Another article by Moghavvemi, Sulaiman, Jaafar, and Kasem (2018) reported the usefulness of social networking technology in an academic setting. YouTube was found to be an effective tool for students. *YouTube* is a video sharing social networking site, it allows users to share their videos with thepublic or share with a specific group of people or keep it private (Halpern & Gibbs, 2013). The use of this type of social media could be helpful for students to enhance their learning situations (Brecht, 2012). It could be used to further their learning on a specific subject, such as mathematics, science, literature, and so on. This type of technology (i.e., social video sharing sites) could also be used to capture learners' attention (Fonseca, Martí, Redondo, Navarro & Sánchez, 2014).

In a later study, Lau (2017) found that students use of technology for academic purposes was not a significant predictor of students' GPA. The study further reported that students using technology for non-academic activities and social media multitasking served as a factor in students having lower GPAs. As noted above, the rate of technological advancements has paved the way for this evolving community. Liu, Kirschner, and Karpinski (2017) concluded that a negative correlation existed between students' academic performance and use of social network sites. The authors of this study found that female students tend to use social media more often than male students, and they also found a stronger relationship between lower GPA and female college students, specifically those who spend more time on social networking sites.

Interviewing students proves to be an excellent way to gain a better understanding of how technology affects students. In one study, Flannigan and Babchuk (2015) wanted to learn about undergraduate students' perceptions of social media use. The students provided that social media was tempting and caused them to falter. Students understood that their time was not well spent as they participated in social media activity. Most could comprehend that engaging in such activity took away time from studying as well as quality work time. Students admitted that they were willing to put in the additional time to be productive in their college studies so that they could also be active in the social community. Many students recognized that they would retain more information if they were not active social media users (Ahn, 2011).

The current literature provides mixed types (i.e., positive and negative relationships with students' academic performance) of findings about the use of social networking sites for students. Ahn (2011) discovered that parental concerns are growing because students are depending on social networking sites for social communications. Additionally, several studies found social networking sites as a threat to students' social communication skills (Abdulahi, Samadi & Gharleghi, 2014; Das & Sahoo, 2011; Moqbel, & Kock, 2018). While using social networking sites helps students to become good at communicating with each other in a virtual environment, they are becoming less comfortable with in-person communication (Vie, 2008). The literature indicates a constant conversation about the use of social media or social networking sites (e.g., Facebook and Twitter) among students and the impact of social networking sites on students' social lives and academic performance (Wang, Jackson, Gaskin & Wang, 2014; Raacke & Bonds-Raacke, 2008; Wang et al., 2018).

Based on the literature review, it becomes clear that little to no research has been undertaken to understand how students' attitude and behavior towards social media use, and

Twitter, in particular, effects academic achievement. Also, no study yet attempted to develop scales to measure undergraduate students' attitudes and behavior towards the use of Twitter, and neither has any study sought to understand the relationship, if any, between students' attitudes and behavior towards Twitter use and their academic performance. As such, with the help of existing studies, in this study, I developed two scales focusing on Twitter use: an attitude scale and a behavioral scale, both of which aid in understanding the relationship between students' attitudes and behavior and their academic performance. The development process of the Twitter scales is described below. As for the first step of the scale development, some themes appeared in the existing literature. The following paragraphs describe the extraction of themes from existing literature and their subsequent use in developing the Twitter scales.

Theme/Scale Development

In an article published by The British Psychological Society, Mc Mahon (2014) provided a psychological stance on why humans have a preference for social media. It identified three specific factors, namely (1) behavioral and cognitive, (2) social, and (3) and self-identity derives people towards social media (e.g. Facebook, Twitter, etc.) use; taken together, these factors are called *cyberpsychological* factors. Yet these factors are thus far not well studied to understand fully humans' appeal to online environments. Most studies focusing on cyberpsychological factors tried to explore a one or two factors at a time, not the combination of all three factors (Arain, et. al., 2013; Kuss & Griffiths, 2017; Ophir, Nass, & Wagner, 2009). Therefore, a study is needed to see relationship between students' academic performance and cyberpsychological factors.

Mc Mahon (2014) posited that humans are driven by interactive spaces in which they live, such as, home, work, and the "*third space*," which is a public space used for informal social interaction outside of home and workspace. Soukup (2006) went on further to say that humans now occupy another "*fourth space*," which is the social online interactive space in their lives. Soukup argued that this new space (i.e., *fourth space*) allows humans to engage in a myriad of activities that might otherwise not be common or acceptable in the other three spaces, and that they have foundations within the three factors: (1) behavioral and cognitive; (2) social; and (3) self-identity.

Examining a subset of the United States population, such as adolescents and young adults between the ages of 10-24 who engage more heavily than most other populations in this fourth space (Anderson & Jiang, 2018) thus becomes necessary. It has been suggested that social media takes a stronger habit-forming effect in this group than with other groups (Kuss & Griffiths, 2017). Adolescent attention is still considered underdeveloped at that stage of growth due to continued brain development, and during that time, habitual behavior is a powerful reinforcer in keeping adolescent attention (Arain, et. al., 2013). The more adolescents invest in carrying out a behavior, the more they will likely persist in repeating it. Adolescent attention can be easily grabbed and kept by social media through constant sounds, vibrations, and/or notifications. The rewards adolescents receive in this fourth space by responding repeatedly to these triggers are (1) social affirmation and validation and (2) constant engagement and no "Fear of Missing Out" (FOMO). In essence, they are indirectly influenced by others and forced to use social networking sites.

The complex lives of American adolescents who are navigating their ways through their years of schooling in K-12 and onto college, along with their increasingly active social lives, could find that it becomes a greater challenge for them to balance their attention between their educational and social needs (Rahman & Stephen, 2016; Lau, 2017). In classrooms, the use of social networking sites (SNSs) are expected to promote active learning, effective communication, and information sharing to potentially increase students' engagement and learning (Koranteng, Wiafe & Kuada, 2018). Despite this, students often find it difficult to separate their use of SNSs between strictly educational and social or find a balance of social use when the priority is educational (Junco, 2015). This can have fundamentally profound effects on students' abilities to function successfully in both educational and social environments while actively engaged with SNSs.

The studies presented in this literature review examine a variety of ways in which students are challenged in achieving academic success when balancing both educational and social use of SNSs. The following themes appeared in the existing literature related to SNSs use among students and are related to their academics.

Multitasking

A host of previous studies (Jacobsen & Forste, 2011; Junco, 2012; Karpinski, Kirschner, Ozer, Mellott & Ochwo, 2013; Ophir, Nass, & Wagner, 2009) have shown that multitasking for students' use of SNSs during study time can be harmful, as it might negatively affect academic performance. In a separate study, Karpinski, Kirschner, Ozer, Mellott, and Ochwo (2013) found that in the United States, multitasking using SNSs and students' academic performance is negatively correlated. Similarly, Lau's (2017) study confirmed that multitasking can indeed negatively impact students' academic performance. Quite interestingly, Karpinski, Kirschner, Ozer, Mellott and Ochwo (2013) noted that this is not the case for European students.

Students' use of Facebook has been found to help them in developing new connections in their transition to college, but multitasking with technologies has been shown to interfere with students' learning processes (Junco, 2012; Junco, 2015). These findings are based on the examination of relationship between multitasking and students' academic performance using their grade point averages (GPAs) as a variable (Junco, 2012). Notably, most literature used GPA as the sole measure of students' academic performance in connection with the use of social networking sites (Junco, 2012; Kirschner & Karpinski, 2010).

Junco (2012) reported that most undergraduate students (92%) spend on average of at least one hour and 40 minutes on Facebook per day. The reason students frequently use Facebook or other social networking sites is because of social communication, and they are most likely to use Facebook or other social networking sites while in class (Fewkes & McCabe, 2012; Rouis, 2012). In addition, Junco (2012) reported that Facebook use in general was a better predictor of academic outcomes rather than the amount of time students spent on their studies. Jacobsen and Forste (2011) reported similar findings, as they discovered a negative association between SNSs use and students' academic performance. Moreover, a negative correlation occurred between students' academic performances and multitasking with SNSs (Jacobsen & Forste, 2011; Lau, 2016). Junco (2012) found that Facebook use and text messaging related negatively to students' GPAs. Junco (2012) also found that male students had low GPAs compared to female students and that Caucasians had higher GPAs compared to Latinos. Junco's (2012) study revealed that students who use Facebook more often than others (i.e., social use of SNSs) are most likely to be unable to focus on their schoolwork and could end up having lower GPAs. In other words, students who multitask during class are more likely to have lower GPAs (Kirschner & Karpinski, 2010).

Hew (2011) found that graduate students spent the least amount of time on Facebook or other SNSs, as well as the least amount of time multitasking with Facebook or other SNSs, as compared to undergraduate students. In general, students interact less with Facebook or other

SNSs as they progress in their class level. Time spent on Facebook served as a significant negative predictor of GPA only for the freshman class level (Junco, 2015). Note that students who earned fewer than 30 credit hours are considered as freshman class level. Junco (2015) reported that multitasking with Facebook was a significant negative predictor of GPA for first three years of college (freshmen, sophomores, and juniors). A claim could be made that by the time students reached to senior class level, then they had a better understanding of how SNSs (e.g., Facebook, Twitter, etc.) use could adversely affect their GPA, and that limiting SNSs use could potentially have a positive effect on desired academic success. Senior students could also have become more skilled at balancing schoolwork and SNSs use.

Kirschner and Karpinski (2010) further examined the concept of multitasking and the assumption that students could work effectively and efficiently in obtaining academic success while engaged with SNSs. Ultimately, their findings showed that students could not work effectively and efficiently while multitasking. Kirschner and Karpinski (2010) found that people made assumptions about multitasking skills; still, they discovered a lack of these skills among students. In most cases, when students multitask, tasks are left entirely incomplete, or they are handled in broken periods of time leaving the work completed to be more inadequate (Jacobsen & Forste, 2011).

Responsibility

In their study, Michikyan, Subramanyam, and Dennis (2015) noted that some students believe that they are responsible for their academics and they should limit using SNSs, such as Facebook and Twitter. Nonetheless, those students also indicated that they know Facebook or other SNSs are some sort of academic distraction, and it is hard to stay away from SNSs. Nowadays, almost every student uses Facebook or some other SNSs for day-to-day information sharing (Samad, Nilashi & Ibrahim, 2019). Because of this, students seem to share mixed opinions about Facebook use and academic responsibility (Gonzalez, Gasco & Llopis, 2019; Michikyan, Subramanyam & Dennis, 2015). Some students mentioned that they avoid Facebook use during their final exams, but some students did not care that much. Ozer, Karpinski, and Kirschner (2014) mentioned that between American and European students, European students are more cautious about when and how to use SNSs and manage their time to be able to study effectively.

Kirschner and Karpinski (2010) wanted to see how many college students use Facebook and how the use of Facebook impacts their academic performance. Their study found that Facebook users invariably have lower GPAs compared to non-users. To explain this phenomenon, Hew (2011) mentioned that on average, Facebook users spend less time studying than non-users. Correspondingly, Park (2010) found that graduate students who use SNSs have a higher mean GPA compared to undergraduate students. It appears that Facebook users are more likely to be involved in extracurricular activities than non-users who work more hours per week (Kirschner & Karpinski, 2010). A few studies suggested that non-users are too busy to use Facebook or other SNSs, usually due to their busy lives (Sheldon, 2012; Wells & Link, 2014).

Perception

A study completed by Samaha and Hawi (2016) revealed that not all students view SNSs as an academic distraction, but some still perceived the use of SNSs to be harmful to academic success. Michikyan, Subramanyam, and Dennis (2015) reported that some students believed that the excessive use of SNSs caused them to have poorer academic outcomes (e.g., a lower midterm grade). Ozer, Karpinski and Kirschner (2014) noted that American students view the use of SNSs as an academic distraction, but that European students perceive it as a source of entertainment. Van Der Schuur, Baumgartner, Sumter, and Valkenburg (2015) identified a negative relationship between the use of SNSs and students' perceptions of learning gains. If students believe that the use of SNSs is not good for them and that it prevents them from learning from either textbooks or the classroom, then they try to reduce the amount of time using SNSs (Ophus & Abbitt, 2009). In contrast, some students use SNSs more often than others because they do not value their education (Samaha & Hawi, 2016).

Schoolwork

In his recent research, Lau (2017) found that the use of SNSs can be both positive and negative for students. As a positive example, it was determined that if students use SNSs to communicate with peers or instructors, they might in turn receive more informal instructions, enabling them to better understand concepts in an easier, generalized manner. As a negative example, however, research found that cyberbullying and challenging social relationships could negatively impact students' academics. According to Ozer, Karpinski, and Kirschner (2014), European students use SNSs primarily for their schoolwork compared to American students. This might contribute to European students' academic success since they seem to find SNSs as a complement to their schoolwork over its use in engaging in social activity.

In another recent study involving Malaysian students in a Teaching English as a Second Language (TESL) program, researchers set out to determine whether students' SNSs addiction influenced academic performance (Rahman & Stephen, 2016). The researchers concluded that students' usage of SNSs was moderate and hat most students connected only between one and four hours per session, even though they logged in daily. Furthermore, their findings indicated that more than half students (54.8%) possessed low addiction to SNSs, which showed that only a minority of students were addicted to SNSs. Perhaps SNSs were the means for minority students to stay connected with friends and family—SNSs were their social outlet.

Doleck, Bazelais, and Lemay (2018) examined two dimensions of student-school characteristics to explain how students' use of SNSs affected academic performance: (1) school life satisfaction and (2) school commitment. They based their research on two underlying assumptions: (1) SNSs used for leisure activity and (2) time was fungible (Doleck, Bazelais, & Lemay, 2018). Participants in the study were pre-university science students recruited from an English *Collège d' enseignement général et professionnel* (CEGEP) (i.e., general and vocational college) in Montreal, Quebec.

The results of Doleck, Bazelais, and Lemay's (2018) study concluded that school life satisfaction had a significant positive effect on school commitment, and school commitment in turn had a significant positive effect on students' academic performance. These results suggested that school commitment exerted greater influence than school life satisfaction in relation to students' academic performance. This study found overall lack of support for a key relationship between SNSs use and students' academic performance.

Procrastination and Leisure

Michikyan, Subramanyam, and Dennis (2015) revealed in their study that some students use SNSs during breaks to help them reduce stress. The researchers found that in the United States, students believed that SNSs use can contribute to academic procrastination and leisure, which can further lead to decreased academic performance (Ozer, Karpinski and Kirschner, 2014). On the other side of the Atlantic, European students believed that the use of SNSs for academic procrastination and leisure can be helpful for them in reducing academic stress.

The effects on students' academic and overall well-being from the use of SNSs was examined by researchers Meier, Reinecke, and Meltzer (2016). They concluded that self-control related negatively and significantly to the frequency of procrastination with Facebook. Habitually checking Facebook and enjoying Facebook related both positively and significantly to the frequency of procrastination with Facebook. In a separate study, Lian, et. al. (2018) examined the mediating role of social networking site (SNS) fatigue and the moderating role of effortful control among participating Chinese undergraduates. The results indicated that addiction to SNSs could significantly predict irrational procrastination in undergraduate students. In addition, communication overload caused by SNSs addiction could interrupt the users' daily tasks, making it harder to concentrate and easier to discontinue activities at hand. The researchers also concluded that SNSs fatigue was an important, underlying psychosocial mechanism in the relation between SNSs addiction and irrational procrastination. SNSs addiction could lead to a subjective feeling of tiredness from SNSs use, which could then result in irrational procrastination (Lian, et. al., 2018).

Procrastination with SNSs could harm productivity and thus lower academic achievement (Junco, 2012; Kirschner & Karpinski, 2010). According to Kirschner and Karpinski (2010), most students felt that using Facebook had a negative impact on their academic performance, and only a few students felt that using Facebook had a positive impact on academic performance. Facebook was found to cause students significant distraction from schoolwork and disrupt their abilities to manage time appropriately. The researchers did indicate that some students reported Facebook as having had a positive impact where it could be used as a networking tool to form study groups.

Facebook use has been measured in different ways, such as through a measure of time spent on Facebook (Junco, 2012) or by splitting users and non-users (Kirschner & Karpinski, 2010). Additionally, grades were measured either through self-report (Kirschner & Karpinski, 2010) or through data collected from the university registrar (Junco, 2012). Although research has been conducted on students' experience using Facebook and its impact on their academic performance, little research exists examining how Twitter relates to students' learning.

The themes above helped to recognize that two types of frequencies might have a relationship with students' academic performance, such as (1) attitudinal or agreement measure and (2) behavioral frequency measure. The following section explains the measures developed for this study.

Measure Development

Previous studies mentioned that multitasking (Rosen, Whaling, Rab, Carrier & Cheever, 2013), responsibility (Furrer et al., 2010; Steelman, Soror, Limayem & Worrell, 2012), and perception (Chaiken & Baldwin, 1981) are attitudinal factors and that schoolwork (Wahler, 1975), procrastination and leisure (Chun Chu & Choi, 2005) are behavioral factors. As such, two constructs were developed for this study to measure the causal connection between Twitter use and students' habits and thinking (i.e., attitudinal measure) and the causal connection between Twitter use and students' actions (i.e., the behavioral measure). The details of these measures appears below.

Measures/Items

The constructs and scales of this study were developed by the nomological network presented below (see Figure 1). The *nomological network* helps to understand the theoretical overview of developed measures and their relationships with each other. Cronbach and Meehl

(1955) encouraged researchers to develop a nomological network before developing any measures. On occasion, a nomological network is used to refine conceptual/theoretical relationships between dependent and independent variables.

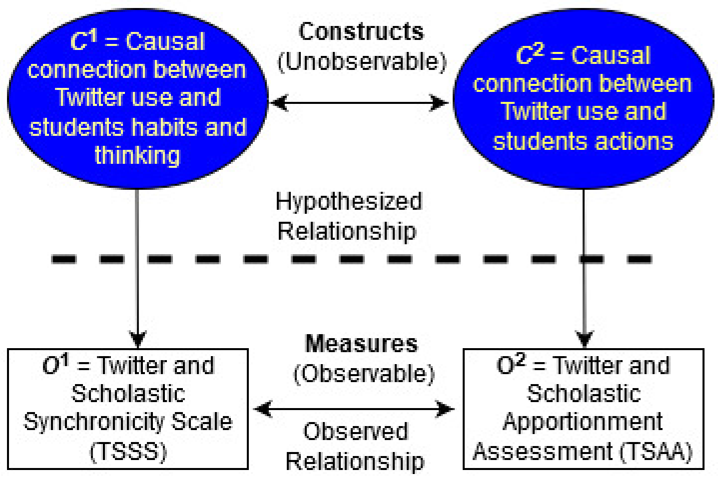

Figure 1. Constructs and scales of the study

The figure above (i.e., the nomological network) shows the measures, scales, and constructs as well as their hypothesized relationship. A *construct* in social science research can be defined as the characteristics that is being assessed using the scale (Cronbach & Meehl, 1955; Messick, 1995).

A nomological network is likewise helpful to analyze a study's construct validity (Byrne, 1984). Often, researchers use a construct key map or a construct map to develop and refine constructs (Wilson, 2004). To explain a construct map, Wilson related that "…it extends from one extreme to another, from high to low, small to large, positive to negative, or strong to weak" (p. 6). Sometimes, a construct map is referred as a unidimensional latent variable. A *latent variable* is not directly observed, but is inferred with the help of other variables, such as intelligence or satisfaction, among others. Note that a latent variable and a construct map can be multidimensional.

This study seeks to examine the psychometric properties of newly developed measures (i.e., examine the psychometric properties of TSSS and TSAA). The Table 1 provides the details. *Psychometric properties* refer to factor structure, internal consistency, and reliability (Devilly & Borkovec, 2000; Schmitt, Langan, Williams & Network, 2007). This study conducted the reliability and validity analysis of newly developed scales, following the guidelines mentioned in Crocker and Algina (2008). For instance, construct validity, content validity, and criterion validity analyze the validity of the scales, and internal consistency analyzes the reliability of the scales. These validity and reliability analysis procedures are explained in detail in later sections.

As a part of construct validity, this study developed measures based on theories and findings of various existing studies (Junco, 2012; Kirschner & Karpinski, 2010; Ozer, Karpinski, & Kirschner, 2014; Rosen, Carrier & Cheever, 2013). There are two measures used in this study, and details of both measures are described directly.

Twitter and Scholastic Synchronicity Scale (TSSS)

As mentioned earlier, authors of various studies indicated that social networking sites (SNSs) impact students' academic performance (Junco, 2012; Kirschner & Karpinski, 2010; Ozer, Karpinski, & Kirschner, 2014; Rosen, Carrier & Cheever, 2013). Conversely, Junco (2012) reported that the impact of SNSs is not same for graduate and undergraduate students. Further, the difference in use of SNSs and academic performance was observed among undergraduate students (e.g., African American and Caucasian students). Seeing this, further investigation is required to measure impacts of SNSs on undergraduate students' scholastic synchronicity.

SNSs comprise a broad area and not all social media might have similar influence in students' academic performance (Ozer, Karpinski & Kirschner, 2014; Rosen, Carrier & Cheever, 2013). Moreover, graduate and undergraduate students' study habits and social interactions often differ. A study is thus needed to focus on a specific SNS, rather than focusing on the use of SNSs in general. Considering the popularity of Twitter among young adults (i.e., undergraduate students), a scale was developed to measure undergraduate students' attitudes towards the use of Twitter for academic purposes. In particular, the developed scale focuses on students' multitasking ability, sense of responsibility, and perception of Twitter use for their academics.

Twitter and Scholastic Apportionment Scale (TSAA)

As indicated, Junco (2012) reported that students' academic performance improved after using SNSs. SNSs helped students to stay in touch with family and peers, and FB made a positive influence in their academic results. Similarly, Ozer, Karpinski, and Kirschner (2014) reported that European students believed that the use of SNSs helped them to stay focused and helped them to improve their grades. These studies suggested that further research is needed to measure students' behavior towards SNSs and explore the relationship between students' behavior towards SNSs and its relationship with their academic performance. As a result, it is important that a study is undertaken to explore a specific SNS (i.e., Twitter) and scholastic apportionment on undergraduate students. With that in mind, this study focused on the social networking site Twitter and developed a measure to understand undergraduate students' behavior towards the use of Twitter for their academics. Hence, the TSAA scale was developed.

Item Characteristics

The conceptual model for the measure includes two scales: (1) the Twitter and Scholastic Synchronicity Scale (TSSS) and (2) the Twitter and Scholastic Apportionment Assessment (TSAA). The first scale (TSSS) intends to measure the causal connection between Twitter use and students' habits and thinking, while the second scale (i.e., TSAA) endeavors to measure the causal connection between Twitter use and students' actions. Based on existing literature, each of these scales were further divided into sub-scales (see Table 1).

Previous studies (Junco, 2012; Kirschner & Karpinski, 2010; Ozer, Karpinski, & Kirschner, 2014; Rosen, Carrier & Cheever, 2013) aided in developing measures of this study. Survey items were developed and grouped into smaller categories or sub-scales in Twitter and Scholastic Synchronicity Scale (TSSS) as well as the Twitter and Scholastic Apportionment Assessment (TSAA) scales. There are two groups of items preliminarily titled (1) the TSSS and (2) the TSAA. As mentioned, multitasking, responsibility, and perception are attitude-related factors, so they are grouped together in the TSSS scale. Schoolwork, procrastination, and leisure are behavior-related factors; resultantly, these two factors are grouped together in the TSAA scale (see Table 1).

Table 1

Scales, Sub-scales, and Items Descriptions

Measure Name	Scale	Items
Twitter and Scholastic Synchronicity Scale (TSSS)	Multitasking (TSSS_M)	I'm good at multitasking with Twitter (M1).
		I multitask with Twitter while studying (M2).
		I have Twitter up while doing homework (M3).
	Responsibility (TSSS_R)	My academics are my main focus (R1).
		I'm a responsible person about schoolwork (R2).
		I only use Twitter when I have the time for it (R3).
	Perception (TSSS_P)	Twitter are time consuming (P1).
		Twitter is an academic distractions (P2).
		Twitter decreases academic performance (P3).

		Twitter takes time away from studying (P4).
Twitter and Scholastic Apportionment Assessment (TSAA)	Schoolwork (TSAA_SW)	I use Twitter for schoolwork (SW1).
		I use Twitter to communicate with my classmates (SW2).
		I use Twitter to communicate for group projects (SW3).
	Procrastination and Leisure (TSAA_PL)	I use Twitter as a break while studying (PL1).
		I use Twitter to procrastinate when I should be studying (PL2).
		I use Twitter to procrastinate if I am struggling/get bored (PL3).
		I use Twitter as a free time activity (PL4).

