12-10-2015 12:00 AM

Social Networking Sites and Personnel Selection: An Initial Validity Assessment

Travis J. Schneider, *The University of Western Ontario*

Supervisor: Dr. Richard Goffin, *The University of Western Ontario*

A thesis submitted in partial fulfillment of the requirements for the Doctor of Philosophy degree in Psychology

Social Networking Sites and Personnel Selection: An Initial Validity Assessment

(Thesis format: Integrated Article)

by

Travis Schneider

Graduate Program in Psychology

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in Industrial/Organizational Psychology

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

The purpose of this dissertation was to add to the literature on the use of social networking sites (SNSs) for personnel selection. The first goal was to evaluate whether SNSs have the potential to be used as a valid source of information for selection. Specific SNS Indicator scales were created to test whether they have better validity evidence than the more traditionally-used Global SNS Rating. In a study of 141 undergraduate students at a large Canadian university, the Specific SNS Indicators demonstrated fairly weak evidence of interrater reliability, but some evidence of structural validity, and construct validity (convergent and discriminant). Specific SNS Indicator scales offered incremental prediction beyond a Global SNS Rating in the prediction of their traditional selection scales. This supported previous research suggesting that making info-processing easier for raters should lead to better validity (Heneman, 1986). The second goal was to investigate some of the practical realities of attaining SNS information. In a study of 892 employed or previously-employed participants, participants were given a hypothetical employer request for their SNS password as part of a job application. The study found that 57.87% applicants would refuse the request, and would thus remove themselves from the applicant pool. This substantial loss of applicants could lower the utility of an organization's selection system and thereby impact an organization's bottom line. Also, the password request could result in adverse impact for protected groups, and lead to legal action on behalf of applicants within these groups. There was only a slight potential benefit to the loss of applicants in that those remaining in the pool were slightly higher on Agreeableness and Conscientiousness, and lower on Psychopathy. However, the effect sizes were small, and those remaining in the applicant pool scored higher on Impression Management, suggesting that their more favorable personality scores may have been dissimulated. I recommend that hiring managers refrain from using SNSs for selection until more validation research can be conducted on SNS information.

Keywords

Social Networking, Social Media, Selection, Personality, Discrimination

Co-Authorship Statement

Richard Goffin

Advisor and professor in Industrial/Organizational Psychology. Second author of the manuscript in Chapter 2 and the publication in Chapter 3.

Kabir Daljeet

Graduate student in Industrial/Organizational Psychology. Third author of the publication in Chapter 3.

Acknowledgments

First and foremost I would like to thank my loved ones for their constant support throughout this process and throughout my entire life. My parents, my brother, and my sister were always there for me, particularly during those times when I was beginning to lose hope and needed their encouragement. My friends, especially those who served as willing expert raters, went above and beyond. In particular, I'd like to thank Matthew McLarnon for his ever sage advice, and for always being willing to help a friend in need. Thank you to my fiancée for her calming nature and her ability to put up with me working a full time job while trying to complete a doctoral dissertation. Kabir Daljeet, thank you for your passion in all things related to Industrial/Organizational Psychology. In addition to your contributions to this work, your enthusiasm reinvigorated my excitement and drive to continue the pursuit of this topic. Finally, I would like to thank my advisor, Richard Goffin, for taking me on as a graduate student and investing considerable amounts of time to develop my skills and abilities. The patience and understanding that he demonstrated throughout this process are just two of the many qualities that make him a great mentor.

Table of Contents

List of Tables

List of Figures

List of Appendices

Chapter 1

1 General Overview

Since the advent of the first social networking site (SNS) in the early 2000's, social media has had a major influence on how individuals choose to interact with others. SNSs offer a unique way of communicating among users. In fact, SNSs are more similar to non-digital human communication than they are to text-based computer-mediated communication (e.g., instant messaging, e-mail), which was once the sole means of communication on the internet (Underwood, Kerline, & Farrington-Flint, 2011). SNSs allow users to create public or semi-public profiles in which others can view their information, such as interests, hobbies, location updates, photos, etc. Lists of users with whom a connection is shared (e.g., friends, co-workers) are created, and users can browse that list and add these users to their own list of connections (e.g., mutual friends; Boyd & Ellison, 2007). Often, a SNS will have a profile that consists of photos, video clips, and biographical/background information created by the user. Most sites also include sections where users can send public or private messages, submit status updates (e.g., what the user is currently doing), and control their privacy settings (e.g., who has access to certain information; Brown & Vaughn, 2011). There is a substantial degree of variation in how this information is created, maintained, and utilized by users.