For the TSSS, which is an attitudinal measure, three items pertain to multitasking, three items involve the constructs of responsibility, and four items inquire as to general perceptions of Twitter (n = 10 items, see Table 1). All items used a 5-point Likert scale indicating participant appraisal of his/her multitasking ability, responsibility involving their academic behaviors and Twitter use, and their perceptions of Twitter in the context of their academic lives (Strongly Disagree = 1, Disagree = 2, Neutral/Mixed Feeling = 3, Agree = 4, and Strongly Agree = 5). Similarly, for the TSAA, which is a behavioral frequency measure, there are three items pertaining to schoolwork and four items involving the construct of leisure uses of Twitter (n = 7 items, see Table 1). Each item used a 5-point Likert scale indicating how often participants use their Twitter while attending to school-related activities or during leisure time (Very Rarely = 1, Rarely = 2, Sometimes = 3, Often = 4, and Very Often = 5).

Construct map

After analyzing the measures/scales, construct maps were developed. A *construct map* is a visual representation of items and their observational range (Wilson, 2004). Construct maps help to formalize the conceptual idea of a measure that is presented using a nomological network. A nomological network presents the brief idea about a construct or measure, and construct map explains how the measure will to evaluate a scenario.

Construct Map for Twitter and Scholastic Synchronicity Scale (TSSS)

Construct maps for this study appear below (see Table 2). As indicated, the TSSS scale has three subscales: multitasking; responsibility; and perception. Each subscale intends to measure undergraduate students' attitudes towards the use of Twitter for their academics. Because of this, three construct maps were developed to understand the outcome of these subscales.

Construct map for multitasking. The multitasking and educational outcomes are related (Jacobsen & Forste, 2011; Junco, 2012). Multitasking divides students' attention when they switch tasks while learning (Junco, 2012; Lau, 2017). For example, if a student watches TV while trying to read a book and the student is often switching his/her attention between these tasks, which amounts to multitasking.

The multitasking subscale can be used to measure students' attitudes towards multitasking with Twitter while studying. With this, the answers to this subscale should vary from strongly disagree to strongly agree. In the present study, strongly agree would indicate that students frequently multitask during study time (see Table 2).

Construct map for responsibility. Academic responsibility helps students to pay more attention towards their schoolwork (Wentzel, 1991) and helps to create an environment that is friendlier towards students' cognitive development (Michikyan, Subramanyam & Dennis, 2015). Lack of academic responsibility could lead to irresponsible behavior and poor interpersonal relationships, which can in turn cause academic disruptions and might lead to poor academic performance (Wentzel, 1991).

The answers for the responsibility subscale (i.e.., academic responsibility) varied from strongly disagree to strongly agree about academic responsibility. Of these options, strongly agree indicated that students' positive attitude towards academic responsibility (see Table 2).

Construct map for perception. The goal of the perception subscale is to measure or capture students' views toward the use of Twitter (i.e., whether students see Twitter is good for their academics or not). The literature indicates that students' perceptions or beliefs could impacts their academic achievements (Paul, Baker & Cochran, 2012). Hence, this study intended to measure students' perceptions towards Twitter and finds the relationship between such perceptions and academic performance.

More specifically, the perception subscale would uncover how undergraduate students view Twitter use in terms of academic work, such as assignments, homework, or group projects, for instance. Seeing this, favorable view towards Twitter use while studying (i.e., strongly agree) and unfavorable view towards Twitter while studying (i.e., strongly disagree) are to be measured using this subscale (see Table 2).

Table 2

Construct map for TSSS subscales

Twitter and Scholastic Synchronicity Scale (TSSS)			
Level	Multitasking	Responsibility	Perception
High	• Engaging other activities while using Twitter.	• Students believe their study should be the priority.	• Students believe that Twitter is an academic destruction.
Low	• No task switching during study time.	• Students think that their study should not be the priority.	• Students think Twitter has nothing to do with their academic performance.

Construct Map for Twitter and Scholastic Apportionment Assessment (TSAA)

Several studies described positive and negative effects on students' academic achievement because of the use of social networking sites (Forbush & Foucault-Welles, 2016; Junco, 2012; Kirschner & Karpinski, 2010; Rosen, Carrier, & Cheever, 2013). But the effects of social networking sites on students' academic achievement will likely depend on how students use social networking sites in their day-to-day lives (see Table 3).

Construct map for schoolwork. This subscale seeks to measure how students use Twitter for their schoolwork or for their academics. An earlier study indicated that some undergraduate students use social networking sites for schoolwork (Subrahmanyam, Reich, Waechter & Espinoza, 2008).

The schoolwork subscale intended to measure how much undergraduate students use Twitter for academic purposes (from strongly disagree to strongly agree). Strongly agree meant that they use Twitter for their academic purpose more frequently (see Table 3).

Construct map for procrastination and leisure. Several studies (Junco, 2012; Kirschner & Karpinski, 2010; Meier, Reinecke, & Meltzer, 2016; Rosen, Carrier, & Cheever, 2013) indicated that procrastination using social networking sites could cause lower academic achievement. The subscale thus measured "procrastination and leisure" using Twitter, or how much undergraduate students use Twitter during break time.

The answers to these subscales range from strongly disagree to strongly agree. Strongly agree meant that students use Twitter to take a break from studying more often than others (see Table 3).

Table 3

Constract map for TSAA subscales

Twitter and Scholastic Apportionment Assessment (TSAA)		
Level	Schoolwork	Procrastination and Leisure
High	• Often use Twitter for school related work.	• Use twitter during study break.
Low	• Little use of Twitter for schoolwork.	• Twitter is not a mean of procrastination and leisure.

Validity Argument

A validity argument can help to provide evidence to evaluate newly developed scales or measures (Kane, 2012). To explain the importance of validity argument, Cronbach (1988) mentioned that a validity argument helps to make things clear in terms of evaluation of a test and to make inferences from the evidences. Kosko (2019) mentioned that a sound validity argument integrates various validity analyses and pieces of evidence to support those validity analyses. Several validity analyses used to evaluate the validity of this study appear in the Table 4, along with what types of evidence were analyzed to support the validity argument.

A measure is only valid if it actually measures what it supposed to measure (Crocker and Algina, 2008). In order to understand how good or bad a measure is, psychometric properties must be analyzed (Brink, Louw & Grimmer-Somers, 2011; Karanicolas et al., 2009). To explain the importance of the analysis of psychometric properties, Brink, Louw and Grimmer-Somers (2011) wrote, "*Psychometric properties* include *validity*, *reliability* and sensitivity to change…" (p. 1) (emphasis original). Karanicolas et al. (2009) listed validity and reliability analysis as two core phases of psychometric properties analysis.

Validation of a statistical measure helps to make a justified decision about the chosen measure (Kane, 2006; Kane, 2011). It is not possible to validate a statistical tool or measure considering all aspects (Anastasi, 1986), but it is important to validate a measure considering the research questions associated with the measure or statistical tool (Kane, 2011). The current study followed the guidelines outlined in the Table 4. Evidence of different statistical analyses were carefully examined to justify the outcome of each validation test.

Table 4

Summary of Validity Evidences, by Form, Provided in Current Study

Form of validity	Terminology based on Wolfe and Smith (2007) and Kosko (2019)	Evidence in current study
Test content	Content: Instrument Purpose Content: Test Specifications & Item Development Content: Item Technical Quality	• Construct map/nomological network • Item table (Scale and subscale and items) Mean square fit statistics (Rasch)
Response process	Sustentative: Behavioral Observation	• Analysis of responses (FA—EFA, and CFA) • Person mean square fit statistics • Wright map
Internal structure	Structural: Unidimensionality Interpretability: Person-Item Maps	• Infit & outfit statistics • PCA of standardized residuals
Generalization	Generalizability: Reliability, Internal Consistency	• Cronbach Alpha • Item separation reliability • Person separation reliability

Note. This table is designed based on the work of Kosko (2019) as well as Wolfe and Smith (2007), and terminologies were adopted from Kane (2011).

Content validity refers to the rationality of the theoretical base of a measure or scale. As a result, content validity can be assessed by analyzing the development of the measure or scale (Crocker & Algina, 2008). In this study, the existing literature greatly aided in the development of the employed measures and scales, which indicates the theoretical base or rationale for developing such measures. Also, the nomological network and construct map, item table (i.e., scale and subscale), and acceptable Rasch Analysis (i.e., in-fit and out-fit) indices can be used as evidence of content validity (Kosko, 2019).

In this study, I developed two new measures or scales; it is thus important to explore the underlying properties of factors. Whether items are properly grouped together can be assessed by Factor Analysis (FA). Exploratory Factor Analysis (EFA) can be utilized to explore and understand the underlying factor structure (Reuterberg & Gustafsson, 1992; Thompson, 2004). Once the factors are explored with the help of EFA, a Confirmatory Factor Analysis (CFA) can confirm the factor structure (Vassend & Skrondal, 1997). In this study, both EFA and CFA were used to explore and confirm the factor structures of the developed measures. Factor structure presented by EFA (i.e., pattern matrix or factor matrix) can be employed as evidence of response process validation. What is more, results obtained from CFA analysis can be used to support the evidence or confirm it.

Rasch Analysis (RA) helps to understand whether the rating scales used in a measure prove appropriate (Brentani & Golia, 2007). RA uses respondents' raw test scores to check performance on the measure. It helps to make the necessary correction of a scale, such as rating scales. RA involves overall psychometric properties analysis (Boone, 2016). For example, RA checks the performance of a measure or a scale based on item difficulty as well as person ability. This study used the RA to analyze the overall psychometric properties of the newly developed scales. Infit and outfit statistics for item and person were used as evidence of internal structure validation (Kosko, 2019). As for the evidence of unidimensionality, acceptable range of infit and outfit was considered.

Furthermore, Cronbach Alpha or Coefficient Alpha, item separation, and person separation from Rasch Analysis can be used to check the generalization validity (Kosko, 2019).

To describe the acceptable range of person separation, Duncan, Bode, Lai and Perera (2003) wrote, "[a] person separation index of 1.50 represents an acceptable level of separation, an index of 2.00 represents a good level of separation, and [an] index of 3.00 represents an excellent level of separation" (p. 593). The items separation index range is the same as the person separation index (Franchignoni, Giordano, Ferriero, Orlandini, Amoresano & Perucca, 2007). A higher value of item and person separation index is better. In addition, item separation and person separation reliability were considered for this study. Reliability value close to 1 is better, whereas close to 0 is undesirable (Smith, 1986). Details of these validity and their analysis processes are described in the following chapter.

Keith (2015) and Petrocelli (2003) mentioned that Hierarchical (sequential) Multiple Regression (HMR) helps to determine which variables serve as important influences on the outcome variable. With this, HMR can be used to understand the relationship between students' academic performance and developed measures (i.e., TSSS and TSAA). Similarly, the HMR was helpful to understand influences of other variables, such as gender, undergraduate level (i.e., freshman, sophomore, junior, and senior), mother's education, etc. One reason for using hierarchical regression is that it detects whether some new variables improve the prediction of outcome over and above an existing set of variables (Keith, 2015).

In this study, I sought to explore the psychometric properties of newly developed scales TSSS and TSAA. The following research questions were investigated in this study:

(1) What are the psychometric properties of the newly developed TSSS items?

(2) What are the psychometric properties of the newly developed TSAA items?

In light of these questions, EFA, CFA, RA, and Coefficient Alpha can be used to explore and understand the validity and reliability aspects of the scales. Further, the objective of this study is to understand the relationship between students' academic achievement and these scales (i.e., TSSS and TSAA). Resultantly, a third research question for this study was investigated:

(3) What is the relationship between Twitter attitude measure and behavior measure and students' academic performance?

HMR is a good statistical analysis by which to understand the investigated relationship (i.e., the relationship between undergraduate students' academic performance and developed measures). In turn, HMR helped to answer the third research question.

Summary

This chapter presents a detailed literature review to help understand the need for this study and a step-by-step procedure of the development of the attitude and behavioral measures. This chapter also presents validity and reliability arguments to display how different statistical tools can be used to check the psychometric properties of the newly developed scales. In this study, validity and reliability argument included (1) EFA and CFA analysis to explore and confirm the underlying factors of these scales, (2) RA analysis to check the overall validity, (3) Coefficient Alpha analysis to check the internal consistency of these scales, and (4) HMR to find the relationship between these scales and students' academic performance. The next chapter will describe the methodological aspects of this study.

CHAPTER III

METHODOLOGY

Introduction

The purpose of this study was to explore psychometric properties of the newly developed Twitter and Scholastic Synchronicity Scale (TSSS) and Twitter and Scholastic Apportionment Assessment (TSAA) items and understand the relationship between TSSS and TSAA on undergraduate students' academic performance. This study tried to answer the following research questions:

(1) What are the psychometric properties of the newly developed TSSS items?

   i. What is the construct validity of TSSS?

   ii. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSSS?

   iii. What are item properties of the subscales of TSSS (i.e., analyze individual subscales: level of difficulty and rating scale)?

(2) What are the psychometric properties of the newly developed TSAA items?

   i. What is the construct validity of TSAA?

   ii. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSAA?

   iii. What are item properties of the subscales of TSAA (i.e., analyze individual subscales: level of difficulty and rating scale)?

(3) What is the relationship between Twitter attitude and behavior measures, and students' academic performance?

i. What is the relationship between TSAA (i.e., a behavioral frequency measure) and academic performance controlling other variables?

ii. What is the relationship between TSSS (i.e., an attitudinal measure) and academic performance controlling other variables?

Figure 2 provides the research methodology. The first box (on the left) presents the data collection procedure. The second box (in the middle) presents different analysis techniques used to analyze the research data, and the third box (on the right) presents the results of this study (see Figure 2).

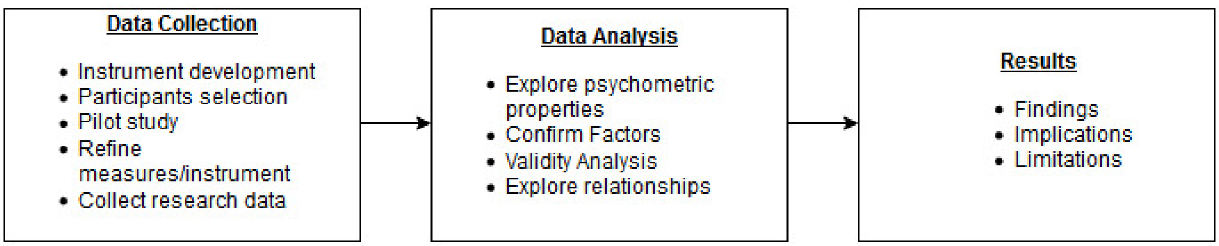

Figure 2. Block diagram of the research method (describes the overview of steps taken to conduct this study)

As to data collection, the instrument was developed after measure development, and then a pilot study took place to receive initial feedback from participants. Items of the measures and the instrument were updated based on the feedback, and then the research data were collected using the updated instrument. The data analysis procedure consisted of Exploratory Factor Analysis, Confirmatory Factor Analysis, Rasch Analysis, and Hierarchical Multiple Regression analysis.

Data Collection Procedure

Instrument Development

The existing literature aided in developing the measures of this study. Items of this study were developed from the ground up and soon revised, based on feedback from the pilot study. Using the developed items, an online survey instrument was designed by way of an online survey platform. Most items in this survey instrument made use of a Likert scale, and closed-ended questions comprised the items. Fowler (2002) mentioned the advantage of closed-ended questions; he wrote, "… ease of response is a priority to maximize returns" (p. 62). Most questions had five to six different options, which helped to increase the variability (see Table 5). Some questions appeared where participants had the option to fill them in just in the event that participants did not think that the available options were enough for them (see Table 5). This survey was implemented in a way that provided an ease to participants. Often, participants might not like to participate in a study if they think that available options do not apply to them; in that case, a fill-in-the-blank the options proved helpful for them (Fowler, 2002).

Table 5

Sample items for the instrument

Item	Options with Accompanying Codes
My academics are my main focus.	(0) Strongly Disagree (1) Disagree (2) Neutral/Mixed Feeling (3) Agree (4) Strongly Agree
What is your mother's/legal guardian's highest level of education?	(0) No High School Diploma (1) High School Diploma (2) Some College (3) Bachelor's Degree (4) Master's Degree (5) Professional Degrees (e.g., Law, Dentistry, Pharmacy) (6) Doctorate (Ph.D.) or Medical (M.D.) (7) Not sure (8) Other_____.

Pilot Study

Items of the associated scales and subscales were modified based on pilot study feedback. A total of 21 participants participated in the pilot study. Students' ages ranged from 18 to 36 (M = 25.62, SD = 5.581). Females made up 61.9% and males accounted for 38.1% of participants. For ethnicity/race, the largest proportion of students were Caucasian (48.1%) followed by Asian (28.6%) and Black (25.3%). Finally, most respondents were Social Studies majors (n = 3; 14.3%). Participants in the pilot study mainly focused on whether items convey the meaning the researcher tried to convey and if any questions confused them. Confusing questions could lead to biased research results and reduce study validity and reliability (Allen & Yen, 1979).

Two (out of 21) participants of the pilot study mentioned that a few items were more general and did not align with the purpose of the study (see Table 6). For example, one mentioned that "Twitter is time consuming." Still, considering that only one participant out of

21 made such a statement, those items were not removed from the measures or the survey instrument before collecting the research data. Interestingly almost all participants (15 out of 21) mentioned that "I am a responsible person about schoolwork" did not fit with the measure. One of the participants wrote about that item, "I am confused about what you meant by this question…this [question] has nothing to do with Twitter." Because of this, the item was removed before collecting the research data.

Fowler (2002) explained the importance of having good survey questions, relating that "Good question and instrument evaluation prior to actually doing a survey is a critical part of good survey practice" (p. 116). It remains important for researchers to learn from their mistakes and ensure that participants can understand the survey. With this, items of this survey were modified based on the pilot study feedback.

Table 6

Problematic items in associated scales and subscales

Scales	Subscales	Items
Twitter and Scholastic Synchronicity Scale (TSSS)	Multitasking	I am good at multitasking with Twitter.
	Responsibility	I only use Twitter when I have the time for it.
	Perception	Twitter is time consuming.
Twitter and Scholastic Apportionment Assessment (TSAA)	School Work	I am a responsible person about schoolwork.

Participants

Population. College students served as the target population of this survey. In particular, all undergraduate (i.e., Associate's or Bachelor's degree) students could have participated. Undergraduate students currently enrolled in an undergraduate program were considered. More specifically, students who have a Twitter account and are using their account on a day-to-day basis comprised the target population. For this study, diversity of participants was desired, with students of different genders, races, and ethnicities encouraged to participate. Also, students with different majors were expected, allowing the survey's distribution via various departmental listserv and student organizations listserv. As to gender, male and female options appeared, as was an option "other" to fill in, so that students could identify with another gender type. Participants completed a series of questions, which was grouped into two scales.

Sample. As for the sample, students from a public university in the Midwest United States (U.S.) were invited to participate. To increase the number of participants in the final study, a few instructors and chairs of different departments were asked to share the invitation email using their student listserv. This study solely focuses on current college students and their Twitter use; those who graduated were not considered. Inclusion of different social groups was expected, and indeed the collected data reflects that diversity.

Participant data ($N = 327$) collection occurred from undergraduate students using Qualtrics (i.e., a survey-hosting website). Total $N = 373$ undergraduate students participated in the study; however, only 87.6% (327 out of 373) indicated that they use Twitter. Students' ages ranged from 18 to 45 ($M = 20.90$, $SD = 3.64$). There were 65.7% females, 32.1% males, and 2.2% identified themselves as other (transgender, genderqueer, agender, bisexual). For ethnicity/race ($N = 327$), the largest proportion of students were 82.8% Caucasian ($n = 271$) followed by 10.5% Asian and 6.7% Black. As to academic discipline, most respondents were Social Science majors (25.8%), followed by "Other" (25.5%), Natural Sciences (19.9%), Business (16.9%), Humanities (7.7%), and Engineering (3.1%).

Research Data Collection and Administration

A survey method was used to collect the research data to investigate how college students used Twitter in their day-to-day lives for academic purposes. An online survey instrument was designed using Qualtrics (i.e., an online survey platform) to collect the research data. A campus-wide email was sent to staff members with access to students' email, then those staff members sent the email invitation to their listserv. The invitation included the Institutional Review Board (IRB) approval number and contact information of investigators. It was noted in the invitation email that all participation would occur on a volunteer basis and was anonymous. As such, no identification information was collected from student participants. In addition, the invitation included the survey weblink so that participants could complete the survey using the weblink. A reminder email was sent a week after the first email was sent. The survey link was kept live (i.e., open for participation) for one month, after which the link became disabled.

As participants completed the survey, they were notified of their progress. A thank-you note was provided upon completion. To maintain the quality of the survey, the Qualtrics system checked and prevented participants from answering survey questions twice. This feature helped to prevent malicious attacks or spam submissions. Before participants began the online survey, the first screen of the survey reminded them about the purpose of the survey and provided brief directions to guide them in completing the survey.

Administration. This was an online survey, causing it to be self-administered. But as mentioned earlier, directions and purpose of the study appeared at the beginning of this questionnaire. Participants thus had some basic ideas about this survey goal and what to expect from it. In addition, an automated progress bar was provided, which helped participants to know where they are in terms of completing the survey. In case anyone had questions regarding the survey or wanted to know more about it, the contact information of the investigators likewise appeared at the beginning of the survey. Couper, Traugott, and Lamias (2001) cautioned that one drawback of online surveys is that they do not allow participants to skip any questions (i.e., even if some questions do not apply to someone). Seeing this, logics were added to move participants to different blocks of the survey instead of forcing them to answer all questions.

Data Analysis Procedure

The following sections outline how data analysis procedures were used to analyze the research data. First, Factor Analysis (FA) was used to uncover the underlying relationship between items using an Exploratory Factor Analysis (EFA). Then Confirmatory Factor Analysis (CFA) was used to check how close the conceptual model was to the hypothesized model. Second, validity aspects of the measure were explored using Classical Test Theory (CTT) and Rasch Analysis (RA). Third, Hierarchical Multiple Regression (HMR) was used to understand the relationship between students' academics and newly developed scales.

Factor Analysis

Factor Analysis is a statistical procedure used to find a small set of unobserved variables (e.g., constructs, latent variables, or factors) that can account for the covariance among a larger set of observed variables (i.e., indicators). Two types underlie the broad statistical family of Factor Analysis (Thompson, 2004). Exploratory Factor Analysis (EFA) is a data-driven approach that aims to discover the factor structure of an instrument (Pett, Lackey, & Sullivan, 2003; Reuterberg & Gustafsson, 1992; Tabachnick & Fidell, 2001). Confirmatory Factor Analysis (CFA) is a theory-driven approach that aims to confirm hypothesized factor structures (Dimitrov, 2013).

There are two major classes for Factor Analysis: (1) Exploratory Factor Analysis (EFA) and (2) Confirmatory Factor Analysis (CFA). EFA helps to determine the number of factors and which observed variables are indicators of each latent variable. EFA is a process primarily used for generating theory and exploration of the data. In this step, the researcher does not have any specific expectations about the number or nature of underlying construct. Also, EFA models can use cross-loadings of indicators, and the errors are assumed to be uncorrelated. In addition, EFA requires a few assumptions, such as homogeneous sample and outliers (Thompson, 2004). For instance, a researcher should start using EFA before CFA if he/she develops the items for the first time. That way, EFA can be helpful to classify items that correlate to each other. Without satisfying the assumptions of FA, analyzed results might not be valid or reliable (Reuterberg & Gustafsson, 1992).

CFA is known as the multivariate statistical procedure (Thompson, 2004). CFA is a theory-testing process using previous theories and results obtained from the EFA (Tinsley, & Tinsley, 1987). In general, CFA is based on a theory and can specify the number of factors and the pattern of indicator factor loadings (Dimitrov, 2013). Under CFA, the variance-covariance matrix of unstandardized variables is analyzed, after which the solution for factor is evaluated to see how well it reproduces the sample covariance matrix of measured variables.