As the popularity of SNSs has continued to increase, organizations have been taking advantage of the opportunity to make use of SNSs as human resource tools. For instance, the Employers Resource Council (ERC) surveyed organizations in Northeast Ohio to see how SNSs were being used (ERC Survey Briefing, 2009). Forty-nine percent of organizations used SNSs for networking and relationship building, 35% for branding/marketing, 35% for external communication, 31% for reaching new customers, 27% for recruiting, 22% for sales,

et cetera. These percentages indicate that organizations have been interested in making use of SNSs for multiple purposes. However, the authors did not specifically mention one important purpose of SNSs that has recently gained popularity: personnel selection (Grasz, 2009).

1.1 History of SNSs in Selection

The use of the internet for selection did not start with SNSs. In fact, using internet search engines to screen potential job hires has been common practice in the recent past (Karl, Peluchette, & Schlaegel, 2010; Taylor, 2006). ExecuNet's survey of 100 executive recruiters in the United States found that 77% of them used search engines to screen job applicants (Forster, 2006; Jones, 2006). A survey by CareerBuilder.com found that 26% of 1,150 hiring managers used search engines for that same purpose (Sullivan, 2006). Given that many human resource managers were already using the internet in the selection of job candidates, they likely saw the use of SNSs as a natural extension of this screening process. Zeidner (2007) reported in a U.S. survey by the Society for Human Resource Management (SHRM) that 15% of human resource professionals checked SNSs for screening purposes, and 40% of those who did not were likely to check them in the near future. In 2009, that number increased, with 45% of 2,600 hiring managers listed as using SNSs to research job applicants (Grasz, 2009). More recent findings indicate that up to 65% to 75% of organizations and recruiters are using SNSs to screen job applicants (Levinson, 2011; Preston, 2011). This emerging trend in the use of SNSs for selection is not unique to the United States, as similar findings have been reported in Germany (Beeger, 2007), the United Kingdom (Peacock, 2008), and here in Canada (Cerasaro, 2008).

Despite this interest in using SNSs for selection, very little peer-reviewed research has been conducted on the topic. Brown and Vaughn (2011) observed that, at the time, no peer-reviewed research had been conducted on the validity or practical/legal consequences of

using SNS in selection. When it comes to developing new selection instruments, the establishment of validity evidence is essential (Society for Industrial and Organizational Psychology [SIOP], 2003). According to the Principles for the Validation and Use of Personnel Selection Procedures (SIOP, 2003), evidence must be gathered in order to support inferences of the instrument's job-relatedness.

Fortunately, there has been some previous research which should help set the stage for work on the validity of SNSs in selection. Researchers have attempted to interpret information on individual differences that is contained within SNSs. Their studies can trace their origins back to the work of Gosling, Ko, Mannarelli, and Morris (2002), who applied the Brunswik (1956) lens model to infer personality from one's physical environments. Gosling et al. proposed two ways that personality can be interpreted from physical environments: identity claims and behavioural residue. Identity claims refer to "symbolic statements made by individuals about how they would like to be regarded" (Gosling et al., p. 124). For example, displaying a photograph of one's family in an office may be done to communicate the importance one places on family values. Behavioural residue refers to unintentional physical traces of one's behaviour. For example, having papers strewn about an office may be indicative of a low degree of tidying behaviours. Relying on this type of information from personal physical environments, Gosling et al. were able to demonstrate that raters' impressions of personality converged with ratees' self-rated personality. After combining self- and peer-ratings of personality, and then correlating them with observer ratings of offices, a mean correlation of .22 across all Five Factor Model (FFM) personality dimensions was obtained. For bedrooms, a mean correlation of .37 was found. In summary, they were able to provide evidence that everyday cues can be used to predict one's personality.

Moving beyond physical environments to the internet, Vazire and Gosling (2004), and Marcus, Machilek, and Schutz (2006) also applied the Brunswik lens model to personal websites. In both studies, website owners rated their own personality, and undergraduate students rated the owners' personality based upon their viewings of the websites. Vazire and Gosling reported an average correlation of .37, $p < .01$, between the five self-reported and observer-rated FFM personality traits for website owners. Also, an average correlation of .34, $p < .01$, between the five self-reported and observer-rated FFM personality traits was found in the Marcus et al. study. Both studies supported the findings that personality inferences can accurately be gleaned from personal websites. More recently, research has started to investigate whether personality can be inferred based on SNSs (Back et al., 2010; Davison, Maraist, & Bing, 2011; Kluemper & Rosen, 2009; Kluemper, Rosen, & Mossholder, 2012; Krämer & Winter, 2008; Ross et al., 2009).

Two focal research articles have explored whether SNSs can be used to infer personality, and also used to infer attributes assessed by other well-established, work-related selection tools. Kluemper and Rosen (2009) investigated whether SNS information is related to an individual's personality and general mental ability (GMA). They asked 63 participants to rate Facebook profiles. Participants had been trained in personality and general mental ability testing, as well as how to use rating forms as part of an employee selection course. They were asked to spend 10 minutes rating six SNS profiles, to "consider multiple aspects of the profiles which could relate to a specific trait, then complete the rating form based on their overall impression of the social networking profile" (p. 572-572). The intraclass correlation coefficient (ICC) for all ratings ranged from .93 to .99, providing evidence of adequate interrater agreement. The 'true scores' of the six ratees consisted of their self-reported personality ratings (Conscientiousness, Emotional Stability, Agreeableness,

Extroversion, Openness to Experience), their general mental ability scores, and their Grade Point Average (GPA). The inclusion of GPA, a measure of academic performance, served as a proxy for job performance. It is worth pointing out, however, that GPA might be a weak proxy of job performance. The strongest meta-analytic coefficient between the two has only been .35, but that was after correction for measurement reliability in the predictor and the criterion, and correction for range restriction in the predictor (observed $r = .16$; Roth, BeVier, Switzer, & Schippman, 1996). Therefore, the following findings should be interpreted with this limitation in mind. Based on their evaluations of the ratees' SNS profiles, raters could significantly distinguish between the ratee with the highest true score and the ratee with the lowest true score for Conscientiousness ($t = 2.77$, $p < .05$), Emotional Stability ($t = 3.00$, $p < .05$), Agreeableness ($t = 6.12$, $p < .05$), Extroversion ($t = 5.99$, $p < .05$), general mental ability ($t = 13.53$, $p < .05$), and GPA ($t = 2.78$, $p < .05$). However, raters could not correctly distinguish between true scores for Openness to Experience, as they assigned higher ratings to those with low Openness to Experience true scores and lower ratings to those with high Openness to Experience true scores ($t = 6.34$, $p < .05$). Again, these findings may be limited by the fact that the authors dichotomized the true scores (see Butts & Ng, 2009, for an explanation of the problems with dichotomization). A more granular approach that takes full advantage of the continuous nature of the variables would have been preferable. Upon reviewing their findings, it is interesting to note that raters were most accurate at rating GMA. It may be that there are more informational cues within SNSs that are indicative of general mental ability than there are of the other attributes. Raters may have been more accurate in their GMA ratings because they had more informational cues at their disposal. Overall, the authors supported that cues ascertained from SNSs may be used to predict (academic) performance and general mental ability, and not just personality.

In another research article, Kluemper et al. (2012) focused on personality traits assessed through SNS profiles, their relation to self-rated personality, and their relation to job performance. In their first study, three trained raters viewed the Facebook profiles of 586 undergraduate students in order to rate their personality on the FFM. These observer-rated personality scores were significantly correlated ($p < .05$) with measures of self-rated personality for each FFM trait (Conscientiousness, $r = .30$; Emotional Stability, $r = .23$; Agreeableness, $r = .40$; Extraversion, $r = .44$; Openness to Experience, $r = .42$). Thus, through viewing someone's Facebook profile, trained raters may be able to rate one's personality in such a way that it may overlap with that person's self-ratings of personality. Observer-rated personality may be advantageous given that self-ratings may be more susceptible to deliberate distortion in more motivated circumstances (i.e., when applying for a job; Connolly, Kavanagh, & Viswesvaran, 2007). To assess criterion-related validity, Kluemper et al. also correlated observer- and self-rated personality with supervisor-rated job performance. Observer-rated Emotional Stability ($r = .27$, $p < .05$), observer-rated agreeableness ($r = .31$, $p < .05$), and self-rated extraversion ($r = .28$, $p < .05$) correlated significantly with performance. The authors did not speculate as to why these particular traits correlated significantly with performance while others did not. Participants' jobs covered a variety of professions (e.g., clerical, customer service, sales), which could explain the pattern of relations reported in the findings. It is surprising that both observer- and self-rated Conscientiousness were not significantly correlated with performance, considering how Conscientiousness is typically applicable to the majority of jobs.