It has been established that a researcher should start with CFA before EFA if he/she develops items based on the literature or existing items (Tinsley, & Tinsley, 1987). If strong evidence appears regarding the construct being examined from previous theory or literature, CFA might be suggested to test the theory of a construct because it is used to determine the model fit (Tinsley, & Tinsley, 1987; Reuterberg & Gustafsson, 1992). On the contrary, if no previous theory exists regarding the construct to be assessed, EFA could be the best option to extract the factors to develop parsimonious instruments (Reuterberg & Gustafsson, 1992). In sum, the interrelationships among variables (i.e., items) will be examined to extract factors in EFA (Thompson, 2004).

The sample was divided into two parts: the first part with 60% participants data for the EFA, and the second part with 40% participants data for CFA. The split-sample is a cross-validation strategy, which helps to increase validation of the analyzed results (Vassend & Skrondal, 1997). The split took place by choosing participants' records randomly. The sub-sample size was determined by using the rule of thumb explained by Gorsuch (1997). According to Gorsuch (1997), EFA requires at least 10-case-per-item, whereas CFA requires less than 10-case-per-item.

Explanatory factor analysis. Data were analyzed using SPSS version 24.0. Descriptive statistics from Exploratory Factor Analysis (EFA) are included. The factor structure and internal consistency reliability (i.e., Coefficient Alpha [α]) for the overall measure and subscales were examined. For both TSSS and TSAA, the factorability of the items was assessed using Principal Axis Factoring (PAF) with Direct Oblimin rotation ($\delta = 0$). PAF is often used in social science research studies as it works better fit to correlated items (De Winter, & Dodou, 2012). Since this is a social science research study, PAF was a good fit as an extraction method. Additionally, it was assumed that items are correlated; therefore, Direct Oblimin rotation was used. For Oblique items (i.e., correlated items), Oblimin or Direct Oblimin rotation is better as it attempts to satisfy simple structure through a parameter called delta (δ), which controls the degree of obliqueness (Jennrich & Sampson, 1966).

First, the Kaiser-Meyer-Olkin (KMO) Measure of Sample Adequacy and Bartlett's Test of Sphericity were examined to investigate the appropriateness of the data for EFA. Bartlett's test helps to identify if the correlation matrix is an identified matrix (Bartlett, 1950). A significant result of Bartlett's test indicates that correlation matrix is not an identity matrix (Hayton, Allen & Scarpello, 2004), which is a desired scenario in factor analysis in the Sample

Adequacy index, also known as Factorial Simplicity index (Kaiser, 1974), the index value varies from 0 to 1. Kaiser (1974) mentioned that a Factorial Simplicity index above .60 is acceptable for factor analysis. Additionally, eigenvalues and scree plots were examined to determine the number of factors to be extracted. Eigenvalues prove helpful in determining which items are good for a factor, and an eigenvalue greater than 1.0 should be retained in factor analysis (Tinsley, & Tinsley, 1987). Nevertheless, the scree plots allow to visually decide how many factors should be retained (Hayton, Allen & Scarpello, 2004). The place where a scree plot starts to shape like an elbow that is the cut-off point for factor retention.

Confirmatory factor analysis. For purposes of this study, CFA was conducted in order to provide confirmatory evidence of the factors explored using EFA. In this phase, approximately 40% of the data was used to run the CFA analysis. After running the test, different goodness-of-fit indices were checked to see how close the conceptual model and hypothesized model was. The Browne's Chi-Square (χ^2) statistics was used to check how closely the hypothesized model fits with the research data (Curran, West & Finch, 1996). The χ^2 is one of the widely-used, goodness-of-fit indeces (Anderson & Gerbing, 1984); however, this index value depends on sample size (Raykov, 1998). A non-significant χ^2 result indicates that the hypothesized model fits the data (Marsh & Hocevar, 1985). Other goodness-of-fit indices, such as Root Mean Square Error of Approximation (RMSEA), Root Mean-square Residual (RMR), Standard Root Mean-square Residual (SRMR), Goodness-of-Fit Index (GFI), and Adjusted Goodness-of-Fit Index (AGFI) were analyzed as well.

RMSEA index is independent of sample size (Anderson & Gerbing, 1984; Raykov, 1998); therefore, along with χ^2 index, it is good to check RMSEA index (Raykov, 1998), An RMSEA index close to 0 is preferred. RMR and SRMR indices are good to check since no known limitations of these indices exist (Anderson & Gerbing, 1984), RMR and SRMR value smaller than .05 is better. Jöreskog and Sörbom (1989) mentioned that GFI and AGFI (i.e., adjusted for the degree of freedom) help to explain how much variance the hypothesized model can explain, and the value of GFI and AGFI index ranged from 0 to 1 (closer to 1 is better).

Commonly, some models might not be able to satisfy all indices (Anderson & Gerbing, 1984; Curran, West & Finch, 1996; Raykov, 1998); should this occur, the voting approach was used to determine the strength of the hypothesized model. Moreover, if any issues in the hypothesized model (i.e., problem to converge) emerged, then modification suggestions were used. Despite this, only one change was made at a time, and then results were checked to see if further modification would be necessary.

Validity and Reliability Analysis

This section explains the analysis of the validity and reliability indices. Classical Test Theory (CTT) methods were employed to further examine the items. To check for possible problematic items, item means and standard deviations for each measure were examined. For this study, validity and reliability were checked. In terms of validity, (1) content validity, (2) criterion validity, and (3) construct validity were analyzed. In addition, measurement of the reliability of this study took place using the Coefficient Alpha.

Evaluation of validity indices. Cronbach (1971) defined validation of a research study as a process, which helps a test developer or researcher to collect evidence, examine it, and verify the evidence to see whether inferences exist that could affect test scores. To describe validity studies, Crocker and Algina (2008) provided three major types of validity studies, namely (1) content validity, (2) criterion validity, and (3) construct validity.

Content validity. For this study, each item was carefully developed to ensure that it fits with the content of the measure. First, development of definitions and constructs. Second, items were developed considering the definitions and constructs. In addition, the relevance of each item was examined. The design of each question enabled the exploration of students' behaviors and attitudes towards Twitter. The pilot study was helpful to understand whether the items express what the measures intend to measure.

Criterion validity. Three questions helped to understand how students use Twitter to communicate with their classmates, work with group members for class projects and homework. Likewise, another four questions helped to measure how students are using Twitter for their leisure during their study break. Allen and Yen (1979) explained, "…test scores can be related to a criterion. The criterion is some behavior that the test scores are used to predict" (p. 97). Due to this, the pilot study took place and data from the pilot study were analyzed for scores to predict hypothesized behavior.

Construct validity. According to Allen and Yen (1979), construct validity is the degree to which a test measures the theoretical construct it intends to measure. A literature review was conducted to develop the constructs for this study (see Figure 1). Specifically, two scales and corresponding items were developed to measure students' behavior and attitude related to Twitter use for their academics.

Cronbach and Meehl (1955) noted that the nomological network helps to test hypotheses about constructs. As indicated, a *nomological network* is a graphical representation of interrelation among and between constructs and observable variables. With this, a nomological network was developed for the construct and hypotheses was tested (see Figure 1).

Evaluation of reliability indices. Reliability refers to the repeatability or consistency of results; often, it is referred as a reliability coefficient (i.e., Coefficient Alpha). To explain the importance of reliability coefficient, Cronbach (1951) related that "[a] reliability coefficient demonstrates whether the test designer was correct in expecting a certain collection of items to yield interpretable statements about individual differences" (p. 297). According to Cronbach, reliability and validity are equally important for a test. The reliability of this study was therefore tested as well, although both scales were tested separately.

As various authors of previous studies have indicated, social network applications have some impact students' academic performance (Junco, 2012; Kirschner & Karpinski, 2010; Ozer, Karpinski, & Kirschner, 2014; Rosen, Carrier & Cheever, 2013). Nonetheless, these studies failed to provide details about how social networking applications impact students' academic performance. It becomes important to explore psychometric properties of the newly developed TSSS and TSAA items, and one must also understand the relationship between TSSS and TSAA on undergraduate students' academic performance.

Validity of the Study

Different types of validity were tested for this study (see Table 7). The Table shows validity type and corresponding terminology, and types of evidence were analyzed to explore those validity. In one possible example, the construct map and the nomological network were used to validate the test specification and item development, while Cronbach's Alpha was used to validate internal consistency and generalizability.

Table 7

Summary of Validity Evidences, by Form, Provided in Current Study

Form of validity	Terminology based on Wolfe and Smith (2007) and Kosko (2019)	Evidence in the current study
Test content	Content: Instrument Purpose Content: Test Specifications & Item Development Content: Item Technical Quality	• Construct a map/nomological network • Item table (Scale and subscale and items) • Mean square fit statistics (Rasch)
Response process	Sustentative: Behavioral Observation	• Analysis of responses (FA—EFA and CFA) • Person mean square fit statistics • Wright map
Internal structure	Structural: Unidimensionality Interpretability: Person-Item Maps	• Infit & outfit statistics • PCA of standardized residuals
Generalization	Generalizability: Reliability, Internal Consistency	• Cronbach Alpha • Item separation reliability • Person separation reliability

Note. This table is designed based on the work of Kosko (2019), Wolfe and Smith (2007), and terminologies were adopted from Kane (2011).

Rasch Analysis

The Rasch Analysis is a psychometric analysis technique that helps to analyze and improve precision data collection instruments. Under the Rasch Analysis, the Rasch Model (RM) is employed, which is a family of measurement models that converts raw scores into linear and reproduceable measurement (Brentani & Golia, 2007). The RM also helps to analyze data collection instrument quality and respondents' performances. The *Winsteps* is a Rasch Analysis and Rasch Measurement software for persons (i.e., participants) and items. Here, Winsteps version 4.3.4 was used to run the Rasch Analysis. The Rasch Analysis helps to understand whether the Rating Scale Model (RSM) of this study is appropriate (Andrich, 1978). Moreover, it helps to check whether all items related to a measure are measured on the same scale as well (i.e., five-point Likert scale).

Unidimensionality. This concept is an explicit requirement for the Rasch Analysis (Linacre, 2006; Smith, 1996; Wright & Masters, 1982). Seeing this, the assumption of unidimensionality of each subscale was tested for this study. In addition, a separate RSM is required for each subscale (Bond & Fox, 2015; Smith, 1996), such as *School Work*, *Procrastination and Leisure*, *Multitasking*, *Responsibility*, and *Perception*.

The Rasch Principal Components Analysis (PCA) was used to examine evidence of the unidimensionality in each subscale. Once item popularity as well as item and person/participant fit were examined (Linacre, 1998), the Rasch PCA was performed on both items and persons and was used in the final model as well. The Rash Analysis helps to examine how similar subscales are compared to the expected model (Wright, 1996). Therefore, results were examined to see if subscales were mirrored with the expected model during the Rasch analysis. Eigenvalues for all subscales were examined to check whether the assumption of unidimensionality was met.

The Grouped Rating Scale Model (GRSM) was considered for the analysis because the Grouped RSM treats each group of items as independent units and analyzes them. This accommodates the differences between the TSAA and TSSS scales, and it fulfills the assumption of unidimensionality (Bond & Fox, 2015). The Grouped RSM model was run with five groups, three *School Work* items, four *Procrastination and Leisure* items, three *Multitasking* items, three

Responsibility items, and four *Perception* items. In this study, the Rasch Analysis helped to understand the appropriateness of the Rating Scale Model (RSM). Moreover, all items related to the measure; both TSSS and TSAA scales (i.e., five-point Likert scale based attitudinal and behavioral frequency measure) were likewise measured on the same scale.

Wright map and construct key map. These maps display a graphical representation of the item and person difficulty; they show item difficulty levels and persons' ability level side-by-side. The Wright map consists of two vertical histograms. The left side histogram shows a person's or participant's ability, and it is represented by "Ɵ," where the person with less ability lies at the bottom and the person with high ability lies at the top. The right-side histogram shows items difficulty, and it is represented by "δ," where least difficult item(s) lies at the bottom and most difficult item(s) lies at the top. The Wright map shows whether the item difficulty and persons' ability are theoretically comparable.

In this study, the Wright map and the construct map were used to show the item difficulty level and expected behavior of participants. The Wright map provides evidence of construct validity as it can plot actual item difficulty and the corresponding person's or participant's ability (Boone, 2016). As such, the Wright map was utilized to provide the construct validity of this study. Like the Wright map, the Construct Key map (or simply construct map) is a visual representation of items and their observational range (Wilson, 2004). The top of the Construct Key map shows the most expected value and the bottom shows the least expect value, whereas the middle of the Construct Key map shows average value. This was used to measure item difficulty and internal structure.

Logit. When person parameter and item parameter are put on a same scale or map, it is called "logit" (Rasch, 1977; Bond & Fox, 2015). In the logit scale, lowest values lie at the bottom and highest values lies at the top, with such a person and item lie at the bottom, indicating least ability and the top indicating highest ability. The logit scale or map has two types of regions: the white region or expected region of the model; and the shaded region or unexpected region of the model. When items and persons fit the model's expected behavior, they fall in the white region; otherwise, they fall in the shaded region. In general, fit values in logit map range $t = \pm 2.0$.

Item fit. Infit and outfit statistics in Rasch analysis help to understand whether items hold unidimensionality. The infit statistics are nothing but weighted standardized residuals; it helps to identify if a person's ability is closer to the item difficulty level. Also, it helps to identify unexpected responses by persons considering items difficulty estimation (that is, it compares persons' ability estimation with items difficulty estimation). The outfit statistics help to understand the irregularities in infit scores (Bond & Fox, 2015, 2007; Linacre, 2016). Infit and outfit statistics are measured using mean-square fit statistics (MNSQ) and t-standardized fit statistics (ZSTD). The MNSQ are the squared residuals of an item, and they show the magnitude of the misfit or amount of distortion of the measure to which it belongs. The ZSTD are the t-tests of hypothesis, which indicates whether the data fit the model (Bond & Fox, 2015). The acceptable values of ZSTD are ±2 (Adams & Khoo, 1993; Bond & Fox, 2007; Wilson, 2005); the acceptable values of MNSQ ranges between 0.5 to 1.5 as productive for measurement (Wright & Linacre, 1994).

This study checked all infit and outfit values based on MSNQ and ZSTD, and most measured items measure fell within the acceptable range of MSNQ and ZSTD. Still, some items did not fall within the range, which explains the issues in CFA analysis. Perhaps a future study could reveal how to fix those items.

Linear Regression Analysis

The Hierarchal Multiple Regression (HMR) was used to understand the relationships between the Twitter attitude measure and behavior measure and students' academic performance. HMR was also used to examine any relationships between students' academic achievement and demographic information.

Multiple Linear Regression (MR) is an extension of simple linear regression analysis (Keith, 2006). *Simple linear regression* is a statistical method that helps to understand relationships between two quantitative variables (usually an independent variable (IV) and a dependent variable (DV)). MR is commonly used in social science studies (typically Psychology, Sociology, or Education), and it describes the relationships between a set of IVs and a DV. This method allows for multiple IVs to be included in the hypothesized model and predict the desired DV.

Method of entry. A few different ways are available to enter variables into a regression model, such as hierarchical, stepwise, forward, backward, and free. Variable entry in a regression model is known as *method of entry*. In this study, variables were entered using "sequential entry" or "hierarchical entry" doing so allows the researcher to enter the IVs into the model in order or in a hierarchy (see Figure 3). In general, IVs are chosen based on existing research studies or hypothesis developed by the researcher (Field, 2014). Regression based on hierarchical entry is often known as Hierarchical Regression (HR); this type of regression is suitable when multiple IVs exists and the researcher wants to understand the influence of one or more subset of IVs (Lomax & Hahs-Vaughn, 2012). HR helps to explain the variance in the DV by an IV or a group of IVs (i.e., a subset of IVs) while controlling for the effects of previously entered IVs into the regression model (Keith, 2006). The subset of IVs or a group of IVs entered in the regression model is known as blocks.

This study consisted of four blocks of IVs: demography; academic information; Twitter use; and Twitter subscales (i.e., TSSS, TSAA). They were used to understand their influence and relationship on students' academic performance (GPA). The first block represents students' demographic information, the second block represents students' academic information, the third block represents students' frequency of checking Twitter, and the fourth block represents newly developed Twitter subscales (see Figure 3). The details of the blocks appear in see Figure 3.

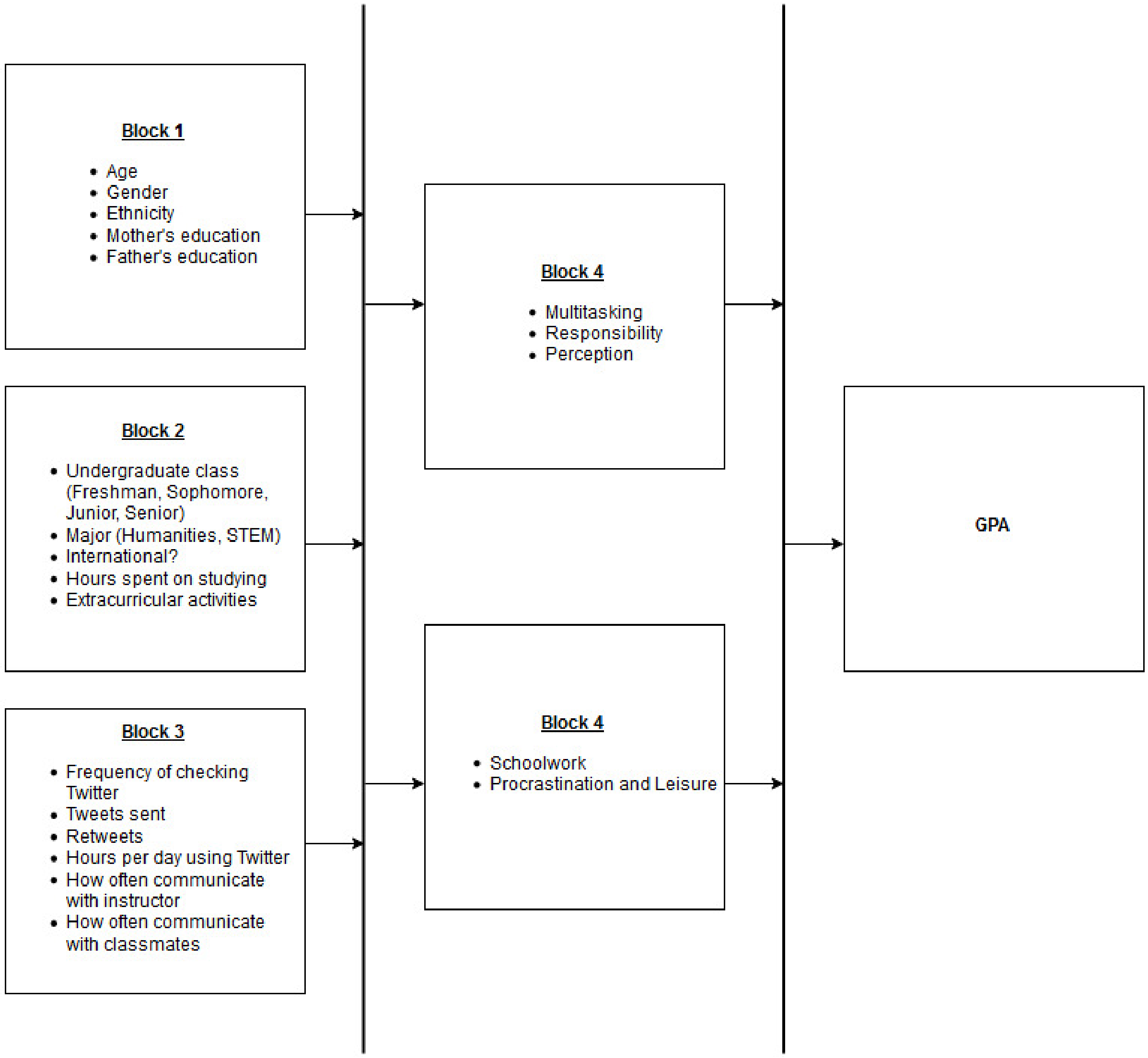

Figure 3. Block-diagram of different blocks for the Hierarchical Regression

The model summary, ANOVA table, and coefficient table were checked to understand the relationship between students' academic performance and Twitter scales (i.e., TSSS and TSAA). In hierarchical multiple regression, the model summary table provides the overview of each model(s) and relationships between dependent variable and independent variables. R^2, the coefficient of determination, indicated the variation of the dependent variable as explained by the independent variables.

The ANOVA Table provides the *F*-statistics and corresponding *p*-value. The *p*-value helps to understand whether the model helps to explain the variation in the dependent variable. Finally, the coefficient table explains the importance of each independent variables- in the model, and it also provides their contribution in explaining the dependent variable. Often, the coefficient table includes Variance Inflation Factor (VIF) values, which help to determine if any multicollinearity issues exist. A VIF value smaller than 10 is expected, which indicates no multicollinearity (Keith, 2006).

Assumptions

The accuracy of the multiple regression result depends on several statistical assumptions (Poole & O'Farrell, 1971). Failure to meet the assumption can be problematic, as Sevier (1957) emphasized the importance of examining the statistical assumptions before conducting any statistical analyses. Violations of assumptions can lead to Type I error or Type II error. A *Type I error* happens when no statistical relationship in the sample population exists, but the researcher mistakenly concludes that there is and rejects the null hypothesis. A *Type II error* occurs when there is in fact a statistically significant relationship in the sample population, but the researcher mistakenly concludes that there is no relationship and retains the null hypothesis. Moreover, violations of assumptions can also impact the effect size (Wilkinson, 1999). Effect size describes the significance of a phenomenon. The correlation coefficient is an example of effect size.

While researchers hold different opinions on how many assumptions should be checked before running a Multiple Regression, Keith (2015) concluded that the most important assumptions are linearity, independence of errors, normality, collinearity, and homogeneity of variance. As such, , missing data, outliers (no outliers found), and the following assumptions were checked before conducting the MR analysis: normality (residuals were normally distributed); linearity (DV and IVs are linearly related); homoscedasticity (variance of errors are similar across the values of IVs); and multicollinearity (IVs are moderately correlated).

Normality of a variable can be checked using Shapiro-Wilk or Kolmogrov-Smirnoff tests. The Shapiro-Wilk test is usually used on a small sample size, while the Kolmogrov-Smirnoff test is for larger sample sizes. In either case, non-significant results indicate normal distribution. Linearity can be tested using scatterplot, where linear data follow the diagonal line (Dimitrov, 2013). Homoscedasticity can also be tested using scatterplot where no points should be outside of ± 3 (on both x-axis and y-axis). The Variance Inflation Factor (VIF) can be used to check multicollinearity; there, a VIF value close to 1 is ideal, while values larger than 10 can be alarming and violation of the assumption (Keith, 2006).

Summary

The methodological approaches used in this study appear in this chapter. This chapter also shows how, step-by-step, different statistical analyses were used in this study. For example, Exploratory Factor Analysis (EFA) was used to explore the underlying factors from the collected data. In other instances, Confirmatory Factor Analysis (CFA) helped to show whether the hypothesized model converged with the research data. How different analysis techniques were used and how the decisions were made from analyzed data also appear in this chapter. Rasch Analysis helped to provide validity aspects of the measure. Classical Test Theory helped to show the reliability of the measure in terms of internal consistency.

Overall, this chapter outlines the measure development, refining the measure, data collection procedure, and data analysis techniques. The findings and implications of this study are discussed in the following chapters.

CHAPTER IV

RESULTS AND DISCUSSION

Introduction

Participant data ($N = 327$) were collected from undergraduate students at a large, public university in the Midwest United States (U.S.) using a survey-hosting website. The data were divided into 60% ($N = 196$) for EFA analysis and 40% ($N = 131$) for CFA analysis. The split-sample is a cross-validation strategy, and it helps to increase the validation of the analyzed results (Frazier, Youngstrom, Kubu, Sinclair & Rezai, 2008; Vassend & Skrondal, 1997). The split took place by choosing participants' records randomly. The sub-sample size was determined by using rule of thumb explained by Gorsuch (1997). To describe the rule of thumb, Gorsuch instructed that EFA requires at least 10-case-per-item, whereas CFA requires less than 10-case-per-item. But in this case, the entire sample was used for Rasch Analysis (RA) and Hierarchical Multiple Regression (HMR).

This study sought to explore the psychometric properties of the newly developed Twitter and Scholastic Synchronicity Scale (TSSS) and Twitter and Scholastic Apportionment Scale (TSAA) items and to understand the relationship between TSSS and TSAA on undergraduate students' academic performance. The study endeavored to answer the following research questions:

(4) What are the psychometric properties of the newly developed TSSS items?

iv. What is the construct validity of TSSS?

v. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSSS?

vi. What are the item properties of the subscales of TSSS (i.e., analyze individual subscales: level of difficulty and rating scale)?