In the second part of Kluemper et al. (2012), a sample of 782 undergraduate students was used. Using the same procedure as Study 1, observer- and self-rated personality were significantly correlated ($p < .05$) with each other for each FFM trait (Conscientiousness, $r =$

.19; Emotional Stability, r = .21; Agreeableness, r = .26; Extraversion, r = .28; Openness to Experience, r = .16). To explore criterion-related validity in this study, Kluemper et al. used academic performance, as measured by cumulative GPA. Observer-rated Conscientiousness (r = .27, p < .05), Emotional Stability (r = .21, p < .05), Agreeableness (r = .19, p < .05), and Openness to Experience (r = .28, p < .05) correlated significantly with academic performance. Also, self-rated Conscientiousness (r = .24, p < .05), Agreeableness (r = .12, p < .05), and Openness to Experience (r = .15, p < .05) correlated significantly with academic performance.

The significant correlations between observer- and self-rated personality in Study 1 and Study 2 in Kluemper et al. (2012) provided supporting evidence of the convergent validity of observer-rated personality via SNSs. Also, the significant correlations between observer-rated personality and performance in both studies offered some supporting evidence of criterion-related validity. In combination, the findings from Kluemper and Rosen (2009) and both studies in Kluemper et al. (2012) suggested that cues obtained from SNSs may have the potential to be valid predictors of job performance. However, a number of the correlations in these studies had only small effect sizes (see Cohen, 1988). Therefore, the current dissertation sought to investigate the validity of SNS information used for selection and explore a new way of rating SNS information by focusing on specific SNS information.

1.2 Validity and Practical Realities

Although there is some potential for using SNS information for selection, there is growing concern amongst researchers and practitioners about its validity (Brown & Vaughn, 2011; Grasz, 2009). In particular, they are concerned that SNS information that is not job-related may be used for selection. In addition, there is growing legal concerns that employers may be breaching employment discrimination laws by accessing protected applicant

characteristics (e.g., age, race, gender; SHRM, 2011). For instance, in order for a selection test to be legally defensible, it must not demonstrate adverse impact against protected groups. Adverse impact is when members of one group are selected at a greater rate than members of another group (e.g., members of a protected group). If adverse impact is found, more work is required by the test publisher or organization to ensure that that the test is not leading to discrimination (for a full list of the work required, see Cascio & Agunis, 2005, p. 185). If using a test that lacks legal defensibility when making hiring decisions, an employer could be subject to human rights tribunals or other legal action.

The current study seeks to investigate some of these concerns by exploring the validity of SNS information used for selection and introducing a new method of rating SNS information. It also seeks to address these concerns by better understanding the practical realities that SNS selection may have on the workforce.

1.3 The Current Dissertation

The first major goal of the current dissertation was investigated in Chapter 2. The goal was to evaluate whether specific information from SNSs has the potential to be used as a valid source of information for personnel selection decisions; to better understand what information SNSs convey (i.e., what construct domains are being tapped), and whether this information can be considered job-relevant.

The second major goal of the dissertation was investigated in Chapter 3. This goal was to go beyond the validity investigation by investigating the practical realities and the effects SNS selection might have on the workforce. For instance, if applicants were asked to give up their SNS password in order to be considered for a job, what percentage would

remove themselves from the selection process? What effect would this have on the remaining applicant pool, and could such a demand lead to discrimination?

1.4 References

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., et al. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science, 21*(3), 372.

Beeger, B. (2007, September 3). Soziale Netzwerke: Karrierekiller im Internet. Retrieved from http://www.stern.de/computer-technik/internet/:Soziale-Netzwerke-Karrierekiller-Internet/596742.html

Boyd, D.M., & Ellison, N.B. (2007). Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, *13*, 210–230.

Brown, V. R., & Vaughn, E. D. (2011). The writing on the (facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology, 26,* 219-225.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Butts, M.M., and Ng, T.W.H. (2009). Chopped liver? OK. chopped data? Not OK. In C.E. Lance and R.J. Vandenberg (Eds.), *Statistical and Methodological Myths and Urban Legends: Doctrine, Verity, and Fable in the Organizational and Social Sciences* (361-386). New York: Taylor & Francis Group.

Cascio, W.F. & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Catano, V.M., Wiesner, W.H., Hackett, R.D., & Methot, L.L. (2010). *Recruitment and Selection in Canada* (4th Ed.). Toronto, ON: Thomson Nelson.

Cerasaro, A. (2008, December 29). Employers defy privacy by using Facebook. *Tennessee Journalist*. Retrieved from http://tnjn.com/2008/apr/08/employers-defy-privacy-by-usin/

Connolly, J. J., Kavanagh, E. J., & Viswesvaran, C. (2007). The convergent validity between self and observer ratings of personality: A meta-analytic review. *International Journal of Selection and Assessment, 15*, 111-117.

Davison, H. K., Maraist, C., & Bing, M. N. (2011). Friend or foe? the promise and pitfalls of using social networking sites for HR decisions. *Journal of Business and Psychology, 26*, 1-7.

ERC Survey Briefing. (2009). *Social Media in the Workplace*. Mayfield Village, OH.

Forster, J. (2006, May 1). Job recruiters dig up dirt on applicants' web pages. *Akron Beacon Journal*, p. 2.

Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*(3), 379.

Grasz, J. (2009, August 19). Forty-five percent of employers use social networking sites to research job candidates, career builder survey finds. *CareerBuilder.com.* Retrieved from http://www.careerbuilder.com/Article/CB-1337-Interview-Tips-More-Employers-Screening-Candidates-via-Social-Networking-Sites/

Jones, K. C. (2006, July 11). 'Digital dirt' derailing job seekers. *InformationWeek.* Retrieved from http://www.informationweek.com/story/showArticle.jhtml?articleID=190302836

Karl, K., Peluchette, J., & Schlaegel, C. (2010). Who's posting Facebook faux pas? A Cross-Cultural examination of personality differences. *International Journal of Selection and Assessment, 18*(2), 174-186.

Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: Evaluating social networking web sites. *Journal of Managerial Psychology*, *24*(6), 567-580.

Kluemper, D. H., Rosen, P. A., & Mossholder, K. W. (2012). Social networking websites, personality ratings, and the organizational context: More than meets the eye? *Journal of Applied Social Psychology.* doi: 10.1111/j.1559-1816.2011.00881.x

Levinson, M. (2011). Social networks: New hotbed for hiring discrimination claims. *Computer World*. Retrieved from http://www.computerworld.com/s/article/9215907/Social_Networks_New_Hotbed_for_Hiring_Discrimination_Claims.

Marcus, B., Machilek, F., & Schütz, A. (2006). Personality in cyberspace: Personal web sites as media for personality expressions and impressions. *Journal of Personality and Social Psychology, 90*(6), 1014.

Peacock, L. (2008, April 4). Employers watch Facebook usage. *Employers Law*.

Preston, J. (2011). Social media history becomes a new job hurdle. *New York Times*.

Ross, C., Orr, E. S., Sisic, M., Arsenault, J. M., Simmering, M. G., & Orr, R. R. (2009). Personality and motivations associated with Facebook use. *Computers in Human Behavior, 25*, 578-586.

Roth, P. L., BeVier, C. A., Switzer, F. S. III, & Schippmann, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology, 81*, 548-556.