(5) What are the psychometric properties of the newly developed TSAA items?

iv. What is the construct validity of TSAA?

v. What is the internal consistency (e.g., Coefficient Alpha, infit-outfit, and unidimensionality) of TSAA?

vi. What are the item properties of the subscales of TSAA (i.e., analyze individual subscales: level of difficulty and rating scale)?

(6) What is the relationship between Twitter attitude and behavior measures, and students' academic performance?

iii. What is the relationship between TSAA (i.e., a behavioral frequency measure) and academic performance controlling other variables?

iv. What is the relationship between TSSS (i.e., an attitudinal measure) and academic performance controlling other variables?

Factor Analysis

First, Factor Analysis (FA) was used to uncover underlying relationship among items using Exploratory Factor Analysis (Pett, Lackey & Sullivan, 2003; Thompson, 2004). Second, the Confirmatory Factor Analysis (CFA) was used to check how close the conceptual model and hypothesized model were to each other (Dimitrov, 2013; Thompson, 2004).

Explanatory Factor Analysis. Data were analyzed using SPSS version 24.0. Descriptive statistics were included. The factor structure and internal consistency reliability (i.e., Coefficient Alpha [α]) for the overall measure and subscales were examined. For both TSSS and TSAA, the factorability of the items was assessed by using Principal Axis Factoring (PAF) with

Direct Oblimin rotation ($\delta = 0$). In order to investigate the appropriateness of the data for EFA, the Kaiser-Meyer-Olkin (KMO) Measure of Sample Adequacy and Bartlett's Test of Sphericity were examined (Kaiser, 1974). Additionally, eigenvalues and scree plots were examined to determine the number of factors that had to be extracted (Hayton, Allen, & Scarpello, 2004).

EFA and Item Analysis of the TSSS

Using recommended data analytic strategies (Thompson, 2004), EFA (PAF for extraction) with Direct Oblimin rotation was conducted to determine the underlying factor structure that explains variation among the item correlations. For the TSSS, the KMO Measure of Sample Adequacy had a value of .732 and was considered adequate (Kaiser, 1974). Also, Bartlett's Test of Sphericity was significant ($\chi^2 = 920.821$, $df = 45$, $p < .001$), indicating that the correlation matrix is not an identity matrix. A three-factor solution was extracted explaining 73.084% of the variance. The initial EFA was conducted (see Table 8).

Table 8

Pattern Matrix for TSSS

	Factor		
	1	2	3
Twitter is an academic distraction.	.906		
Twitter takes time away from studying.	.871		
Twitter decreases academic performance.	.808		
Twitter is time consuming.	.605		
I multitask with Twitter while studying.		.868	
I have Twitter up while doing homework.		.812	
I am good at multitasking with Twitter.		.657	
I am a responsible person about schoolwork.			.831
My academics are my main focus.			.643
I only use Twitter when I have the time for it.			.448

A ten-item, three-factor model was retained. The first factor had four items and was termed "Perception" (i.e., Range = 4 to 16). The second factor was labeled "Multitasking" with three items (i.e., Range = 4 to 12). The third factor had three items and was labeled

"Responsibility" (i.e., Range = 4 to 12). Two of the three factors had high internal consistency reliability of .888 for "Perception," and .802 for "Multitasking." Despite this, the third factor had slightly low internal consistency reliability of .644 for "Responsibility." Additionally, upon examination of item means, the range fell between 2.07 and 4.23 (see Table 9). Such correlations revealed that all items were moderately correlated and statistically significant ($p < .01$).

Table 9

Descriptive Statistics TSSS

	Mean	*SD*	*N*
I am good at multitasking with Twitter.	2.79	1.203	196
I multitask with Twitter while studying.	2.39	1.229	196
I have Twitter up while doing homework.	2.07	1.226	196
My academics are my main focus	4.21	.874	196
I am a responsible person about schoolwork.	4.23	.825	196
I only use Twitter when I have the time for it.	3.31	1.172	196
Twitter is time consuming.	3.24	1.203	196
Twitter is an academic distraction.	3.20	1.166	196
Twitter decreases academic performance.	2.74	1.047	196
Twitter takes time away from studying.	3.10	1.172	196

After analyzing the pattern matrix for TSSS scale, some items had smaller loadings compared to other items. For example, the "Twitter is time consuming" item has the lowest loadings in the Perception subscale (i.e., the difference with the closest loading is .203), "I am good at multitasking with Twitter" item has the lowest loadings in Multitasking subscale (i.e., the difference with the closest loading is .155), and "I only use Twitter when I have the time for it" item has the lowest loadings in Responsibility subscale (i.e., the difference with the closest loading is .195). King, Leskin, King, and Weathers (1998) cautioned that the higher saturation in factor loading indicates a weaker relationship. This scenario also indicates that there could be a correlation error among items. This situation might not be theoretically pleasing (Cattel, 1963;

King, Leskin, King, & Weathers, 1998). One plausible explanation for this scenario is that each of these three items are more general in nature, whereas other items are more specific in nature.

Considering the lower loading, mentioned items were removed and EFA analysis was repeated a second time. For the updated TSSS, the KMO Measure of Sample Adequacy had a value of .720 and was considered adequate (Kaiser, 1974). Moreover, Bartlett's Test of Sphericity was significant ($\chi^2 = 622.432$, $df = 21$, $p < .001$), indicating that the correlation matrix is not an identity matrix. A three-factor solution was extracted, which explained 83.279% of the variance. A second EFA was conducted (see Table 10). Removing the mentioned items caused an increase in the percentage of the variance.

Table 10

Pattern Matrix II for TSSS

	Factor		
	1	2	3
Twitter decreases academic performance.	.863		
Twitter is an academic distraction.	.862		
Twitter takes time away from studying.	.855		
I have Twitter up while doing homework.		.937	
I multitask with Twitter while studying.		.895	
My academics are my main focus.			.760
I am a responsible person about schoolwork.			.696

A seven-item, three-factor model was retained. The first factor had three items and was termed "Perception" (i.e., Range = 4 to 12). The second factor was labeled "Multitasking" with two items (i.e., Range = 4 to 8). The third factor had two items and was labeled "Responsibility" (i.e., Range = 4 to 8). The evidence suggested that after deleting the problematic items, internal consistency reliability increased slightly for all three factors: .893 for "Perception"; .855 for "Multitasking"; and .673 for "Responsibility." Additionally, upon the examination of item means, the range (i.e., 2.07 to 4.23) remained the same. These correlations revealed that all items were moderately correlated and statistically significant ($p < .01$).

EFA and Item Analysis of the TSAA

For the TSAA, the KMO Measure of Sample Adequacy had a value of .794 and was considered adequate (Kaiser, 1974). Further, Bartlett's Test of Sphericity was significant (χ^2 = 1233.767, $df = 21$, $p < .001$), indicating that the correlation matrix is not an identity matrix. A two-factor solution was extracted, explaining 80.49% of the variance. A seven-item, two-factor model was retained (see Table 11).

Table 11

Pattern Matrix of TSAA

	Factor	
	1	2
I use Twitter to procrastinate if I am struggling/get bored.	.976	
I use Twitter to procrastinate when I should be studying.	.945	
I use Twitter as a break while studying.	.937	
I use Twitter as a free time activity.	.880	
I use Twitter to communicate for group projects.		.802
I use Twitter for schoolwork.		.623
I use Twitter to communicate with my classmates.		.414

In the TSAA scale, the first factor was labeled "Procrastination and Leisure," and it had four items (i.e., Range = 4 to 16). The second factor had three items and was termed "Schoolwork" (i.e., Range = 4 to 12). The two factors had high internal consistency reliability of .964 for "Procrastination and Leisure" and .654 for "Schoolwork." In addition, upon examination of item means, the range was 1.16 to 3.07 (see Table 12). These correlations revealed that all items were moderately correlated and statistically significant ($p < .01$). One of the items in "Schoolwork" appeared problematic. For instance, the "I use Twitter for schoolwork" item has the lowest loadings in Schoolwork subscale (i.e., the difference with the closest loading is .209). As a result, this item was removed before running the analysis again.

Table 12

Descriptive Statistics TSAA

	Mean	*SD*	*N*
I use Twitter for schoolwork.	1.27	.665	196
I use Twitter to communicate with my classmates.	1.64	1.060	196
I use Twitter to communicate for group projects.	1.16	.507	196
I use Twitter as a break while studying.	2.96	1.517	196
I use Twitter as a free time activity.	3.07	1.502	196
I use Twitter to procrastinate when I should be studying.	2.82	1.524	196
I use Twitter to procrastinate if I am struggling/get bored.	2.89	1.490	196

A six-item, three-factor model was retained (see Table 13). The first factor had four items and was termed "Procrastination and Leisure" (i.e., Range = 4 to 16). The second factor was labeled "Schoolwork" and had two items (i.e., Range = 4 to 8). The evidence suggested that after deleting the problematic item, the internal consistency reliability remained the same for first factor and increased slightly for the second factor: .964 for "Perception" and .673 for "Schoolwork." What is more, upon the examination of item means, the range (i.e., 1.16 to 3.07) remained the same. The correlations revealed that all items were moderately correlated and statistically significant ($p < .01$).

Table 13

Pattern Matrix II of TSAA

	Factor	
	1	2
I use Twitter to procrastinate if I am struggling/get bored.	.977	
I use Twitter to procrastinate when I should be studying.	.942	
I use Twitter as a break while studying.	.936	
I use Twitter as a free time activity.	.877	
I use Twitter to communicate for group projects.		.771
I use Twitter to communicate with my classmates.		.674

Confirmatory Factor Analysis. For purposes of this study, CFA was conducted in order to provide confirmatory evidence of the factors explored using EFA (Pett, Lackey & Sullivan, 2003; Thompson, 2004). In this phase, approximately half of the data were used to run the CFA analysis, following a rule of thumb by Gorsuch (1997). The *rule of thumb* refers to the general guidelines followed by most researchers. After running the test, different indices (e.g., Chi-Square, RMSEA, SRMR, GFI, AGFI, etc.) were checked to examine the comparability of the conceptual model and hypothesized model.

CFA and Item Analysis of the TSSS

The CFA analysis was done using the correlation matrix based on the seven items and three factors obtained from the EFA analysis (Marsh & Hocevar, 1985). The initial results indicated that the data supported the hypothesized model. Furthermore, the model identification table (shown in Table 14) indicated that the initial model was over-identified. As the number of distinct values in *S* is higher than the number of free parameters, the model should be considered as over-identified.

Table 14

Model Identification TSSS

Number of Distinct Values in *S*	Number of Free Parameters
$p(p + 1)/2$	7 Factor Loadings
Here, p is the number of independent	7 Measurement error variances
variables.	3 Latent covariance correlation
7(7+1)/2	0 Correlated measurement error
7(8)/2	
28 total	17 total

After adding the suggested changes i.e., allowing correlation between two items in perception subscale, the following CFA model was obtained (see Figure 4). Also, the Chi-

Square value (χ^2 = 2.81, p >.05) was not significant. The non-significant value of Chi-Square (χ^2) indicated that the predicted or hypothesized model is congruent with the observed data (Marsh & Hocevar, 1985). Often, the Chi-Square (χ^2) test is criticized as a size dependent goodness-of-fit index (Raykov, 1998). Seeing this, other goodness-of-fit indices were analyzed, such as RMSEA, RMR, SRMR, GFI, and AGFI. The obtained results indicated that the RMSEA index was .001, and a smaller value of RMSEA index indicates data is good fit for the hypothesized model (Raykov, 1998). RMR index was .04, and SRMR index was .04 as well. According to Anderson and Gerbing (1984), RMR and SRMR values of less than .05 is better. Jöreskog and Sörbom (1989) noted that a higher GFI and AGFI index value is better. Higher GFI and AGFI indexes indicate closeness between the hypothesized model and the model obtained from research data. GFI and AGFI value of 0 means no match, and 1 means that there is a perfect match between the hypothesized model and the actual model (i.e., model based on research data). The obtained results show that GFI index was .95, and AGFI index was .98; both values are an indication of model fit (MacCallum & Hong, 1997). These indices (i.e., RMSEA, RMR, SRMR, GFI, and AGFI) indicated that the obtained model was a good model and that the sample data supports the hypothesized model (see Figure 4).

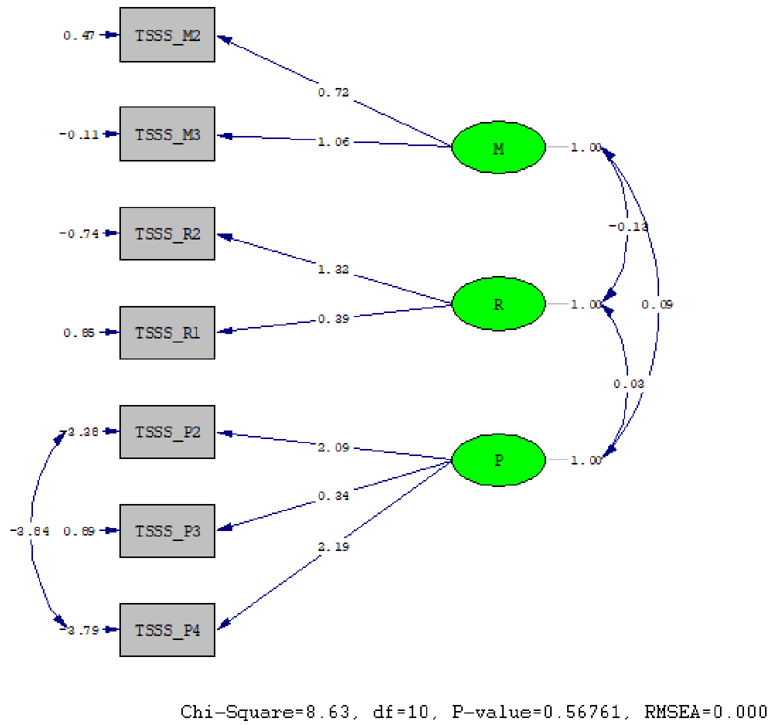

Figure 4. CFA model for TSSS. “M” refers to Multitasking, “R” refers to Responsibility, and “P” refers to Perception

In the CFA model for the TSSS scale, two covariances in the “Perception” subscale were allowed to correlate. This was done in order to improve the results of the CFA model. According to Floyd and Widaman (1995), if two items share factor loadings or share the same latent variable, the covariance should be allowed to correlate. In this scenario, items TSSS_P2 and TSSS_P4 share same latent variable in the “Perception” subscale; therefore, any correlations between these items were allowed.

CFA and Item Analysis of the TSAA

The CFA analysis was done using the correlation matrix based on the six items and two factors obtained from the EFA analysis. Initial results indicated that the hypothesized model and the theoretical model differ. Still, the model identification table indicated that the initial model was over-identified (see Table 15). The number of distinct values in *S* is higher than the number of free parameters; thus, the model should be considered as over-identified.

Table 15

Model Identification TSAA

Number of Distinct Values in *S*	**Number of Free Parameters**
$p(p + 1)/2$	6 Factor Loadings
Here, p is the number of independent	6 Measurement error variances
variables.	2 Latent covariance correlation
6(6+1)/2	0 Correlated measurement error
6(7)/2	
21 total	14 total

Like the TSSS model analysis using CFA, the TSAA model was analyzed and the same goodness-of-fit indices were considered. After adding the suggested changes in the TSAA model, the following CFA model was obtained. Interestingly, the Chi-Square ($\chi^2 = 5.37$, $p > .05$) value was non-significant. With this, the Chi-Square (χ^2) result suggested that the hypothesized model fits the data (Marsh & Hocevar, 1985). The RMSEA index was .001, and smaller value of RMSEA index indicates that the data is a good fit for the hypothesized model (Raykov, 1998). The RMR index was.01 and the SRMR index was.01. According to Anderson and Gerbing (1984), RMR and SRMR values less than .05 is better. Jöreskog and Sörbom (1989) mentioned that a higher GFI and AGFI index value is more desirable. The obtained results show that the GFI index was .99 and that the AGFI index was .97; both values are an indication of model fit (MacCallum & Hong, 1997). This indicates that the obtained model was a good model (see Figure 5).

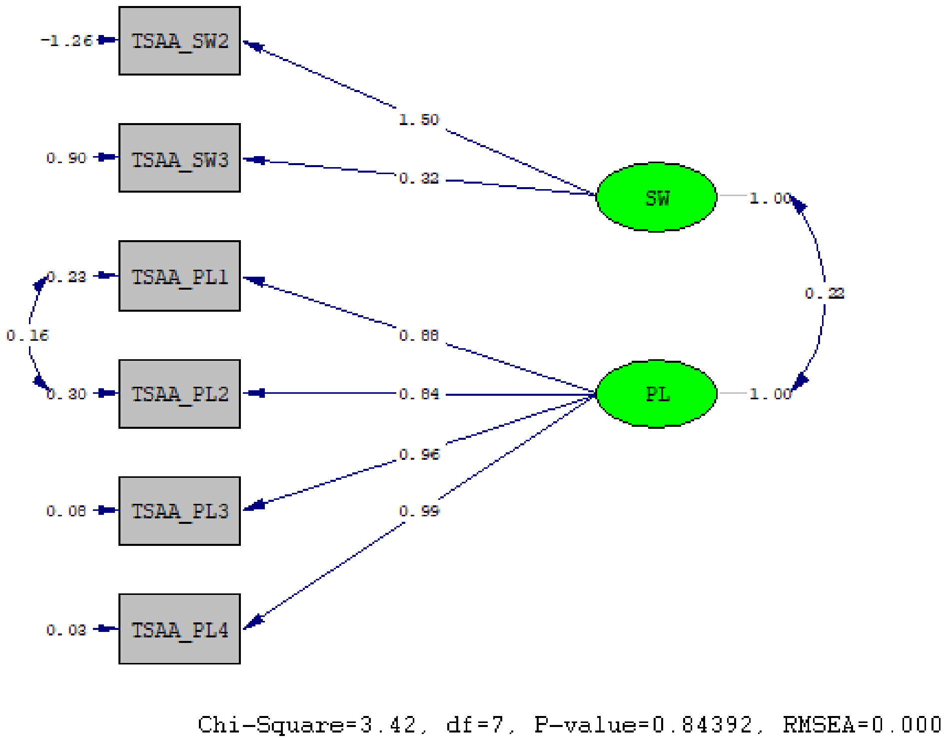

Figure 5. CFA model for TSAA. "SW" refers to Schoolwork, and "PL" refers to Procrastination and Leisure subscale

Note that in the CFA model, two covariances in the "Procrastination and Leisure" subscale were allowed to correlate. This was done in order to improve the results of the CFA model. According to Floyd and Widaman (1995), if two items share factor loadings or share the same latent variable, the covariance should be allowed to correlate. In this scenario, items TSAA_PL1 and TSAA_PL2 share same latent variable in the "Procrastination and Leisure" subscale; as such, the correlation between these items was allowed.

Rasch Analysis

Unidimensionality is considered the most important assumption of a scale (de Jong-Gierveld & Kamphuls, 1985). Presence of unidimensionality indicates that items of the scale measure a single trait or latent variable (Brentani & Golia, 2007; Holmes, 1982). A *latent variable* is not directly observed, but is inferred with the help of other variables, such as happiness, intelligence, or satisfaction, for example. Rasch Analysis (RA) is a good statistical tool by which to check the unidimensionality of a scale (Holmes, 1982). The following sections describe how RA was used to verify unidimensionality of newly developed Twitter scales.

Unidimensionality

The unidimensionality of an assessment is one of the primarily analyses examined in Rasch analysis (Bond & Fox, 2007; Linacre, 2010; Wright & Stone, 1979). *Unidimensionality* represents a single dominant construct, even though multiple items are used to collect participants' data. In sum, unidimensionality ensures the measure of a single topic (Haley, McHorney & Ware Jr, 1994). In this study, construct keymap, wright map, logits, and Principal Component Analysis (PCA) items' standardized residuals fit statistics were used to assess the unidimensionality of the newly developed Twitter measure for both scales: TSSS and TSAA.

Item Fit Statistics. The items' standardized residuals statistics were analyzed for the unidimensionality assessment; the average mean square infit (MNSQ =.98, Z = -.62) and outfit (MNSQ = 1.00, Z = -.39) statistics both appear well within the generally accepted range of 0.5 to 1.5 (Bond & Fox, 2007; Wilson, 2005). Also, the ZSTD statistics were acceptable as average ZSTD scores were between ±2 (Adams & Khoo, 1993; Bond & Fox, 2007; Wilson, 2005)

The PCA residual loading for items indicated acceptable infit and outfit MNSQ values in all five contrasts, which were between .5 to 1.5 (see Table 16). Literature based on fit statistics provides that an MNSQ value range between 0.5 and 1.5 is a good range (Linacre, 2016; Wright & Linacre, 1994). Resultantly, the unidimensionality of both TSSS and TSAA scales held. Furthermore, the average infit and outfit MNSQ values were .98 and 1.00, which support the overall unidimensionality of the whole measure (i.e., both TSSS and TSAA scales). The acceptable range of infit outfit statistics also suggested the construct validity of both TSSS and TSAA scales (Baghaei & Amrahi, 2011).

Table 16

Item Analysis Statistics for TSSS and TSAA Scales

				Infit		Outfit	
		Item	SE	MNSQ	ZSTD	MNSQ	ZSTD
TSSS	Perception	TSSS_P2	.06	1.08	1.07	1.05	.70
		TSSS_P3	.06	1.35	4.25	1.45	5.04
		TSSS_P4	.06	1.06	0.77	1.01	.18
		Mean	.06	1.16	2.03	1.17	1.97
	Multitasking	TSSS_M2	.06	.87	-1.70	.91	-1.09
		TSSS_M3	.06	.77	-3.11	.75	-2.66
		Mean	.06	.82	-2.40	.83	-1.87
	Responsibility	TSSS_R1	.08	1.19	1.84	1.38	3.33
		TSSS_R2	.08	1.17	1.66	1.42	3.63
		Mean	.08	1.18	1.75	1.40	3.48
TSAA	Schoolwork	TSAA_SW2	.09	1.27	2.17	1.01	.11
		TSAA_SW3	.21	1.00	.08	1.06	.29
		Mean	.15	1.13	1.12	1.03	.20
	Procrastination and Leisure	TSAA_PL1	.06	.82	-2.52	.80	-2.76
		TSAA_PL2	.06	.91	-1.29	.88	-1.53
		TSAA_PL3	.06	.64	-5.63	.65	-5.10
		TSAA_PL4	.06	.64	-5.63	.65	-5.16
		Mean	.06	.75	-3.77	.74	-3.63
		Total Mean	.08	.98	-.62	1.00	-.39

In terms of reliability, person/participant reliability was .83 (with 2.24 separation) and item reliability was 1.00 (with 15.33 separation). In addition, the Global Root-Mean-Square Residual value was .85, whereas the expected value was .86. With this, the overall measure fits the hypothesized model.

Item difficulty and rating scales of individual subscales. In order to explore item difficulty, rating scales were analyzed. Item difficulty for individual subscales indicated MNSQ infit and outfit scores were below 2.0. According to Linacre (2005), an acceptable MNSQ value is less than 2; under a stricter rule, it is less than 1.5. Values higher than 2.0 represent higher item difficulty. Moreover, the acceptable item difficulty value is considered as evidence of construct validity. Because of this, the acceptable item difficulty level indicates construct validity of Twitter scales.

To analyze the rating scale of individual subscales, each item MSNQ and mean MNSQ of each subscale were analyzed. The analyzed results indicated that each subscale had an acceptable rating scale, which is less than 2 MSNQ infit and outfit values (Linacre, 2005). Correspondingly, SE range of items were .08 to .14, which indicated that the rating scale was good (Table 17 and Table 18).