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection* (4th ed.) Bowling Green, OH: Author (Available on-line)

Society for Human Resource Management (2011, October). *SHRM Research Spotlight: Social Networking Websites and Online Search Engines for Screening Job Candidates.* Retrieved from http://www.shrm.org/Research/SurveyFindings/Articles/Documents/PART%202_Social%20Media%20Flier_FINAL.pdf

Sullivan, S. (2006, October 26). One-in-four hiring managers have used Internet search engines to screen job candidates; one-in-ten have used social networking sites, CareerBuilder.com survey finds. Retrieved from http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?id=pr331&ed=12%2F31%2F2006&sd=10%2F26%2F2006&cbRecursionCnt=1&cbsid=a5015667d80f4b599c46d2b08f406b67-241548812-RI-4&ns_siteid=ns_us_g_One%2din%2dFour_Hirin_

Taylor, S. (2006). Seeking secrets in cyberspace. *Staffing Management, 2*. Retrieved from http://www.shrm.org/ema/sm/articles/2006/julysept06cover.asp

Underwood, J. D. M., Kerlin, L., & Farrington-Flint, L. (2011). The lies we tell and what they say about us: Using behavioural characteristics to explain Facebook activity. *Computers in Human Behavior*, *27*, 1621-1626.

Vazire, S., & Gosling, S. D. (2004). E-perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, *87*, 123–132.

Zeidner, R. (2007). How deep can you probe? *HRMagazine*, *52*, 57–62.

Chapter 2

2 Social Media in Selection: A Construct Validity Approach

In this chapter, I will investigate whether SNSs can potentially be used as valid sources of information for personnel selection purposes. The goal is to better understand what construct domains are being tapped by SNS information, and whether this information is job-relevant. This chapter is a manuscript that has yet to be submitted for publication.

2.1 Abstract

With the increasing popularity of social networking sites (SNSs), hiring managers are looking to SNSs such as Facebook to aid with their hiring decisions (Levinson, 2011). However, previous research has focused on a Global SNS Rating of applicants' SNS profiles that may not be construct valid. Thus, the current study created a set of Specific SNS Indicators, which consist of specific informational cues ascertained from SNS profiles, and correlated them, along with a Global SNS Rating, with scales of traditional selection constructs in order to test their construct validity. Findings from the current study of 141 undergraduate students at a large Canadian university demonstrated that Specific SNS Indicators had fairly weak reliability evidence, but had some evidence of structural validity and of construct validity. Only Extraversion, Openness to Experience, and Verbal Ability Specific SNS Indicator scales were related to traditional selection scales, and the effect sizes were only small to moderate. Extraversion and Verbal Ability Specific SNS Indicator scales offered incremental prediction beyond a Global SNS Rating in the prediction of the traditional selection scales. This additional incremental prediction could be of value when considering the construct validity of SNSs in selection contexts in the future. Specific SNS Indicators evidenced discriminant validity with respect to scales of response distortion. Given

the fairly limited reliability and validity evidence, it is recommended that practitioners avoid the use of SNS information for selection unless future research supports its use.

Keywords

Social Networking Sites, Social Media, Selection, Individual Differences, Personality, General Mental Ability, Integrity, Discrimination

2.2 Introduction

Social Networking Sites (SNSs) such as Facebook, LinkedIn, and Twitter have become enormously popular, even to the point of changing social interaction (Subrahmanyam, Reich, Waechter, & Espinoza, 2008**)**. As a result of this popularity, hiring managers have started considering their use for assessment purposes, with some research showing that up to 65% of companies are using SNSs to screen job applicants (Levinson, 2011). In fact, the outreach manager for the Equal Employment Opportunity Commission's office in Houston has suggested that 75% of recruiters have researched their applicants online, with 70% of them reporting that they have rejected a candidate as a result (Preston, 2011).

The use of SNSs for selection may lead to major benefits, which may explain their popularization as a selection tool. For example, they incur relatively little cost to the organization. Also, there is an abundance of information about job candidates available on SNSs that is easily accessible. This allows even small organizations the opportunity to attempt to verify information contained in résumés, reference letters, interviews, and other selection information sources typically drawn from applicants. Previously, advanced background checks were unavailable to small organizations due to their substantial costs.

Now, those making hiring decisions may have access to "detailed information that would allow them to draw conclusions or make inferences about the applicant's character or personality that might not be as easily or economically obtained through traditional means" (Brown & Vaughn, 2011, p. 220). For example, interviews are expensive to conduct. Instead of (or in addition to) asking applicants about their hobbies and interests in an interview, an employer can easily find their hobbies and interests listed on their SNS profile pages. Also, because there is a near-unlimited amount of space to describe this information on SNSs, more of this information may be available than in resumes, curricula vitae, or application forms, which may be limited by space restrictions.