Table 17

Level of difficulty of individual subscales of TSAA subscales

Item	label	δ (delta)	SE	Infit	Outfit
TSAA_SW2	0	-.69	.06	1.40	1.20
	1	-.15	.08	.90	.70
	2	.39	.13	1.10	.90
	3	.64	.21	1.60	1.40
TSAA_SW3	0	-.50	.06	1.10	1.10
	1	-.03	.12	2.10	1.00
	2	.40	--	2.90	1.20
TSAA_PL1	0	-1.41	.09	.80	.70
	1	-.93	.07	.40	.40
	2	-.40	.05	.40	.30
	3	.01	.06	.70	.60
	4	.66	.09	.80	.80
TSAA_PL2	0	-1.46	.10	.70	.70
	1	-.89	.09	.80	.80
	2	-.49	.06	.60	.50
	3	.04	.08	.70	.70
	4	.48	.07	.80	.90
TSAA_PL3	0	-1.39	.08	.60	.70
	1	-.81	.07	.50	.50
	2	-.24	.04	.30	.20
	3	.08	.05	.60	.60
	4	.89	.08	.70	.70
TSAA_PL4	0	-1.43	.09	.60	.60
	1	-.89	.06	.30	.30
	2	-.27	.04	.30	.20
	3	.02	.04	.60	.60
	4	.86	.08	.70	.70

Table 18

Level of difficulty of individual subscales of TSSS subscales

Item	label	δ (delta)	SE	Infit	Outfit
TSSS_M2	0	-1.17	.09	1.00	1.00
	1	-.41	.06	.80	.80
	2	-.30	.09	1.50	1.40
	3	.30	.07	.80	.70
	4	1.47	.05	.40	.50
TSSS_M3	0	-1.06	.07	.80	.80
	1	-.35	.07	.70	.60
	2	.04	.09	.90	.90
	3	.40	.07	.80	.70
	4	1.55	.05	.50	.50
TSSS_P2	0	-1.95	.29	.80	.80
	1	-.82	.09	1.40	1.40
	2	-.57	.07	1.20	1.10
	3	.05	.08	.80	.80
	4	-.03	.14	1.50	1.50
TSSS_P3	0	-1.73	.25	.80	.80
	1	-.34	.08	2.20	2.30
	2	-.40	.08	1.80	1.90
	3	-.19	.12	1.60	1.60
	4	-.14	.20	1.80	1.90
TSSS_P4	0	-1.76	.20	.80	.70
	1	-.65	.12	1.80	1.80
	2	-.53	.06	1.10	1.00
	3	.00	.07	.90	.90
	4	.03	.19	1.60	1.60
TSSS_R1	0	-5.17	.00	.00	.00
	1	-1.12	.25	1.50	1.50
	2	-.71	.17	1.70	1.80
	3	-.40	.07	1.60	1.70
	4	-.37	.07	1.30	1.30
TSSS_R2	0	-5.17	.00	.00	.00
	1	.19	.00	6.30	5.00
	2	-.80	.16	1.50	1.60
	3	-.41	.06	1.50	1.50
	4	-.37	.08	1.40	1.40

Principal Component Analysis (PCA) of Residuals. The Principal Component Analysis (PCA) results demonstrates that the Rasch model for Twitter measure explained 65.2% of the variance (24.39 eigenvalue) in the data for the subscale (see Figure 6). The unexplained total variance was 34.8% of the variance (13 eigenvalue). The eigenvalues of all five contrasts were 4.85, 1.83, 1.62, 1.42, and .65. It appeared that the first eigenvalue was above the acceptable range of eigenvalue (see Figure 6). An acceptable eigenvalue for first contrast is less than 3 (Linacre, 2005).

Table of STANDARDIZED RESIDUAL variance in Eigenvalue units = ITEM information units

		Eigenvalue	Observed		Expected
Total raw variance in observations	=	37.3923	100.0%		100.0%
Raw variance explained by measures	=	24.3923	65.2%		63.3%
Raw variance explained by persons	=	7.1868	19.2%		18.7%
Raw Variance explained by items	=	17.2055	46.0%		44.7%
Raw unexplained variance (total)	=	13.0000	34.8%	100.0%	36.7%
Unexplned variance in 1st contrast	=	4.8566	13.0%	37.4%	
Unexplned variance in 2nd contrast	=	1.8357	4.9%	14.1%	
Unexplned variance in 3rd contrast	=	1.6221	4.3%	12.5%	
Unexplned variance in 4th contrast	=	1.4228	3.8%	10.9%	
Unexplned variance in 5th contrast	=	.6579	1.8%	5.1%	

Figure 6. Standard Residuals

The obtained eigenvalue for first contrast proved too problematic to support the unidimensionality of the scale. As unidimensionality was not fully satisfied, further analysis became necessary and a deattenuated correlation was considered. Often, deattenuated correlation can verify unidimensionality (Nam, Yang, Lee, Lee & Seol, 2011; Rogers, 1976). The *deattenuated correlation* helps to explain whether two subsets of items are uncorrelated or uncorrelated because of measurement error. Obtained results suggested that the deattenuated correlation values ranged from .68 to 1.00 (most of the values were close to 1.00 while only 3 of them were close to .68, which indicates moderate correlation). According to Linacre (2005), an acceptable range of deattenuated correlation value falls between .50 and 1.00, which occurred during analysis. Due to this, unidimensionality was not a problem.

Perhaps smaller sample size caused the first contrast value to be higher than 3 eigenvalues. As Chou and Wang (2010) mentioned, at least 500 participants in a study is good for Rasch Analysis; in this study, however, only 327 participants were considered, which was a small sample size for RA. Other studies also indicated that the eigenvalue of first contrast depends on the sample size (Linacre & Tennant, 2009; Raiche, 2005). As a result, considering an eigenvalue of 24.39 compared to 4.85 does not seem like very much (see Figure Residuals). Based on the evidence obtained from deattenuated correlation, the measure was considered as unidimensional.

Wright Map, Construct Key Map, and Logit

The Wright map and the Construct keymap were used to compare the predicted order of item difficulty with the actual order of item difficulty in the data set. These comparisons help to establish the validity of evidence towards the item difficulty hierarchy or response processes as well as unidimensionality or internal structure (Rittle-Johnson, Matthews, Taylor & McEldoon, 2011). When person parameter and item parameter are placed on the same scale or map such as the Wright map, that scenario is called "logit" (Rasch, 1977; Bond & Fox, 2015).

Wright map. The Wright Map indicates that most item and person difficulties range from -2.00 to +2.00; this is a good range of item difficulty (see Figure 7). It can thus be concluded that the items of the newly developed measure are good and that the difficulty level is acceptable.

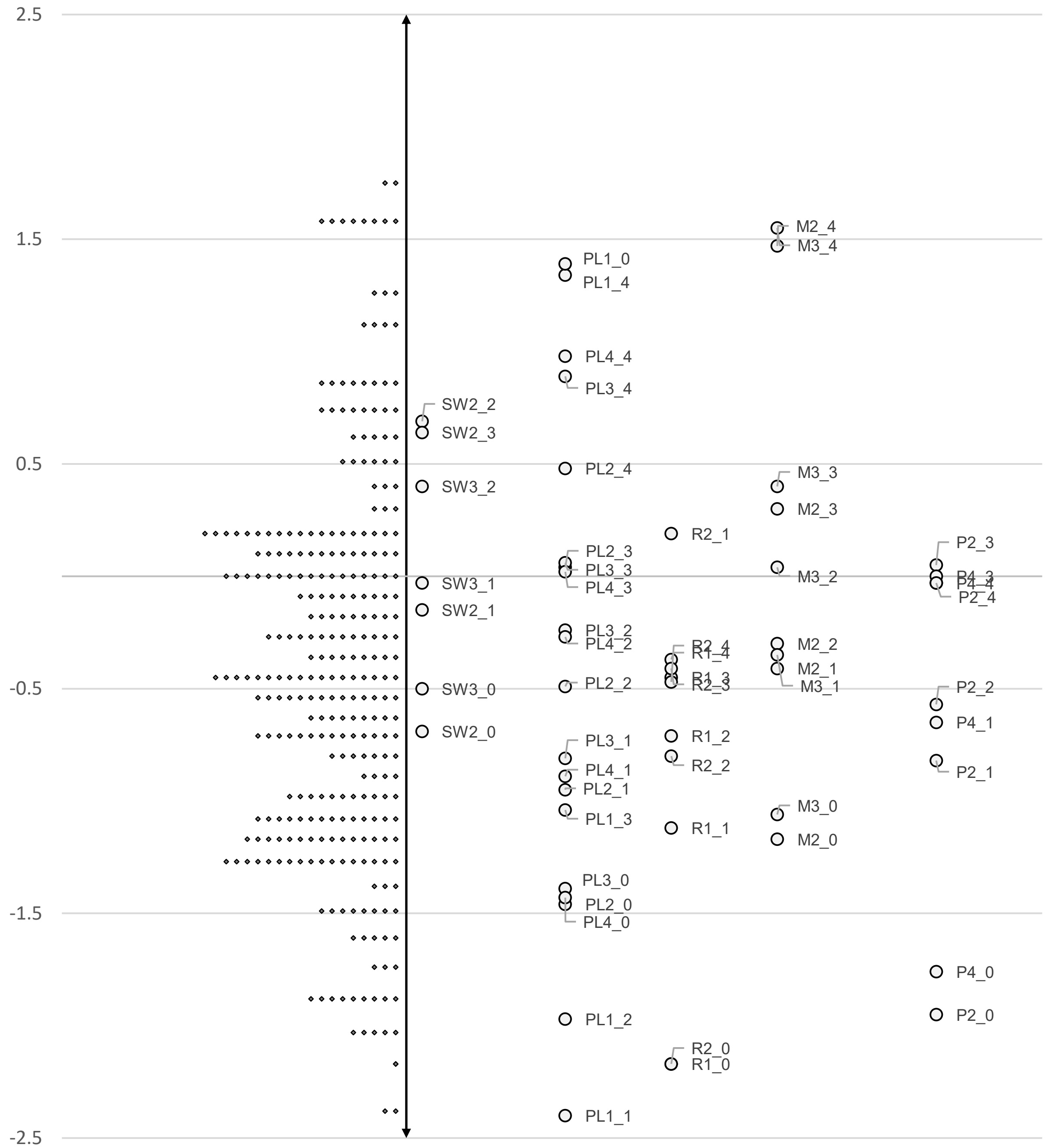

Figure 7. Wright map of the newly developed Twitter scales.

In this measure, it was expected that item difficulty would increase from lower to higher levels. All items in the measure were coded 0 to 4; under this coding, 2 was the middle value of the Likert scale, and 2 was considered during the analysis. Note that here, 0 referred to lower levels and 4 referred to higher levels. The average difficulty of the items should range from -

3.00 to +3.00. Note that, there were two participants outside the range. For simplicity those participants were removed.

Moreover, all subscales were plotted in the Wright map, which indicated that logit values were ±2.5 (see Figure 7). According to Linacre (2005), logits between ±3.0 indicates an acceptable rating scale and level of difficulty. With this, the evidence suggests that no problem exists in terms of level of difficulty and rating scales of individual subscales (see Figure 7). Note that the axis of the Wright Map was adjusted to ±2.5 for simplicity as most data points are within the range of ±2.5. Notably, one data point fell outside of the range of ±2.5 (i.e., -5.17), and that data point was removed from the Wright Map as that amounted to an outlier.

Construct keymap. The construct keymap analysis results suggested that all items in the measure support the expected behavior (see Figure 8). The construct keymap allows a comparison between the predicted order of item difficulty and actual order of item difficulty in a data set. In the present study, comparisons from the construct keymap help to establish the validity evidence toward item difficulty. An acceptable level of item difficulty provides evidence of a scale being valid and unidimensional.

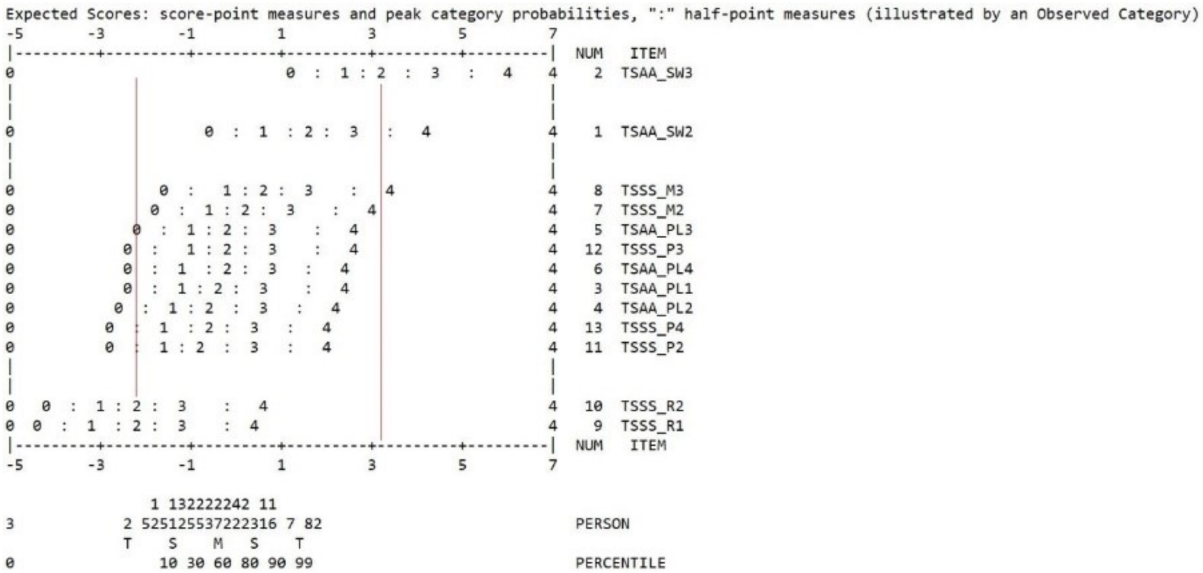

Figure 8. Construct Keymap of the newly developed Twitter scales

Construct keymap results indicated that each item's difficulty level falls between -3.50 to + 3.50. The range is acceptable, and no items fell outside this acceptable range.

Multiple Regression

TSAA and GPA

The multiple regression assumptions, such as, (1) normality, (2) linearity, (3) outliers, (4) homoscedasticity, and (5) multicollinearity were tested before running the analysis. The histogram and Q-Q plot showed that the normality assumption for dependent variable was non-significant, meaning that it was not violated, $p > .05$. According to the scatterplot of raw scores of both variables, the assumption of linearity was non-significant, and neither was it violated, $p > .05$. According to Cook's test, no outliers existed, since the range fell between .001 to .566 and none of the values were above 1; therefore, the assumption of outliers was not violated. Assessing the scatterplot of standardized predicted values versus standardized residuals, no values are outside +3 or -3; this means that the assumption of homoscedasticity was not violated. If the variance inflation factor (VIF) has large values, it signals the presence of multicollinearity

(Keith, 2015); the VIF should not be greater than ten (Lomax & Hahs-Vaughn, 2012). Based on the obtained results, VIF values for Model 1 and Model 2 (i.e., less than 10) indicated that the assumption of multicollinearity was not violated. Despite this, the VIF values for Model 3 (i.e., greater than 10) indicated that the assumption of multicollinearity was violated. Overall, only one assumption (i.e., multicollinearity) was violated whereas the other assumptions were not. Perhaps further study could reveal which item in those mentioned scales is causing the multicollinearity; removing that item could resolve the violation of the multicollinearity assumption.

In Block 1 – Demographics: only "Ethnicity" was inserted. In Block 2 – Academic Information: (1) Major (two groups), (2) Undergraduate Status (Freshman, Sophomore, Junior, Senior), (3) Study Time, and (4) Extracurricular Activities (ECAs) (Yes or No) were inserted. In Block 3 – Twitter Use: (1) Retweets Sent and (2) Twitter Time were inserted. In Block 4 – Twitter TSAA Subscales: (1) TSAA_PL and (2) TSAA_SW were inserted. The dependent variable was GPA. The ANOVA table (see Table 19) indicates that when the Ethnicity was entered by itself, it was not a significant predictor of college GPA, F (1, 325) = 1.379, p = .241. In the second model, Major, Undergraduate Class, Study Time, and Extracurricular Activities were not significant predictors of college GPA, F (5, 321) = 24.293, p = .132. In the third model, Retweets Sent and Twitter Time were not significant predictors of college GPA, F (7, 319) = 45.828, p = .061. But in the fourth model, TSAA_PL and TSAA_SW did serve as significant predictors of college GPA, F (9, 317) = 39.785, p < .001.

Table 19

ANOVA Table TSAA and GPA

Model		Sum of Squares	*df*	Mean Square	*F*	*p*
1	Regression	.308	1	.308	1.379	.241
	Residual	72.610	325	.223		
	Total	72.918	326			
2	Regression	20.018	5	4.004	24.293	.218
	Residual	52.901	321	.165		
	Total	72.918	326			
3	Regression	36.562	7	5.223	45.828	.067
	Residual	36.357	319	.114		
	Total	72.918	326			
4	Regression	38.677	9	4.297	39.785	.000
	Residual	34.241	317	.108		
	Total	72.918	326			

As shown in the first model in Table 20, the demographics contributed non-significantly to the regression model, F (1, 325) = 1.379, p = .241, and accounted for .4% of the variance in college GPA. The academic information in the second model explained an additional 27% of the variance in college GPA, but this change in R^2 was not statistically significant, F (4, 321) = 29.899, p = .218. Nonetheless, the magnitude of R^2 and F-statistics increased substantially. The Twitter use in the third model explained an additional 22.7% of the variance in college GPA, but this change in R^2 was not statistically significant as well, F (2, 319) = 72.579, p = .067. Finally, adding TSAA subscales to the regression model explained an additional 2.9% of the variance in college GPA, with this change in R^2 being statistically significant, F (2, 317) = 9.79, $p < .001$.

Table 20

Hierarchical Multiple Regression Analysis Summary Predicting College GPA (N = 327) using TSSS subscales

	Variable	R^2	ΔR^2	F change	df_1	df_2
Model 1	Demographic	.004	.001	1.379	1	325
Model 2	Demographic Academic Info.	.275	.263	29.899	4	321
Model 3	Demographic Academic Info. Twitter Use	.501	.490	72.579	2	319
Model 4	Demographic Academic Info. Twitter Use TSAA Subscales	.530	.517	9.793	2	317

According to the "Coefficients" table (see Table 21), the unstandardized coefficient in Model 1 for the constant is 2.847 ($p < .001$) and ethnicity is -.020 ($p = .241$). The unstandardized coefficient in Model 2 for the constant is 2.316 ($p < .001$), ethnicity is -.011 ($p = .440$), undergraduate status is -.054 ($p = .004$), major is -.017 ($p = .586$), study time is .216 ($p < .001$), and extracurricular activities is -.044 ($p = .341$). The unstandardized coefficient in Model 3 for the constant is 2.941 ($p < .001$), ethnicity is -.031 ($p = .012$), undergraduate status is -.047 ($p = .002$), major is -.018 ($p = .466$), study time is .135 ($p < .001$), extracurricular activities is -.103 ($p = .009$), retweets sent is .029 ($p = .152$), and Twitter time is -.214 ($p < .001$).

For Model 4, however, the unstandardized coefficient for the constant, ethnicity, undergraduate status, study time, extracurricular activities, Twitter time, TSAA_SW, and TSAA_PL were statistically significant ($p < .001$), except major and retweets sent ($p = .763$). In sum, the unstandardized coefficient for the constant is 2.504, ethnicity is -.025, undergraduate status is -.021, major is -.020, study time is .082, extracurricular activities is -.098, retweets sent is -.005, Twitter time is -.145, TSAA_SW is .097, and TSAA_PL is .032. More specifically, the unstandardized coefficient for most variables in Model 4 was found to be significantly different from zero. Other variables introduced in chapter 3 were removed from the model due to higher VIF value, which indicated multicollinearity.

Table 21

Coefficient Table GPA and TSAA

Model		Unstandardized Coefficients		Standardized Coefficients	*t*	*p*	Collinearity Statistics	
		B	SE	*β*			Tolerance	VIF
1	(Constant)	2.847	.028		10.970	.000		
	Ethnicity	-.020	.017	-.065	-1.174	.241	1.000	1.000
2	(Constant)	2.316	.082		28.242	.000		
	Ethnicity	-.011	.015	-.037	-.772	.440	.978	1.022
	UG Class	-.054	.018	-.140	-2.935	.004	.987	1.013
	Major	-.017	.030	-.026	-.546	.586	.991	1.009
	Study Time	.216	.020	.511	1.619	.000	.975	1.026
	Extracurricular Activities	-.044	.046	-.046	-.954	.341	.973	1.027
3	(Constant)	2.941	.088		33.479	.000		
	Ethnicity	-.031	.013	-.103	-2.513	.012	.924	1.083
	UG Class	-.047	.015	-.122	-3.057	.002	.982	1.018
	Major	-.018	.025	-.029	-.729	.466	.983	1.017
	Study Time	.135	.018	.320	7.369	.000	.831	1.204
	Extracurricular Activities	-.103	.039	-.107	-2.638	.009	.958	1.044
	Retweets Sent	.029	.020	.060	1.436	.152	.894	1.119
	Twitter Time	-.214	.018	-.530	-12.017	.000	.804	1.244
4	(Constant)	2.504	.105		23.888	.000		
	Ethnicity	-.025	.010	-.081	-2.419	.016	.913	1.095
	UG Class	-.021	.013	-.055	-1.692	.092	.952	1.050
	Major	-.020	.020	-.032	-.981	.327	.980	1.021
	Study Time	.082	.016	.194	5.280	.000	.756	1.322
	Extracurricular Activities	-.098	.031	-.102	-3.127	.002	.953	1.050
	Retweets Sent	-.005	.017	-.011	-.302	.763	.824	1.213
	Twitter Time	-.145	.015	-.359	-9.461	.000	.706	1.417
	TSAA_SW	.097	.031	.131	3.142	.002	.856	1.169
	TSAA_PL	.032	.015	.095	2.163	.031	.764	1.309

TSSS and GPA

In Block 1 – Demographics only "Ethnicity" was inserted. In Block 2 – Academic Information: (1) Major (two groups), (2) Undergraduate Status (Freshman, Sophomore, Junior, Senior), (3) Study Time, and (4) Extracurricular Activities (ECAs) (Yes or No) were inserted. In Block 3 – Twitter Use: (1) Retweets Sent and (2) Twitter Time were inserted. In Block 4 – Twitter TSSS Subscales: (1) TSSS_M, (2) TSSS_R, and (3) TSSS_P were inserted. There, the dependent variable was GPA (see Table 23).

The ANOVA table (see Table 22) indicates that when the Ethnicity was entered by itself, it was not significant predictor of college GPA, F (1, 325) = 1.379, p = .241. In the second model, Major, Undergraduate Class, Study Time, and Extracurricular Activities were not significant predictors of college GPA, F (5, 321) = 24.293, p = .132. Nevertheless, the magnitude of R^2 and F-statistics increased substantially in the second model. In the third model, Retweets Sent and Twitter Time were not significant predictors of college GPA, F (7, 319) = 45.828, p = .061. Again, the magnitude of R^2 and F-statistics increased substantially in the third model. In the fourth model, TSSS_M, TSSS_R, and TSSS_P were significant predictors of college GPA, F (10, 316) = 66.744, p = .001. Substantial change in F-statistics and the model not being significant but decreasing in p-value indicate that new variables are impacting on the model, but results are not statistically significant. This situation justifies the need for at least one new variable, and possibly more.

Table 22

ANOVA Table TSSS and GPA

Model		Sum of Squares	*df*	Mean Square	*F*	*p*
1	Regression	.308	1	.308	1.379	.241
	Residual	72.610	325	.223		
	Total	72.918	326			
2	Regression	20.018	5	4.004	24.293	.218
	Residual	52.901	321	.165		
	Total	72.918	326			
3	Regression	36.562	7	5.223	45.828	.067
	Residual	36.357	319	.114		
	Total	72.918	326			
4	Regression	49.488	10	4.949	66.744	.000
	Residual	23.430	316	.074		
	Total	72.918	326			

As the first model in Table 23 displays, the demographic contributed non-significantly to the regression model, F (1, 325) = 1.379, p = .241, and accounted for .4% of the variance in college GPA. The academic information in the second model explained an additional 27% of the variance in college GPA, although this change in R^2 was not statistically significant, F (4, 321) = 24.293, p = .132. The Twitter use in the third model explained an additional 22.7% of the variance in college GPA and this change in R^2 was not statistically significant, F (2, 319) = 45.828, p = .061. Finally, adding TSSS subscales to the regression model explained an additional 17.7% of the variance in college GPA, with this change in R^2 being statistically significant, F (3, 316) = 58.114, p = .001.

As the results indicated that a substantial change in F-statistics for subsequent models are not being significant, and decrease in p-value indicated that the new independent variables are important for the final model. This situation justifies the need for a new variable or variables (Petrocelli, 2003).

Table 23

Hierarchical Multiple Regression Analysis Summary Predicting College GPA (N = 327) using TSSS subscales

	Variable	R^2	ΔR^2	*F* change	df_1	df_2
Model 1	Demographic	.004	.001	1.379	1	325
Model 2	Demographic Academic Info.	.275	.263	29.899	4	321
Model 3	Demographic Academic Info. Twitter Use	.501	.490	72.579	2	319
Model 4	Demographic Academic Info. Twitter Use TSSS Subscales	.679	.669	58.114	3	316

According to the "Coefficients" table (see Table 24), the unstandardized coefficient in Model 1 for the constant is 2.847 ($p < .001$) and ethnicity is -.020 ($p = .241$). The unstandardized coefficient in Model 2 for the constant is 2.316 ($p < .001$), ethnicity is -.011 ($p = .440$), undergraduate status is -.054 ($p = .004$), major is -.017 ($p = .586$), study time is .216 ($p < .001$), and extracurricular activities is -.044 ($p = .341$). The unstandardized coefficient in Model 3 for the constant is 2.941 ($p < .001$), ethnicity is -.031 ($p = .012$), undergraduate status is -.047 ($p = .002$), major is -.018 ($p = .466$), study time is .135 ($p < .001$), extracurricular activities is -.103 ($p = .009$), retweets sent is .029 ($p = .152$), and Twitter time is -.214 ($p < .05$).

Table 24

Coefficient Table GPA and TSSS

Model		Unstandardized Coefficients		Standardized Coefficients	t	p	Collinearity Statistics	
		B	SE	β			Tolerance	VIF
1	(Constant)	2.847	.028		10.970	.000		
	Ethnicity	-.020	.017	-.065	-1.174	.241	1.000	1.000
2	(Constant)	2.316	.082		28.242	.000		
	Ethnicity	-.011	.015	-.037	-.772	.440	.978	1.022
	UG Class	-.054	.018	-.140	-2.935	.004	.987	1.013
	Major	-.017	.030	-.026	-.546	.586	.991	1.009
	Study Time	.216	.020	.511	1.619	.000	.975	1.026
	Extracurricular Activities	-.044	.046	-.046	-.954	.341	.973	1.027
3	(Constant)	2.941	.088		33.479	.000		
	Ethnicity	-.031	.013	-.103	-2.513	.012	.924	1.083
	UG Class	-.047	.015	-.122	-3.057	.002	.982	1.018
	Major	-.018	.025	-.029	-.729	.466	.983	1.017
	Study Time	.135	.018	.320	7.369	.000	.831	1.204
	Extracurricular Activities	-.103	.039	-.107	-2.638	.009	.958	1.044
	Retweets Sent	.029	.020	.060	1.436	.152	.894	1.119
	Twitter Time	-.214	.018	-.530	-12.017	.000	.804	1.244
4	(Constant)	2.504	.105		23.888	.000		
	Ethnicity	-.025	.010	-.081	-2.419	.016	.913	1.095
	UG Class	-.021	.013	-.055	-1.692	.092	.952	1.050
	Major	-.020	.020	-.032	-.981	.327	.980	1.021
	Study Time	.082	.016	.194	5.280	.000	.756	1.322
	Extracurricular Activities	-.098	.031	-.102	-3.127	.002	.953	1.050
	Retweets Sent	-.005	.017	-.011	-.302	.763	.824	1.213
	Twitter Time	-.145	.015	-.359	-9.461	.000	.706	1.417
	TSSS_M	.198	.015	.480	13.007	.000	.747	1.339
	TSSS_R	.090	.022	.135	4.084	.000	.935	1.070
	TSSS_P	-.035	.016	-.074	-2.191	.029	.891	1.122

As Model 4 indicates, the unstandardized coefficient for the constant, ethnicity, study time, extracurricular activities, Twitter time, TSSS_M, TSSS_R, and TSSS_P were statistically significant (p < .001), except undergraduate status, major and retweets sent (p = .763). They were as follows: the unstandardized coefficient for the constant is 2.504, ethnicity is -.025, undergraduate status is -.021, major is -.020, study time is .082, extracurricular activities is -.098, retweets sent is -.005, Twitter time is -.145, TSSS_M is .198, TSSS_R is .090, and TSSS_P is -.035. In other words, the unstandardized coefficient for most variables in Model 4 was found to be significantly different from zero. Other variables mentioned in chapter 3 were removed from the model due to higher VIF value, which indicated multicollinearity.

Summary

Detailed analysis processes are described in this chapter. This chapter also presents how step-by-step different statistical analysis were used in this study. First, Exploratory Factor Analysis (EFA) helped to uncover two factors in the TSAA scale and three factors in the TSSS scale. Second, Confirmatory Factor Analysis (CFA) helped to confirm those factors found in EFA. Classical Test Theory (CTT) provided the evidence of internal consistency reliability of individual subscales of TSAA and TSSS. The internal consistency reliability of two factors of TSAA scale were .964 for “Procrastination and Leisure” and .654 for “Schoolwork.” What is more, the internal consistency reliability of three factors of TSSS scale were .893 for “Perception,” .855 for “Multitasking,” and .673 for “Responsibility.”

Third, Rasch Analysis (RA) provided evidence of internal consistency, construct validity, level of difficulty, and rating scale. Both TSAA and TSSS scales were unidimensional, indicating the construct validity of the scales (Baghaei & Amrahi, 2011). The Wright Map, the construct map, and MNSQ values (i.e., infit and outfit) provided evidence of level of difficulty and rating scales of individual subscales. Fourth, Hierarchical Multiple Regression (HMR) helped to uncover the relationship between students’ academic performance and TSAA scale (i.e., 53% variance can be explained by TSAA). Additionally, the relationship between students’ academic performance and TSSS scale (i.e., 68% variance can be explained by TSSS) was discovered. The following chapter discusses the implications of this research.

CHAPTER V

DISCUSSION, IMPLICATIONS, AND CONCLUSION

Introduction

This final chapter provides the summary of this study along with findings and implications, both of which build on the results provided in Chapter IV. Following the summary, the next section discusses the findings obtained from various statistical analyses in relation to each research question. Finally, this chapter concludes with suggestions for future research as well as this study's limitations.

Summary of the Study

In this study, a measure was developed consisting of two scales: (1) the Twitter and Scholastic Synchronicity Scale (TSSS) and (2) the Twitter and Scholastic Apportionment Assessment (TSAA). The purpose of the measure is to understand how undergraduate students use Twitter for their academics, in terms of attitude and behavior. Data collection for this study occurred using a survey hosting website (i.e., Qualtrics), and 327 undergraduate students from a large, public university in the Midwest United States (U.S.) participated.

The purpose of this study was to explore psychometric properties of the two newly developed scales (i.e., TSSS and TSAA) and understand the relationship between those scales on undergraduate students' academic performance. *Psychometric properties* refers to the validity and reliability (i.e., factor structure, internal consistency, prediction of outcome, and test-retest reliability) of a measure (Devilly & Borkovec, 2000; Schmitt, Langan, Williams & Network, 2007).

To explore and confirm the underlying factors or summary variables of the scales (e.g., blood pressure and body mass index can be used to summarize a person's health condition), Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA) were used. Classical Test Theory (CTT) was to test internal consistency reliability and Rasch Analysis (RA) was used to test the unidimensionality of the scales. In other words, CTT and RA were used to check the reliability and validity of the scales. In addition, RA was used to check construct validity, item difficulty, and rating scale of each subscale. RA focused on infit and outfit statistics, while the Wright map and the construct keymap explored the validity aspects of the newly developed scales. Finally, Hierarchical Multiple Regression was used to understand the relationship between undergraduate students' academic performance and the Twitter scales (i.e., TSSS and TSAA).

Findings of the Study

The findings in this study are discussed in terms of the research questions. The details of the analysis appear in Chapter IV, and this section presents an interpretation of those results. The study endeavored to answer the following research questions:

(1) What are the psychometric properties of the newly developed TSSS items?

(2) What are the psychometric properties of the newly developed TSAA items?

(3) What is the relationship between Twitter attitude and behavior measures and students' academic performance?

Research Question 1: Psychometric Properties of TSSS Scale

Both EFA and CFA indicated that three factors appear in the attitude scale (i.e., TSSS). EFA suggested "Perception," "Multitasking," and "Responsibility" factors with factor loading value ranged from .696 to .937. Moreover, the results obtained from the CFA analysis suggested that the sample data supports the hypothesized model (Vassend & Skrondal, 1997). Different indices (e.g., Chi-Square, RMSEA, SRMR, GFI, AGFI, etc.) were checked to examine the comparability of both the conceptual model and hypothesized model. With this, the TSSS scale proved valid and reliable based on the EFA and CFA results (Reuterberg & Gustafsson, 1992; Thompson, 2004). This means that the TSSS scale can be used to measure students' attitudes towards Twitter use for their academics.

Infit and outfit statistics from RA revealed that construct validity was satisfied (Brentani & Golia, 2007). In sum, this means that the scales are theoretically sound rather than showing any pertinent bias towards participants. More specifically, acceptable infit and outfit statistics indicates that the items contained in a measure are equally difficult or easier for all participants (Engelhard, 1994). Each of the infit and outfit statistics for each of the examined items ranged from .50 to 1.50. Further, the mean infit and outfit statistics of each subscale ranged from .82 to 1.42. In particular, the Wright map logits between ±3 confirmed the construct validity of the TSSS scale. When construct map logits of items fall within the acceptable range, those items belong to same subject or concept (Hiller, Refshauge, Bundy, Herbert & Kilbreath, 2006; Lai, Fisher, Magalhães & Bundy, 1996). In this instance, the subject was Twitter use for undergraduate students' academics. Infit statistics, outfit statistics, and the construct keymap provided the evidence of item difficulty and rating scale. So, the evidence suggests that the newly developed scales measure what they were intended to measure (i.e., students' attitude towards the use of Twitter for academics).

The internal consistency reliability for the three factors were .893 for "Perception," .855 for "Multitasking," and .673 for "Responsibility." This means that replication of this study should show similar results. The average infit MNSQ values ranged from .82 to 1.18, and the average outfit MNSQ values ranged from .83 to 1.40. Both average infit and outfit MSNQ value ranges support the overall unidimensionality of the TSSS scale (Boone, 2016). Again, acceptable ranges of infit and outfit statistics suggested that item difficulty was acceptable, and construct keymap logit with ±3 provided the evidence of acceptable rating scale.

Evidence of unidimensionality suggests that the items contained in the scale are used to measure a single subject or concept. Similarly, an acceptable range of infit and outfit statistics indicates that items are valid and unbiased for participants. The construct keymap logits present additional validity evidence of each subscale within a particular scale. Also, the logits of the construct keymap suggests that the internal structure of an instrument is valid.

Research Question 2: Psychometric Properties of the TSAA Scale

Both EFA and CFA indicated that the TSAA behavior scale consists of two factors. The EFA suggested "Procrastination and Leisure" and "Schoolwork" factors with factor loading values ranging from .674 to .977. Such results suggest that these subscales are moderate- to strongly-related to the measure (Kline, 2014). In addition, results obtained from the CFA analysis suggested that sample data helped to hold the hypothesized model (Reuterberg & Gustafsson, 1992; Thompson, 2004). Different indices (e.g., Chi-Square, RMSEA, SRMR, GFI, AGFI, etc.) were checked to examine the comparability of both the conceptual model and hypothesized model. Based on this evidence from both EFA and CFA, it can be deduced that that the TSAA scale can be used to measure students' behavior when using Twitter for their academics.

Infit and outfit statistics from RA suggested that construct validity was satisfied, as all infit and outfit statistics for all items fell between .50 and 1.50. Additionally, the mean infit and outfit statistics of each subscale ranged from .74 to 1.13. Likewise, the Wright map logits between ±3 confirmed the construct validity of the TSAA scale. Infit statistics, outfit statistics, and the construct keymap provided evidence of both item difficulty and rating scale. Also, as indicated, these results suggest that items are not biased towards any particular group of participants and that the TSAA scale is valid.

The internal consistency reliability for three factors were .964 for “Procrastination and Leisure” and .654 for “Schoolwork.” The average infit MNSQ values ranged from .75 to 1.13, while average outfit MNSQ values ranged from .74 to 1.03. Both average infit and outfit MSNQ value ranges support the overall unidimensionality of the TSAA scale (Boone, 2016). Once again, acceptable range of infit and outfit statistics suggested that item difficulty was acceptable and that the construct keymap logit with ±3 provided evidence of an acceptable rating scale. These results showed that the TSAA scale would behave the same way if used again in the future and will help to measure students’ behavior towards Twitter use for academics.

Once again, the evidence of unidimensionality suggests that the items of the scale were used successfully to measure a single subject or concept; here, it applies to students’ behavior when using Twitter for academics. With this, the acceptable range of infit and outfit statistics indicates that items are valid and not biased towards any group (e.g., male and female) of participants. The construct keymap logits presents additional validity evidence of each subscales (i.e., “Procrastination and Leisure” and “Schoolwork”) within the TSAA scale. Likewise, the logits as they appear within the acceptable range on the construct keymap suggest that the internal structure of instrument is valid (see Figure 8).

Research Question 3: Academic Performance and Twitter Scales

Hierarchical Multiple Regression (HMR) suggested that a positive relationship exists between the Twitter attitude measure and the behavior measure and students’ academic performance. The TSAA scale explained the 53% variance in student’s academic performance, and this relationship was positive. The TSSS scale was likewise able to explain the 68%

variance in students' academic performance, and again, the relationship was positive. Both results suggest that the TSAA and TSSS scales can play important roles in predicting students' academic performance, as reflected in students' GPA (Grade Point Average). These two scales can help to explain most of the variability that occurs in students' GPAs.

Implications and Future Research Directions

This study's results suggest that a positive relationship exists between undergraduate students' academic performance and the responsibility subscale. Quite similarly, a comparable relationship appeared for academic performance and multitasking. Evidence suggested, however, the existence of a negative relationship between academic performance and students' perception about Twitter use for their academics. Yet again, there is a positive relationship between students' academic performance and behavior subscales (i.e., schoolwork, and procrastination and leisure). The implications of this study for undergraduate student users, collegiate faculty and administrators, and future research appear in the sections below.

Implications for undergraduate student users

Evidence from this study suggest that students who believed that academics should be a priority and that they should not use Twitter during study time often had higher GPAs than students who indicated that they care less about their academic performances. This result remains consistent with the existing literature (Michikyan, Subramanyam & Dennis, 2015; Ozer, Karpinski & Kirschner, 2014). In these studies, the authors found a positive relationship between students' academic performance and SNS use. The evidence suggests that students who prioritize tasks and reserve their time to study leads to better academic performance. Using Twitter during leisure time or study breaks positively affects their GPAs. In all, limited or controlled use of Twitter does not hamper students' academic performance.

The results of this study suggested that students who believed that Twitter is an academic distraction have lower GPAs than those who believed that Twitter serves as a helpful tool for their academics. Because those who believed that Twitter is helpful for their academics used Twitter to help them complete their schoolwork, the results indicated that there is a positive relationship between those students' Twitter use and their GPAs. Samaha and Hawi (2016) mentioned that some students view social networking sites as an academic distraction. The present study's results support the findings of Samaha and Hawi (2016). Moreover, the findings of this study suggest that it is not what students believe that matters the most, but rather how they use Twitter. There is evidence that when students use Twitter as a communication tool to communicate with peers and instructors, it improves their academic performance.

The evidence also proposes that students should avoid task switching while studying. The results of this study as well as previous studies (Jacobsen & Forste, 2011; Junco, 2012; Karpinski, Kirschner, Ozer, Mellott & Ochwo, 2013; Ophir, Nass, & Wagner, 2009) indicate that focusing on one task at a time fosters productivity in students. In this study, the items appearing in the multitasking subscale were reverse coded; the results confirmed that students who do not multitask while studying performed better academically. This implies that multitasking might often lead to poor quality work as well as incomplete tasks.

A recent study reported that the more college students spend time on social networking sites, the less progress they make in academics (Gonzalez, Gasco & Llopis, 2019). Nonetheless, the findings in the present study indicate that spending time on social networking sites is not detrimental as long as students use it to improve their knowledge or use it for schoolwork. In other words, where and how students spend time on social media is what matters the most in regard to academic achievement.

As noted above, results from this study provide cautionary information about Twitter use for students during their study break or leisure. As such, students can use Twitter during study breaks or leisure time, but not to the extent that they procrastinate from completing their schoolwork. Similarly, this study's results indicate that Twitter can be used as a good form of technology by which to communicate with peers and instructors. In addition, students might benefit from focusing on specific discussion blogs to have deliberate conversation with peers and instructors or even other scholars. This implication supports the findings of some existing literature (Ellison, Steinfield & Lampe, 2007; Pempek, Yermolayeva, & Calvert, 2009) in which the authors asserted that students should use social networking sites as means by which to communicate with classmates.

In a separate study, Ahn (2011) mentioned that students could retain more information if they avoid using social media. Yet the findings of this study suggest otherwise. How students use Twitter might affect their academic performance, and Twitter use can be characterized as either a positive or a negative factor in students' academic success. The findings of this study suggest that if students use Twitter to communicate with their peers and instructors for schoolwork, then Twitter may have a positive impact on their academic performance. When improving students' academic performance, evidence from this study suggests that Twitter can be helpful for them to acquire and retain knowledge.

Implications for Collegiate Faculty and Administrators

The findings of this study have implications for both faculty and administrators. The results here assert that Twitter could be helpful for students' academics, particularly when they use it to communicate with other students and instructors. Perhaps where feasible, faculty and administrators should encourage students to make better use of Twitter and use it more often, such as for deliberate conversation and as an information-sharing medium. Literature that has investigated the use of Twitter suggested that doing so allows users to follow specific users, discussions, and Special Interest Groups (SIGs) in real-time. A *SIG* is a community within a large organization that shares some interest on a specific topic or knowledge area. For example, a SIG can be statistics club where its members like to find applications of statistical concepts. In such cases, faculty and administrators could post important information to student-based discussion blogs or SIGs. Perhaps this could be the fastest and most verifiable way for them to communicate with students.

The findings of this study also suggest that Twitter can be a good tool to develop networks and communicate with other professionals. A few studies mentioned that Twitter can be used to advance professional development (Carpenter & Krutka, 2015; Davis, 2015; Visser, Evering & Barrett, 2014). In those instances, college faculty and administrators could use Twitter to share survey links and collect participants' quick responses since most Twitter users check their accounts for updates at least a few times each day (Vis, 2013). Often, Twitter is also used to share breaking news. When that occurs, Twitter can serve as a quick information sharing and reporting platform. For example, employees could use it quickly to report or update their employers or administrators about their progresses or concerns related to their work. Of course, a business's or college's policies would have to be updated to reflect Twitter's use for official business.

Implications for Research

Prior research has attempted to investigate the relationship between students' use of SNSs and their academic performance (see, e.g., Jacobsen & Forste, 2011; Junco, 2012; Karpinski, Kirschner, Ozer, Mellott & Ochwo, 2013; Ophir, Nass, & Wagner, 2009). The findings of this study suggest that the relationship between Twitter as a SNS and academic performance is not straightforward. A few studies reported that a negative correlation exists between the amount of time students spend on SNS and their academic performance (Hew, 2011; Jacobsen & Forste, 2011; Junco, 2012). Evidence from this study indicate, however, that it is not the amount of time students spend in SNS, but rather how they spend it is related to their academic performance. The findings here reveal that researchers who focus on technology use and students' academics should explore what different ways, if any, in which students are learning as well as if and how technological tools aid the learning process. In sum, the relationship between use of technology and students' academic performance does not have a straightforward relationship. Identifying the ways that students use other technology to learn will foster researchers' understanding of the relation between SNS and academic performance.

Earlier studies that focused on students' academic performance claimed that students' academic performance relates to their attitude and behavior (Ajzen, 1987; Chun Chu & Choi, 2005; Credé & Kuncel, 2008). This study likewise suggests that students' attitude and behavior towards Twitter for academics indeed relates to academic performance. This finding indicates that attitude and behavior remain the most predictive factors of students' academic performance and these traits help to explain the variance in their self-reported GPAs.

Chapter IV presented factor structures of the newly developed Twitter scale. During the analysis, the present study revealed that some items had low factor loadings compared to other factor loadings. The reason for lower factor loadings is that an item was generic in nature, rather than a specific, topic-related item. One example would be, "I am good at multitasking with Twitter." If the words "with Twitter" are omitted, the item is quite general in nature. Hence, this item was not worded properly, and the Exploratory Factor Analysis (EFA) results indicated lower factor loading for this item. This situation implies that a researcher should be careful when developing an item. Problematic items can mislead participants and result in less-meaningful research findings.

Factor Analysis (FA) results indicated three factors in the TSSS or in the behavior scale and two factors in the TSAA of in the attitude scale. Then again, the RA of the overall Twitter scale (i.e., both TSSS and TSAA scale) indicated unidimensionality. At first glance, it can be assumed that the FA and RA results contradict each other. Upon further examination, however, these results actually advance the same idea. In particular, FA intended to uncover the underlying factor structure, and the FA results indicated subscales within each scale. RA intended to show that each of the items endeavored to measure a single topic, which was "Twitter use among undergraduate students for the academics." Further, the RA of individual subscales for the rating scale and difficulty level provided evidence of individual subscales and the validity of those scales. Both FA and RA assisted in evaluating the Twitter scales from two different angles (i.e., the exploratory and confirmatory angles).

Perhaps researchers could evaluate their scales from different points of view and use different statistical tools to evaluate the scales. Different statistical tools could each have specific limitations of their own. For example, some statistical tools work best with large sample sizes (e.g., Chi-Square), whereas some are not sample size dependent (e.g., RMSEA). Even if the initial results suggest that a scale is good, using different statistical tools could aid in strengthening the findings, and they could likewise cross-validate the initial findings. Correspondingly, a study's initial results can suggest that the resulting measure is valid and reliable. Still, a researcher could continue the study with further and more in-depth analyses to better understand the validity and reliability aspects contained within the measure. Besides, each statistical analysis has some advantages and disadvantages; analyzing research data using different statistical tools could therefore help to obtain more robust and accurate findings.

The findings of this study suggest that the newly developed Twitter scales can be adopted easily as a technology tool to find any relationship between students' academic performances and their attitudes and behavior when they use the tool. The items of the Twitter scales can be modified as needed to find and assess any relationships between other technology-oriented tools and students' academic performance. Seeing that Twitter is a technology-oriented tool, its acceptance can be assessed using the Technology Acceptance Model (TAM), focusing primarily on Perceived Usefulness (PU) and Perceived Ease of Use (PEOU).

Although this study focused on neither PU nor PEOU, an increasing number of students come to use Twitter as the number of Twitter users increases every day (Basak, Sural, Ganguly & Ghosh, 2019; Evans, Hackney, Rauniar, Rawski, Yang & Johnson, 2014; Gao, Cao, Li, Yao, Chen & Tang, 2019). Resultantly, the growth in the number of users can be interpreted as both PU and PEOU. In addition, studies have suggested that Twitter is topic-oriented (i.e., micro-blogging), uses a simple User Interface (UI), and provides simpler navigability (see, e.g., Kwon, Park & Kim, 2014). Seeing this, researchers and professionals could use Twitter as a communication tool and then use the Twitter scales to assess attitudes and behavior and the use of this technological medium.

Finally, the findings here suggest that while investigating students' academic performance, researchers should consider both attitude and behavior together. Not considering these two aspects simultaneously could lead instead to incomplete or inaccurate findings. Because students' self-reported behavior and attitude are correlated and are better predictors of their academic performance, these two traits should not be considered separately (Ostroff, 1992;

Shrigley, 1990; Slavin, 1978; Tanaka, Murakami, Okuno & Yamauchi, 2002). The findings here support the findings of studies that focused on students' academic performance and attitude and behavior together (Ajzen, 1987; Chun Chu & Choi, 2005; Credé & Kuncel, 2008).

Limitations

There are several limitations of this study. The sample was collected from only one academic institution. Having a sample from multiple institutions might be viewed as more appropriate to the generalizability of the findings that this study contains. A similar study can include institutions from a larger geographic region or even from all the States in the U.S. Although this study was conducted at only one college, it still consists of an adequate sample size; however, in the future, researchers should find a way in future studies to make the sample more representative of the actual demographic composition in the U.S. (i.e., race/ethnicity, gender, college major, etc.).

A more representativeness and diverse sample could help to reduce participation bias (Deming, 1990). Some participants could provide different answers than their fellow peers, but they did not participate, which results in *participation bias*. For example, some students might have been busy in their studies and as a result had different experiences than those who participated. In such instances, those who did not fully participate by devoting enough time to the study could lead to different findings than the findings produced in this study. In an attempt to minimize or even eliminate participation bias, keeping survey links open for longer periods of time than usual could be key to attaining such a result.

Self-reported surveys is another limitation of this study, because it could suffer from self-report bias. Participants' various emotional states can cause *self-report* bias, which in turn could lead to contaminated results (Donaldson & Grant-Vallone, 2002). In this study, when participants took the survey, how they felt at that moment in general could have positively or negatively impacted their answers. A longitudinal study might be a potential solution to control any self-report bias. In a longitudinal study, each participant may be asked the same study questions different times, which could allow researchers to learn if and why variation occurred in participants' answers.

This study only tested the measures in a public university located in the Midwest U.S. Testing the measures at two different universities (e.g., private, public, religious, community college, research-oriented college, and so on) could help to discern whether any differences exist based on region and institution type. For instance, students at one type of university might be savvier and more comfortable using Twitter for academic work. As an aside, graduate students could be included in a future study to see if there is a statistically significant difference between undergraduate and graduate students' use of Twitter.

Some constructs of this study might be viewed differently by different students. For example, some students could think that their social lives should be their first and foremost priority whereas others might think that education is more important. Considering this, the subscale that focused on students' "Perception about Twitter use for their academics" could have conveyed a different meaning and in turn affected their answers while completing the survey.

Different options in the Likert scale might not have been so obvious to some participants (e.g., mixed feeling, disagree or strongly disagree). This situation might have led to cultural bias. A *cultural bias* occurs when participants do not focus on the variations on provided answers (i.e., Likert scale). Cultural differences could have an impact on response bias (Smith, 2004); for instance, while one group thinks "strongly agree" is appropriate, another can think that "agree" is more suitable. Further study is needed to understand whether subtle differences in Likert scale options are appropriate for most participants. This could include a Differential Item Functioning (DIF), which accounts for potential effects related to race/ethnicity, gender or college major, for example (Abad, Colom, Rebollo & Escorial, 2004; Crane, Belle & Larson, 2004; Van Dam, Earleywine & Forsyth, 2009).

Another potential limitation of this study could be response bias. A *response bias* occurs when participants knowingly or unknowingly report incorrect or less accurate data. In this study, participants reported their study time and amount of time they spent in Twitter for their academics and for non-academic purposes. Even when they provide reasonable estimations for these amounts of time, students might not know exactly how many minutes or hours they spent studying or using Twitter for academic and non-academic reasons. A future study should account closely for time spend on Twitter using some types of observations. A potential solution could be students' self-observation and recording of their study time and Twitter use time using a cell-phone app.

Summary and Conclusion

This study has developed measures to examine the use of Twitter based on a literature review. It also explored and confirmed the factor structures of the measures using both EFA and CFA, as detailed in Chapters II, III, and IV. Finally, this study checked both the validity and reliability of the newly developed measures using multiple statistical analyses. The TSSS and TSAA scales can serve as supplemental tools for correlational research studies in this field, as this will improve the validity of the scales. This study can likewise serve as a guideline by which future researchers can check validity and reliability of new scales. In all, this study's results could likewise serve as a reference study in the field of SNS as well as measuring students' academic performance in relation to use of SNS.

This study contributes to the literature. First, this study validates previous studies that reported that students' behavior towards SNS can help improve academic performance (Junco, Elavsky & Heiberger, 2013; Junco, Heiberger & Loken, 2011; Lin, Hoffman & Borengasser, 2013; Mao, 2014). Even so, some studies have contended that students' attitude and SNS use can negatively affect students' academic performance (Flanigan & Babchuk, 2015; Kirschner & Karpinski, 2010). This study provides evidence to the contrary. Ultimately, this study has demonstrated that both behavior and attitude toward Twitter use for academic purposes ought to be considered *together* so that this phenomenon can be more fully understood.

APPENDICES

APPENDIX A

SAMPLE DATA FROM PILOT STUDY

APPENDIX A

Sample Data from Pilot Study

Table 25

Sample data from the pilot study demographic information (gender, age, ethnicity)

Items	Options	Coding/(Answer)
Demography and Socioeconomic status		
Sex/Gender	Male	0
	Female	(1)
	Transgender	2
	Prefer not to respond	3
	Other____	4
What is your age in years? Please round in nearest whole point	--	(19)
What is your race/ethnicity?	White/Caucasian	(0)
	Black/African American	1
	American Indian/Native American	2
	Hispanic/Latino	3
	Asian/Pacific Islander	4
	Multi/Bi-racial	5
	Other________(please specify)	6

Table 26

Sample data from the pilot study demographic information (parents' education)

Items	Options	Coding/(Answer)
Demography and Socioeconomic status		
What is your mother's/legal guardian's highest level of education?	No High School Diploma	0
	High School Diploma	1
	Some College	(2)
	Bachelor's Degree	3
	Master's Degree	4
	Professional Degree (e.g., Ed.S.) or certificate	5
	Doctorate (Ph.D.) or Medical (M.D.) Degree	6
	Not sure	7
What is your father's/legal guardian's highest level of education?	No High School Diploma	0
	High School Diploma	1
	Some College	(2)
	Bachelor's Degree	3
	Master's Degree	4
	Professional Degree (e.g., Ed.S.) or certificate	5
	Doctorate (Ph.D.) or Medical (M.D.) Degree	6
	Not sure	7

Table 27

Sample data from the pilot study Academic Information

Items	Options	Coding/(Answer)
Academic Information		
If answer undergraduate then, what is your class level/status	Freshman	0
	Sophomore	(1)
	Junior	2
	Senior	3
Major	Social Science/ Humanities	(0)
	Science/technology/ engineering/math	1
Are you an international student?	No	(0)
	Yes	1
Approximately how many hours do you spend studying per day?	--	(4)
Are you involved in any extracurricular activities?	No	(0)
	Yes	1

Table 28

Sample Data Pilot Study Twitter Use

Items	Options	Coding/(Answer)
Twitter use		
How frequently do you check your Twitter account?	Never (0 time a day)	0
	Rarely (1 to 5 times a day)	1
	Sometimes (6 to 10 times a day)	(2)
	Frequently (11 to 15 times a day)	3
	Almost always (16+ times a day)	4
How many tweets do you post/send every day?	0	0
	1-5	(1)
	6-10	2
	11-15	3
	16-20	4
	20+	5
How many retweet/reply do you (send) every day?	0	0
	1-5	(1)
	6-10	2
	11-15	3
	16-20	4
	20+	5
Approximately how many hours do you spend on Twitter per day?	Not applicable	0
	0-1 hours	1
	2-5 hours	(2)
	6-10 hours	3
	11-15 hours	4
	16-20 hours	5
	20+ hours	6
How often do you use Twitter to communicate with your instructor per week?	Very often (11+ times)	3
	Often (6-10 times)	(2)
	Sometimes (1-5 times)	1

	Never (0 time)	0
How often do you use Twitter to communicate with your classmates/peers per week?	Very often (11+ times)	3
	Often (6-10 times)	2
	Sometimes (1-5 times)	(1)
	Never (0 time)	0

Table 29

Sample Data Pilot Study TSSS and Academic Performance

Items	Options	Coding/(Answer)
	Twitter and TSSS	
Multitasking		
I'm good at multitasking with Twitter	Strongly disagree	0
	Disagree	(1)
	Neutral/Mixed feeling	2
	Agree	3
	Strongly agree	4
I multitask with Twitter while studying	Strongly disagree	0
	Disagree	(1)
	Neutral/Mixed feeling	2
	Agree	3
	Strongly agree	4
I have Twitter up while doing homework	Strongly disagree	0
	Disagree	(1)
	Neutral/Mixed feeling	2
	Agree	3
	Strongly agree	4
Responsibility		
My academics are my main focus	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	3
	Strongly agree	(4)
I'm a responsible person about schoolwork	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
I only use Twitter when I have the time for it	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
Perception		
Twitter are time consuming	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
Twitter is an academic distractions	Strongly disagree	0
	Disagree	1

	Neutral/Mixed feeling	(2)
	Agree	3
	Strongly agree	4
Twitter decreases academic performance	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
Twitter takes time away from studying	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
Academic performance		
What is your current cumulative GPA?	--	(3.8)

Table 30

Sample Data Pilot Study TSAA and Academic Performance

Items	Options	Coding/(Answer)
Twitter and TSAA		
Schoolwork		
I use Twitter for schoolwork.	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)
	Strongly agree	4
I use Twitter to communicate with my classmates	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	(2)
	Agree	3
	Strongly agree	4
I use Twitter to communicate for group projects	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	(2)
	Agree	3
	Strongly agree	4
Procrastination and Leisure		
I use Twitter as a break while studying	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	(2)
	Agree	3
	Strongly agree	4
I use Twitter to procrastinate when I should be studying	Strongly disagree	0
	Disagree	(1)
	Neutral/Mixed feeling	2
	Agree	3
	Strongly agree	4
I use Twitter to procrastinate if I am struggling/get bored.	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	(2)
	Agree	3
	Strongly agree	4
I use Twitter as a free time activity.	Strongly disagree	0
	Disagree	1
	Neutral/Mixed feeling	2
	Agree	(3)

	Strongly agree	4
Academic performance		
What is your current cumulative GPA?	--	(3.8)

APPENDIX B

RATING SCALE AND WRIGHT MAP

APPENDIX B

Rating Scale and Wright Map

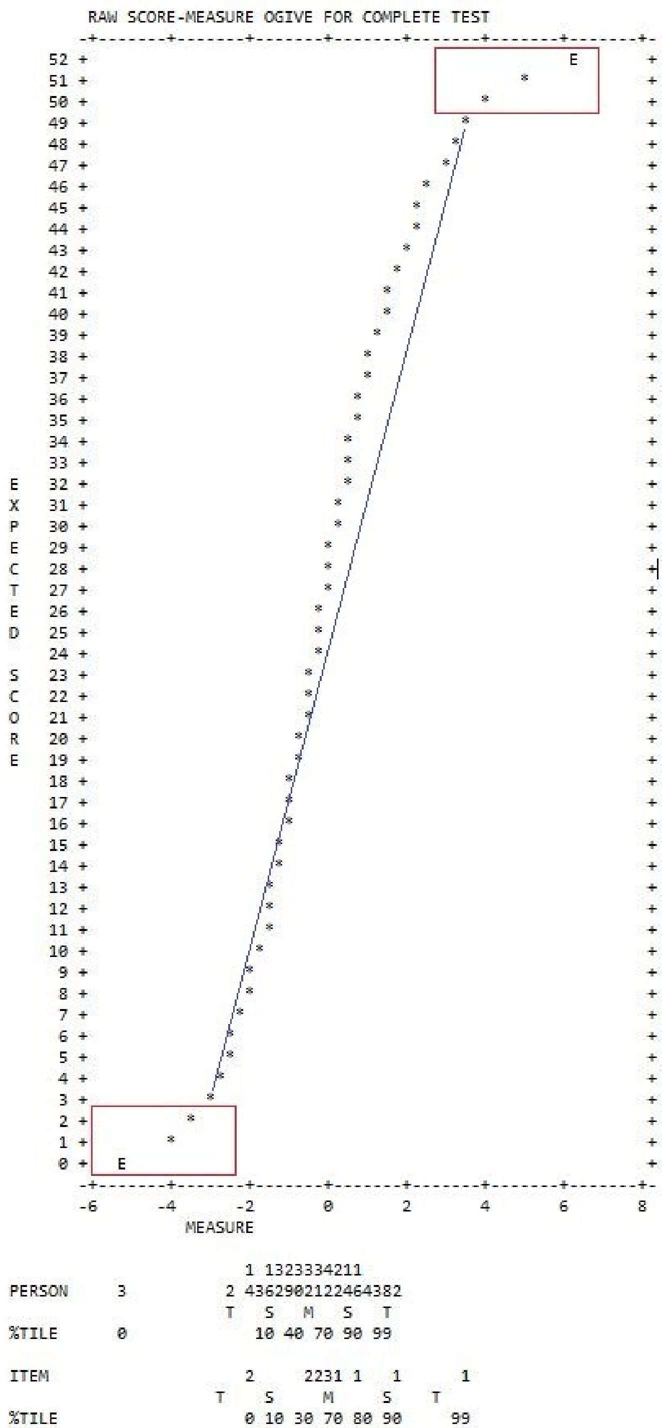

Figure 9. Rating scale of Twitter Scale Measures

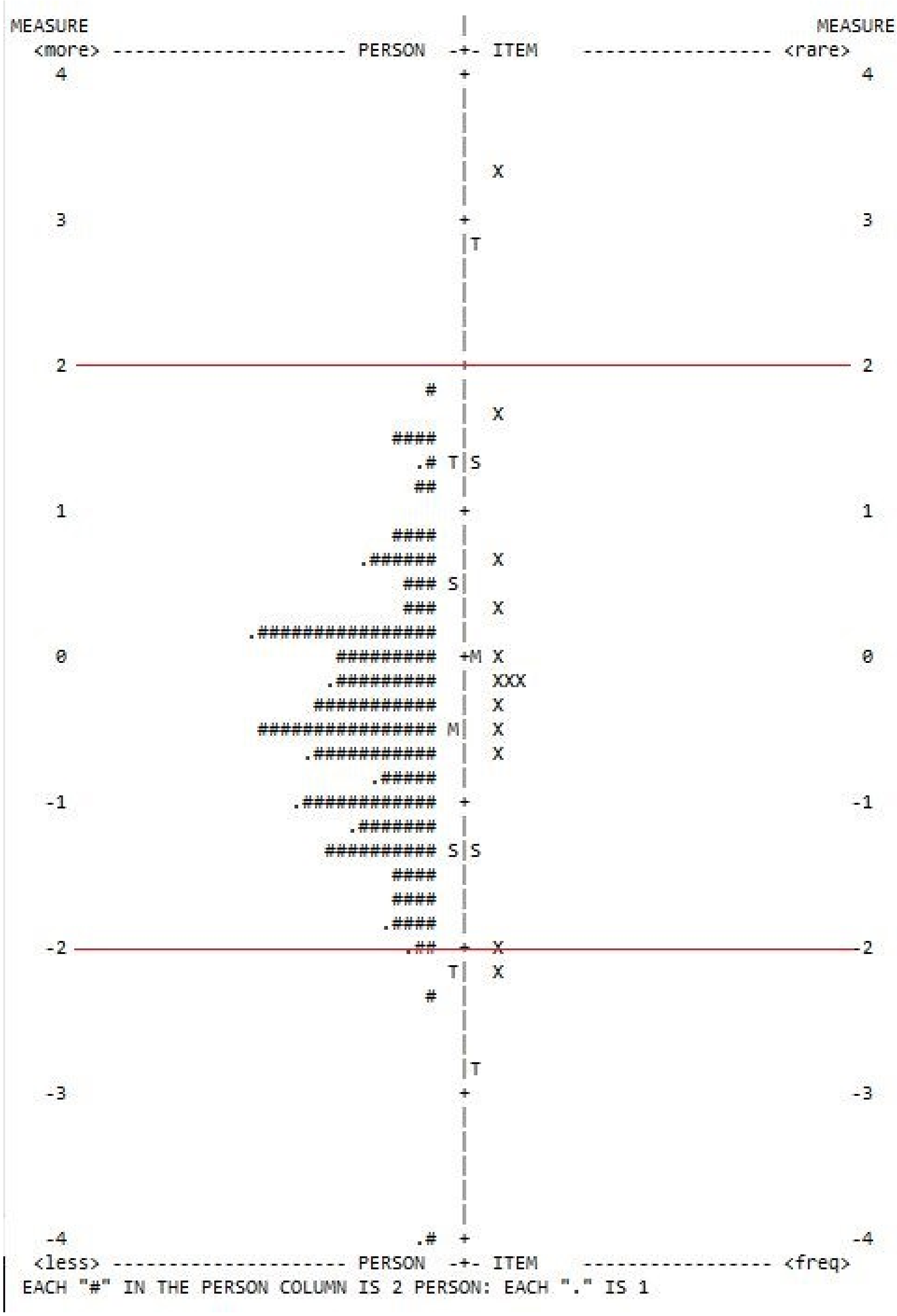

Figure 10. Wright map of the Twitter scales

REFERENCES

REFERENCES

Abad, F. J., Colom, R., Rebollo, I., & Escorial, S. (2004). Sex differential item functioning in the Raven's Advanced Progressive Matrices: Evidence for bias. *Personality and individual differences*, *36*(6), 1459-1470.

Abdulahi, A., Samadi, B., & Gharleghi, B. (2014). A study on the negative effects of social networking sites such as facebook among asia pacific university scholars in Malaysia. *International Journal of Business and Social Science*, *5*(10).

Aharony, N. (2010). Twitter use in libraries: An exploratory analysis. *Journal of Web Librarianship*, *4*(4), 333-350.

Ahn, J. (2011). The effect of social network sites on adolescents' social and academic development: Current theories and controversies. *Journal of the Association for Information Science and Technology*, *62*(8), 1435-1445.

Ainin, S., Naqshbandi, M. M., Moghavvemi, S., & Jaafar, N. I. (2015). Facebook usage, socialization and academic performance. *Computers & Education*, *83*, 64-73.

Ajzen, I. (1987). Attitudes, traits, and actions: Dispositional prediction of behavior in personality and social psychology. In *Advances in experimental social psychology* (Vol. 20, pp. 1-63). Academic Press.

Ajzen, I., & Fishbein, M. (1977). Attitude-behavior relations: A theoretical analysis and review of empirical research. *Psychological bulletin*, *84*(5), 888.

Alhabash, S., & Ma, M. (2017). A tale of four platforms: Motivations and uses of Facebook, Twitter, Instagram, and Snapchat among college students?. *Social Media+ Society*, *3*(1), 2056305117691544.

Allen, M.J., & Yen, W.M. (1979). *Introduction to measurement theory.* Monterey, CA: Brooks/Cole Publishing.

Amiruzzaman, M., Amiruzzaman, S., Ozer-Guclu, I., & Karpinski, A.C. (2019, April). *Twitter and academic performance multitasking: Frequency and attitude scale development*. Paper presented at the meeting of the American Educational Research Association, Toronto, Canada.

Anastasi, A. (1986). Evolving concepts of test validation. *Annual Review of Psychology*, *37*, 1–15.

Anderson, J. C., & Gerbing, D. W. (1984). The effect of sampling error on convergence, improper solutions, and goodness-of-fit indices for maximum likelihood confirmatory factor analysis. *Psychometrika*, *49*(2), 155-173.

Anderson, M., & Jiang, J. (2018). Teens, social media & technology 2018. *Pew Research Center*, *31*.

Arceneaux, N., & Schmitz Weiss, A. (2010). Seems stupid until you try it: Press coverage of Twitter, 2006-9. *New media & society*, *12*(8), 1262-1279.

Aslam, S. (2017). Twitter by the numbers: Stats, demographics & fun facts. Retrieved from https://www.omnicoreagency.com/twitter-statistics/. Retrieved on July 3, 2017.

Baghaei, P., & Amrahi, N. (2011). Validation of a multiple choice English vocabulary test with the Rasch Model. *Journal of Language Teaching & Research*, *2*(5).

Bartlett, M. S. (1950). Tests of significance in factor analysis. *British Journal of Psychology, Statistical Section*, *3*, 77-85.

Basak, R., Sural, S., Ganguly, N., & Ghosh, S. K. (2019). Online Public Shaming on Twitter: Detection, Analysis, and Mitigation. *IEEE Transactions on Computational Social Systems*.

Beldarrain, Y. (2006). Distance education trends: Integrating new technologies to foster student interaction and collaboration. *Distance education*, *27*(2), 139-153.

Bond, T. G., & Fox, C. M. (2015). *Applying the Rasch model: Fundamental measurement in*

Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how?. *CBE—Life Sciences Education*, *15*(4), rm4.

Brecht, H. D. (2012). Learning from online video lectures. *Journal of Information Technology Education*, *11*(1), 227-250.

Brentani, E., & Golia, S. (2007). Unidimensionality in the Rasch model: how to detect and interpret. *Statistica*, *67*(3), 253-261.

Brink, Y., Louw, Q., & Grimmer-Somers, K. (2011). The quality of evidence of psychometric properties of three-dimensional spinal posture-measuring instruments. *BMC musculoskeletal disorders*, *12*(1), 93.

Byrne, Barbara M. (1984). The general/academic self-concept nomological network: A review of construct validation research. *Review of educational research*. *54*(3). 427-456.

Cain, J. (2008). Online social networking issues within academia and pharmacy education. *American journal of pharmaceutical education*, *72*(1), 10.

Carpenter, J. P., & Krutka, D. G. (2015). Engagement through microblogging: Educator professional development via Twitter. *Professional development in education*, *41*(4), 707-728.

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research*, *1*(2), 245-276.

Chaffey, D. (2019). Global social media research summary 2019. Retrieved from, https://www.smartinsights.com/social-media-marketing/social-media-strategy/new-global-social-media-research/, Retrieved on, May 17, 2019.

Chaiken, S., & Baldwin, M. W. (1981). Affective-cognitive consistency and the effect of salient behavioral information on the self-perception of attitudes. *Journal of Personality and Social Psychology*, *41*(1), 1.

Chang, T. S., & Hsiao, W. H. (2014). Time spent on social networking sites: Understanding user behavior and social capital. *Systems Research and Behavioral Science*, *31*(1), 102-114.

Chun Chu, A. H., & Choi, J. N. (2005). Rethinking procrastination: Positive effects of “active” procrastination behavior on attitudes and performance. *The Journal of social psychology*, *145*(3), 245-264.

Coad, D. T. (2017). “That’s My Face to the Whole Field!”: Graduate Students’ Professional Identity-Building through Twitter at a Writing Studies Conference. *Computers and Composition*, *45*, 51-66.

Crane, P. K., Belle, G. V., & Larson, E. B. (2004). Test bias in a cognitive test: differential item functioning in the CASI. *Statistics in Medicine*, *23*(2), 241-256.

Credé, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on psychological science*, *3*(6), 425-453.

Crocker, L., & Algina, J. (2008). *Introduction to classical & modern test theory* (2nd ed.). Independence, KY: Cengage Learning.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*(3), 297-334.

Cronbach, L. J. (1971). *Test validity*. In R. L. Thorndike (ed.), Educational Measurement (2nd ed., pp. 440-507). Washington, DC: American Council on Education.

Cronbach, L.J., & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281 – 302.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological methods*, *1*(1), 16.

Danciu, E., & Grosseck, G. (2011). Social aspects of web 2.0 technologies: Teaching or teachers' challenges?. *Procedia-Social and Behavioral Sciences*, *15*, 3768-3773.

Das, B., & Sahoo, J. S. (2011). Social networking sites–a critical analysis of its impact on personal and social life. *International Journal of Business and Social Science*, *2*(14), 222-228.

Davis, F. D. (1985). *A technology acceptance model for empirically testing new end-user information systems: Theory and results* (Doctoral dissertation, Massachusetts Institute of Technology).

Davis, K. (2015). Teachers' perceptions of Twitter for professional development. *Disability and rehabilitation*, *37*(17), 1551-1558.

Deming, W. Edwards. *Sample design in business research*. Vol. 23. John Wiley & Sons, 1990.

de Jong-Gierveld, J., & Kamphuls, F. (1985). The development of a Rasch-type loneliness scale. *Applied psychological measurement*, *9*(3), 289-299.

De Winter, J. C., & Dodou, D. (2012). Factor recovery by principal axis factoring and maximum likelihood factor analysis as a function of factor pattern and sample size. *Journal of Applied Statistics*, *39*(4), 695-710.

DeAndrea, D. C., Ellison, N. B., LaRose, R., Steinfield, C., & Fiore, A. (2012). Serious social media: On the use of social media for improving students' adjustment to college. *The Internet and higher education*, *15*(1), 15-23.

Denker, K. J., Manning, J., Heuett, K. B., & Summers, M. E. (2018). Twitter in the classroom: Modeling online communication attitudes and student motivations to connect. *Computers in Human Behavior*, *79*, 1-8.

Devilly, G. J., & Borkovec, T. D. (2000). Psychometric properties of the credibility/expectancy questionnaire. *Journal of behavior therapy and experimental psychiatry*, *31*(2), 73-86.

Dietz, S., & Henrich, C. (2014). Texting as a distraction to learning in college students. *Computers in Human behavior*, *36*, 163-167.

Dimitrov, D. M. (2013) *Quantitative research in education: Intermediate & advanced methods*. Oceanside, NY: Whittier Publications.

Donaldson, S. I., & Grant-Vallone, E. J. (2002). Understanding self-report bias in organizational behavior research. *Journal of business and Psychology*, *17*(2), 245-260.

Duncan, P. W., Bode, R. K., Lai, S. M., & Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: the Stroke Impact Scale. *Archives of physical medicine and rehabilitation*, *84*(7), 950-963.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook "friends:" Social capital and college students' use of online social network sites. *Journal of computer-mediated communication*, *12*(4), 1143-1168.

Engelhard Jr, G. (1994). Examining rater errors in the assessment of written composition with a many-faceted Rasch model. *Journal of Educational Measurement*, *31*(2), 93-112.

Evans, C., Hackney, R., Rauniar, R., Rawski, G., Yang, J., & Johnson, B. (2014). Technology acceptance model (TAM) and social media usage: an empirical study on Facebook. *Journal of Enterprise Information Management*.

Fewkes, A. M., & McCabe, M. (2012). Facebook: Learning tool or distraction?. *Journal of Digital Learning in Teacher Education*, *28*(3), 92-98.

Field, A. (2014). *Discovering statistics using IBM SPSS statistics 4th edition.* London: Sage Publications.

Flanigan, A. E., & Babchuk, W. A. (2015). Social media as academic quicksand: A phenomenological study of student experiences in and out of the classroom. *Learning and Individual differences*, *44*, 40-45.

Flanigan, A. E., & Kiewra, K. A. (2018). What college instructors can do about student cyber-slacking. *Educational Psychology Review*, *30*(2), 585-597.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological assessment*, *7*(3), 286.

Fonseca, D., Martí, N., Redondo, E., Navarro, I., & Sánchez, A. (2014). Relationship between student profile, tool use, participation, and academic performance with the use of Augmented Reality technology for visualized architecture models. *Computers in human behavior*, *31*, 434-445.

Forbush, E., & Foucault-Welles, B. (2016). Social media use and adaptation among Chinese students beginning to study in the United States. *International Journal of Intercultural Relations, 50*, 1-12.

Fowler, F.J. (2002). *Survey research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Franchignoni, F., Giordano, A., Ferriero, G., Orlandini, D., Amoresano, A., & Perucca, L. (2007). Measuring mobility in people with lower limb amputation: Rasch analysis of the mobility section of the prosthesis evaluation questionnaire. *Journal of rehabilitation medicine*, *39*(2), 138-144.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*, *38*(3), 474-480.

Fuligni, A. J. (1997). The academic achievement of adolescents from immigrant families: The role of family background, attitudes, and behavior. *Child development*, *68*(2), 351-363.

Furrer, O., Egri, C. P., Ralston, D. A., Danis, W. M., Reynaud, E., Naoumova, I., ... & Furrer-Perrinjaquet, A. (2010). Attitudes toward corporate responsibilities in Western Europe and in Central and East Europe. *Management International Review*, *50*(3), 379-398.

Gabre, H., & Kumar, G. (2012). The effects of perceived stress and Facebook on accounting students' academic performance. *Accounting and Finance Research, 1*(2), 87-100.

Gao, X., Cao, Z., Li, S., Yao, B., Chen, G., & Tang, S. (2019). Taxonomy and Evaluation for Microblog Popularity Prediction. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, *13*(2), 15.

Gonzalez, R., Gasco, J., & Llopis, J. (2019). University students and online social networks: Effects and typology. *Journal of Business Research*.

Gorsuch, R. L. (1997). Exploratory factor analysis: Its role in item analysis. *Journal of personality assessment*, *68*(3), 532-560.

Grosseck, G., & Holotescu, C. (2008, April). Can we use Twitter for educational activities. In *4th international scientific conference, eLearning and software for education, Bucharest, Romania*.

Haley, S. M., McHorney, C. A., & Ware Jr, J. E. (1994). Evaluation of the MOS SF-36 physical functioning scale (PF-10): I. Unidimensionality and reproducibility of the Rasch item scale. *Journal of clinical epidemiology*, *47*(6), 671-684.

Halpern, D., & Gibbs, J. (2013). Social media as a catalyst for online deliberation? Exploring the affordances of Facebook and YouTube for political expression. *Computers in Human Behavior*, *29*(3), 1159-1168.

Hargittai, E. (2010). Digital na(t)ives? Variation in internet skills and uses among members of the “net generation”. *Sociological inquiry*, *80*(1), 92-113.

Hargittai, E., & Hsieh, Y. L. P. (2010). Predictors and consequences of differentiated practices on social network sites. *Information, Communication & Society*, *13*(4), 515-536.

Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational research methods*, *7*(2), 191-205.

Hew, K. F. (2011). Students’ and teachers’ use of Facebook. *Computers in human behavior*, *27*(2), 662-676. *human sciences* (3rd ed.). New York, NY: Routledge.

identity-building through Twitter at a writing studies conference. *Computers and Composition*, *45*, 51-66.

Hiller, C. E., Refshauge, K. M., Bundy, A. C., Herbert, R. D., & Kilbreath, S. L. (2006). The Cumberland ankle instability tool: a report of validity and reliability testing. *Archives of physical medicine and rehabilitation*, *87*(9), 1235-1241.

Holmes, S. E. (1982). Unidimensionality and vertical equating with the Rasch model. *Journal of Educational Measurement*, 139-147.

Jacobsen, W. C., & Forste, R. (2011). The wired generation: Academic and social outcomes of electronic media use among university students. *Cyberpsychology, Behavior, and Social Networking*, *14*(5), 275-280.

Jansen, B. J., Zhang, M., Sobel, K., & Chowdury, A. (2009). Twitter power: Tweets as electronic word of mouth. *Journal of the American society for information science and technology*, *60*(11), 2169-2188.

Jennrich, R. I., & Sampson, P. F. (1966). Rotation for simple loadings. *Psychometrika*, *31*(3), 313-323.

Jöreskog, K. G., & Sörbom, D. (1989). *LISREL 7: A guide to the program and applications*. Spss.

Junco, R. (2012). The relationship between frequency of Facebook use, participation in Facebook activities, and student engagement. *Computers & Education*, *58*(1), 162-171.

Junco, R., & Cotten, S. R. (2012). No A 4 U: The relationship between multitasking and academic performance. *Computers and Education, 59*(2), 505.

Junco, R., Elavsky, C. M., & Heiberger, G. (2013). Putting twitter to the test: Assessing outcomes for student collaboration, engagement and success. *British Journal of Educational Technology*, *44*(2), 273-287.

Junco, R., Heiberger, G., & Loken, E. (2011). The effect of Twitter on college student engagement and grades. *Journal of computer assisted learning*, *27*(2), 119-132.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, *39*(1), 31-36.

Kane, M. T (2006). Validation. In R. Brennan (Ed.), *Educational measurement,* 4th ed. (pp. 17–64), Westport, CT: American Council on Education and Praeger.

Kane, M. T. (2011). Validating score interpretations and use: Messick lecture. In *Language Testing Research Colloquium, Cambridge: SAGE*.

Karanicolas, P. J., Bhandari, M., Kreder, H., Moroni, A., Richardson, M., Walter, S. D., ... & Collaboration for Outcome Assessment in Surgical Trials (COAST) Musculoskeletal Group. (2009). Evaluating agreement: conducting a reliability study. *JBJS*, *91*(Supplement_3), 99-106.

Karpinski, A. C., Kirschner, P. A., Ozer, I., Mellott, J. A., & Ochwo, P. (2013). An exploration of social networking site use, multitasking, and academic performance among United States and European university students. *Computers in Human Behavior*, *29*(3), 1182-1192.

Kassens-Noor, E. (2012). Twitter as a teaching practice to enhance active and informal learning in higher education: The case of sustainable tweets. *Active Learning in Higher Education*, *13*(1), 9-21.

Keith, T. (2006). *Multiple regression and beyond.* Boston: Pearson.

Kim, B., & Kim, Y. (2019). Growing as Social Beings: How Social Media Use for College Sports Is Associated with College Students' Group Identity and Collective Self-esteem. *Computers in Human Behavior*.

King, D. W., Leskin, G. A., King, L. A., & Weathers, F. W. (1998). Confirmatory factor analysis of the clinician-administered PTSD scale: Evidence for the dimensionality of posttraumatic stress disorder. *Psychological Assessment*, *10*(2), 90.

Kirschner, P. A., & Karpinski, A. C. (2010). Facebook® and academic performance. *Computers in Human Behavior, 26*(6), 1237-1245.

Kline, P. (2014). *An easy guide to factor analysis*. Routledge.

Koranteng, F. N., Wiafe, I., & Kuada, E. (2018). An Empirical Study of the Relationship Between Social Networking Sites and Students' Engagement in Higher Education. *Journal of Educational Computing Research*, 0735633118787528.

Kosko, K. W. (2019). A multiplicative reasoning assessment for fourth and fifth grade students. *Studies in Educational Evaluation*, 60, 32-42.

Kuss, D., & Griffiths, M. (2017). Social networking sites and addiction: Ten lessons learned. *International journal of environmental research and public health*, *14*(3), 311.

Kwon, S. J., Park, E., & Kim, K. J. (2014). What drives successful social networking services? A comparative analysis of user acceptance of Facebook and Twitter. *The Social Science Journal*, *51*(4), 534-544.

Lai, J. S., Fisher, A. G., Magalhães, L. C., & Bundy, A. C. (1996). Construct validity of the sensory integration and praxis tests. *The Occupational Therapy Journal of Research*, *16*(2), 75-97.

Lassen, S. R., Steele, M. M., & Sailor, W. (2006). The relationship of school-wide positive behavior support to academic achievement in an urban middle school. *Psychology in the Schools*, *43*(6), 701-712.

Lau, W. W. (2017). Effects of social media usage and social media multitasking on the academic performance of university students. *Computers in human behavior*, *68*, 286-291.

Lee, A. R., Son, S. M., & Kim, K. K. (2016). Information and communication technology overload and social networking service fatigue: A stress perspective. *Computers in Human Behavior*, *55*, 51-61.

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2014). The relationship between cell phone use, academic performance, anxiety, and satisfaction with life in college students. *Computers in Human Behavior*, *31*, 343-350.

Lian, S. L., Sun, X. J., Zhou, Z. K., Fan, C. Y., Niu, G. F., & Liu, Q. Q. (2018). Social networking site addiction and undergraduate students' irrational procrastination: The mediating role of social networking site fatigue and the moderating role of effortful control. *PloS one*, *13*(12), e0208162.

Lin, M. F. G., Hoffman, E. S., & Borengasser, C. (2013). Is social media too social for class? A case study of Twitter use. *TechTrends*, *57*(2), 39.

Linacre, J. M. (2005). *User's guide to Winsteps*. Chicago: Mesa Press.

Linacre, J.M. (2006), WINSTEPS Rasch measurement computer program, Chicago: Winsteps.com.

Liu, Y., & Tsai, Y. R. (2012). The impact of social networking services (SNS) on college students' social relationship and private life. *International Journal of Arts and Commerce*, *1*(4), 1-10.

Lomax, R., & Hahs-Vaughn, D. (2012). *An introduction to statistical concepts.* Taylor & Francis Group: New York, NY.

Lubis, S. H., Ridzuan, S., Ishak, I. Y., Othman, H. F., Mohammed, N., Hamid, Z. A., & Izham, M. (2012). The relationship between time spent on Facebook and cumulative grade point average (CGPA) among third year biomedical science students in faculty health sciences. *Procedia-Social and Behavioral Sciences, 60*, 590-595.

MacCallum, R. C., & Hong, S. (1997). Power analysis in covariance structure modeling using GFI and AGFI. *Multivariate Behavioral Research*, *32*(2), 193-210.

Moghavvemi, S., Sulaiman, A., Jaafar, N. I., & Kasem, N. (2018). Social media as a complementary learning tool for teaching and learning: The case of youtube. *The International Journal of Management Education*, *16*(1), 37-42.

Mao, J. (2014). Social media for learning: A mixed methods study on high school students' technology affordances and perspectives. *Computers in Human Behavior*, *33*, 213-223.

Marsh, H. W., & Hocevar, D. (1985). Application of confirmatory factor analysis to the study of self-concept: First-and higher order factor models and their invariance across groups. *Psychological bulletin*, *97*(3), 562.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149-174.

Mc Mahon, C. (2014). Why We Need a New Theory of Cyberbullying. *Available at SSRN 2531796*.

Mehmood, S., & Taswir, T. (2013). The effects of social networking sites on the academic performance of students in college of applied sciences, Nizwa, Oman. *International Journal of Arts and Commerce*, *2*(1), 111-125.

Meier, A., Reinecke, L., & Meltzer, C. E. (2016). "Facebocrastination"? Predictors of using Facebook for procrastination and its effects on students' well-being. *Computers in Human Behavior*, *64*, 65-76.

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.

Michikyan, M., Subrahmanyam, K., & Dennis, J. (2015). Facebook use and academic performance among college students: A mixed-methods study with a multi-ethnic sample. *Computers in Human Behavior*, *45*, 265-272.

Minocha, S. (2009). A case study-based investigation of students' experiences with social software tools. *New review of Hypermedia and Multimedia*, *15*(3), 245-265.

Moqbel, M., & Kock, N. (2018). Unveiling the dark side of social networking sites: Personal and work-related consequences of social networking site addiction. *Information & Management*, *55*(1), 109-119.

Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E., & Becker, T. (2011). Feeling bad on Facebook: depression disclosures by college students on a social networking site. *Depression and anxiety*, *28*(6), 447-455.

Networking, 14(5), 275-280.

Nam, S. K., Yang, E., Lee, S. M., Lee, S. H., & Seol, H. (2011). A psychometric evaluation of the career decision self-efficacy scale with Korean students: A Rasch model approach. *Journal of Career development*, *38*(2), 147-166.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences*, *106*(37), 15583-15587.

Ophus, J. D., & Abbitt, J. T. (2009). Exploring the potential perceptions of social networking systems in university courses. *Journal of Online Learning and Teaching*, *5*(4), 639-648.

Ostroff, C. (1992). The relationship between satisfaction, attitudes, and performance: An organizational level analysis. *Journal of applied psychology*, *77*(6), 963.

Ozer, I., Karpinski, A. C., & Kirschner, P. A. (2014). A cross-cultural qualitative examination of social-networking sites and academic performance. *Procedia-Social and Behavioral Sciences*, *112*, 873-881.

Park, J. H. (2010). Differences among university students and faculties in social networking site perception and use: Implications for academic library services. *The Electronic Library*, *28*(3), 417-431.

Parra, D., Trattner, C., Gómez, D., Hurtado, M., Wen, X., & Lin, Y. R. (2016). Twitter in academic events: A study of temporal usage, communication, sentimental and topical patterns in 16 computer science conferences. *Computer Communications*, *73*, 301-314.

Paul, J. A., Baker, H. M., & Cochran, J. D. (2012). Effect of online social networking on student academic performance. *Computers in Human Behavior*, *28*(6), 2117-2127.

Pempek, T. A., Yermolayeva, Y. A., & Calvert, S. L. (2009). College students' social networking experiences on Facebook. *Journal of applied developmental psychology*, *30*(3), 227-238.

Petrocelli, J. V. (2003). Hierarchical multiple regression in counseling research: Common problems and possible remedies. *Measurement and evaluation in counseling and development*, *36*(1), 9-22.

Pett, M.A., Lackey, N.R., & Sullivan, J.J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research.* Thousand Oaks, CA: Sage Publications, Inc.

Poole, M. A., & O'Farrell, P. N. (1971). The assumptions of the linear regression model. *Transactions of the Institute of British Geographers*, 145-158.

Raacke, J., & Bonds-Raacke, J. (2008). MySpace and Facebook: Applying the uses and gratifications theory to exploring friend-networking sites. *Cyberpsychology & behavior*, *11*(2), 169-174.

Rahman, S. H., & Stephen, J. (2016). OMG (oh my grade)! Social networking sites ruin my academic grades. *Pertanika Journal of Social Sciences & Humanities*, *24*(1), 483-494.

Rankin, M. (2009). *The Twitter experiment at UT Dallas*. Retrieved from, http://www.utdallas.edu/~mrankin/usweb/twitterconclusions.htm. Retrieved on, August 29, 2019.

Raykov, T. (1998). On the use of confirmatory factor analysis in personality research. *Personality and Individual Differences*, *24*(2), 291-293.

Reuterberg, S. E., & Gustafsson, J. E. (1992). Confirmatory factor analysis and reliability: Testing measurement model assumptions. *Educational and Psychological Measurement*, *52*(4), 795-811.

Rittle-Johnson, B., Matthews, P. G., Taylor, R. S., & McEldoon, K. L. (2011). Assessing knowledge of mathematical equivalence: A construct-modeling approach. *Journal of Educational Psychology*, *103*(1), 85.

Rogers, W. T. (1976). Jackknifing disattenuated correlations. *Psychometrika*, *41*(1), 121-133.

Roopchund, R., Ramesh, V., & Jaunky, V. (2019). Use of Social Media for Improving Student Engagement at Université des Mascareignes (UDM). In *Information Systems Design and Intelligent Applications* (pp. 11-20). Springer, Singapore.

Rosen, L. D., Carrier, L. M., & Cheever, N. A. (2013). Facebook and texting made me do it: Media-induced task-switching while studying. *Computers in Human Behavior*, *29*(3), 948-958.

Rosen, L. D., Whaling, K., Rab, S., Carrier, L. M., & Cheever, N. A. (2013). Is Facebook creating “iDisorders”? The link between clinical symptoms of psychiatric disorders and technology use, attitudes and anxiety. *Computers in Human Behavior*, *29*(3), 1243-1254.

Rouis, S. (2012). Impact of cognitive absorption on Facebook on students' achievement. *Cyberpsychology, Behavior, and Social Networking*, *15*(6), 296-303.

Samad, S., Nilashi, M., & Ibrahim, O. (2019). The impact of social networking sites on students' social wellbeing and academic performance. *Education and Information Technologies*, 1-14.

Samaha, M., & Hawi, N. S. (2016). Relationships among smartphone addiction, stress, academic performance, and satisfaction with life. *Computers in Human Behavior*, *57*, 321-325.

Schmitt, J., Langan, S., Williams, H. C., & Network, E. D. E. (2007). What are the best outcome measurements for atopic eczema? A systematic review. *Journal of Allergy and Clinical Immunology*, *120*(6), 1389-1398.

Selwyn, N. (2009). Faceworking: Exploring students' education-related use of Facebook. *Learning, Media and Technology, 34*(2), 157-174.

Sevier, F. A. (1957). Testing the assumptions underlying multiple regression. *The Journal of Experimental Education*, *25*(4), 323-330.

Sheldon, P. (2012). Profiling the non-users: Examination of life-position indicators, sensation seeking, shyness, and loneliness among users and non-users of social network sites. *Computers in Human Behavior*, *28*(5), 1960-1965.

Shrigley, R. L. (1990). Attitude and behavior are correlates. *Journal of research in Science Teaching*, *27*(2), 97-113.

Singh, R., & Teoh, J. B. P. (2000). Impression formation from intellectual and social traits: Evidence for behavioural adaptation and cognitive processing. *British Journal of Social Psychology*, *39*(4), 537-554.

Slavin, R. E. (1978). Student teams and comparison among equals: Effects on academic performance and student attitudes. *Journal of Educational Psychology*, *70*(4), 532.

Smith, P. B. (2004). Acquiescent response bias as an aspect of cultural communication style. *Journal of Cross-Cultural Psychology*, *35*(1), 50-61.Smith, R. M. (1986). Person fit in the Rasch model. *Educational and Psychological Measurement*, *46*(2), 359-372.

Smith, R. M. (1996). A comparison of methods for determining dimensionality in Rasch measurement. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(1), 25-40.

Soukup, C. (2006). Computer-mediated communication as a virtual third place: building Oldenburg's great good places on the world wide web. *New Media & Society*, *8*(3), 421-440.

Steelman, Z., Soror, A., Limayem, M., & Worrell, D. (2012). Obsessive compulsive tendencies as predictors of dangerous mobile phone usage.

Subrahmanyam, K., Reich, S. M., Waechter, N., & Espinoza, G. (2008). Online and offline social networks: Use of social networking sites by emerging adults. *Journal of applied developmental psychology*, *29*(6), 420-433.

Tabachnick, B. G., & Fidell, L. S. (2001). Principal components and factor analysis. *Using multivariate statistics*, *4*, 582-633.

Tanaka, A., Murakami, Y., Okuno, T., & Yamauchi, H. (2002). Achievement goals, attitudes toward help seeking, and help-seeking behavior in the classroom. *Learning and Individual Differences*, *13*(1), 23-35.

Tang, Y., & Hew, K. F. (2017). Using Twitter for education: Beneficial or simply a waste of time?. *Computers & Education*, *106*, 97-118.

Thompson, B. (2004). *Exploratory and confirmatory factor analysis: Understanding concepts and applications.* Washington, D.C.: American Psychological Association.

Thompson, C. (2007). How Twitter creates a social sixth sense. *Wired Magazine*, *15*(7).

Tiernan, P. (2014). A study of the use of Twitter by students for lecture engagement and discussion. *Education and Information Technologies*, *19*(4), 673-690.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. *Journal of counseling psychology*, *34*(4), 414.

Van Dam, N. T., Earleywine, M., & Forsyth, J. P. (2009). Gender bias in the sixteen-item Anxiety Sensitivity Index: An application of polytomous differential item functioning. *Journal of Anxiety Disorders*, *23*(2), 256-259.

Van Der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior*, *53*, 204-215.

Vassend, O., & Skrondal, A. (1997). Validation of the NEO Personality Inventory and the five-factor model. Can findings from exploratory and confirmatory factor analysis be reconciled?. *European Journal of Personality*, *11*(2), 147-166.

Vie, S. (2008). Digital divide 2.0:“Generation M” and online social networking sites in the composition classroom. *Computers and Composition*, *25*(1), 9-23.

Villavicencio, F. T., & Bernardo, A. B. (2013). Positive academic emotions moderate the relationship between self-regulation and academic achievement. *British Journal of Educational Psychology*, *83*(2), 329-340.

Vis, F. (2013). Twitter as a reporting tool for breaking news: Journalists tweeting the 2011 UK riots. *Digital journalism*, *1*(1), 27-47.

Visser, R. D., Evering, L. C., & Barrett, D. E. (2014). #TwitterforTeachers: The implications of Twitter as a self-directed professional development tool for K–12 teachers. *Journal of Research on Technology in Education*, *46*(4), 396-413.

Wahler, R. G. (1975). Some structural aspects of deviant child behavior 1. *Journal of Applied Behavior Analysis*, *8*(1), 27-42.

Wang, J. L., Jackson, L. A., Gaskin, J., & Wang, H. Z. (2014). The effects of social networking site (SNS) use on college students' friendship and well-being. *Computers in Human Behavior*, *37*, 229-236.

Wang, P., Wang, X., Wu, Y., Xie, X., Wang, X., Zhao, F., ... & Lei, L. (2018). Social networking sites addiction and adolescent depression: A moderated mediation model of rumination and self-esteem. *Personality and Individual Differences*, *127*, 162-167.

Wells, T., & Link, M. (2014). Facebook user research using a probability-based sample and behavioral data. *Journal of Computer-Mediated Communication*, *19*(4), 1042-1052.

Wentzel, K. R. (1991). Social competence at school: Relation between social responsibility and academic achievement. *Review of educational research*, *61*(1), 1-24.

Wilson, M. (2004). *Constructing measures: An item response modeling approach*. Routledge.

Wilkinson, L., & Task Force on Statistical Inference, American Psychological Association, Science Directorate. (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*(8), 594-604. http://dx.doi.org/10.1037/0003-066X.54.8.594

Wise, K., Alhabash, S., & Park, H. (2010). Emotional responses during social information seeking on Facebook. *Cyberpsychology, Behavior, and Social Networking*, *13*(5), 555-562.

Wolfe, E. W., & Smith, J. E. (2007). Instrument development tools and activities for measure validation using Rasch models: part II--validation activities. *Journal of applied measurement*, *8*(2), 204-234.

Wright, B. D. (1996). Comparing Rasch measurement and factor analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *3*(1), 3-24.

Wright, B. D., & Linacre, J. M. (1994). Reasonable mean-square fit values. Rasch Measurement Transactions, 8, 370. http://www.rasch.org/rmt/rmt83b.htm.

Wright, B.D., & Masters, G.N. (1982), Rating Scale Analysis, MESA Press, Chicago.

Young, J. (2009a). *Teaching with Twitter: not for the faint of heart.* Retrieved from, http://chronicle.com/article/Teaching-With-Twitter-Not-for/49230/. Retrieved on, August 29, 2019.