Despite the potential for valid prediction of performance, a major concern among researchers and hiring managers is that SNS information that is not job-related may be used for selection (Brown & Vaughn, 2011; Grasz, 2009; SHRM, 2011). As a result, recently, the percentage of organizations using SNS information for selection has dropped slightly (SHRM, 2011). Human resource professionals reported an increase in their legal concerns about discovering information about protected applicant characteristics (e.g., age, race, gender, religious affiliation; SHRM, 2011). Employers may purposely or accidentally be accessing protected group status information to which they may not otherwise have access. For example, a job applicant's national origin or religion may be easily ascertained or inferred from his or her SNS profile. If this information were to affect an employer's rating of the applicant's SNS profile and affect the selection decision, this could be a breach of employment discrimination laws (for information on the Canadian Human Rights Act and provincial laws, see Catano, Wiesner, Hackett, & Methot, 2010; for information on employment discrimination laws in the U.S., visit http://www.eeoc.gov/). Unfortunately, these concerns have not been assuaged by the selection literature. With a few notable

exceptions (Bohnert & Ross, 2010; Kluemper & Rosen, 2009; Kluemper, Rosen, & Mossholder, 2012; Van Iddekinge, Lanivich, Roth, & Junco, 2013), there has been very little primary research on the viability of SNSs in selection. In fact, Brown and Vaughn (2011) noted in a review of the literature at the time that no peer-reviewed research had been done on the practical and legal consequences of using SNSs, or even on the validity of using SNSs in selection. In an attempt to address the lack of validity studies, the current study provides a preliminary investigation of the construct validity of using SNSs to assess job-relevant information, and introduces a new way of assessing this information. In addition, hiring managers likely have focused on using SNSs as a screening tool to remove the "bad apples". Hiring managers may be searching for illegal activities on SNSs to avoid embarrassment or the endangerment of employees or customers (Slovensky & Ross, 2012). The goal of the present work was to consider SNSs from a personnel selection "best practices" perspective. Rather than focusing solely on screening out bad applicants, the focus was on selecting the best applicant, which is characteristic of most traditional selection systems.

2.2.1 A New Way to Assess SNS Information

Researchers (e.g., Kluemper & Rosen, 2009; Kluemper et al., 2012), and likely hiring managers as well, have focused primarily on a Global SNS Rating. A Global SNS Rating involves a simple evaluation process whereby all informational cues are considered simultaneously (Thompson, Iacovou, & Shirland, 2006). By contrast, specific informational cues consist of those individual behaviours/observations referred to in a SNS profile (e.g., how many racist jokes did the SNS user post on his/her profile). Brown and Vaughn (2011) suggested that specific indicators of profile information might be helpful in systematically organizing the information collected from SNSs (Brown & Vaughn, 2011). Systematically organizing this information according to the attributes that are being inferred could lead to

less biased information being collected from individuals, and may help document this information in case of future litigation (Brown & Vaughn). Also, the specific indicators may improve reliability and validity over overall judgments, as the consistency of overall judgments has been found to decrease as the number of informational cues increases (Cook & Stewart, 1975; Keeley & Doherty, 1972). This follows from research by Heneman (1986) in the performance appraisal literature which specifies that simplifying the information processing required by raters and providing clear standards increases the validity of ratings.

For the current study, I refer to these specific informational cues as Specific SNS Indicators, which I define as specific cues ascertained from SNS profiles that are believed to be conceptually linked to job-relevant human attributes. These Specific SNS Indicators may be evaluated by raters (e.g., the careful articulation of thoughts/ideas in the SNS user's posts) or may simply be tabulated (e.g., how many friends are on his/her social network). The indicators may be multidimensional, and may tap more than one trait or attribute.

As in Brown and Vaughn (2011), I theorized that specific indicators of profile information would be helpful in systematically organizing the information collected from SNSs. However, a Global SNS Rating was also included to allow empirical comparison of the validity of both the Global SNS Rating and the Specific SNS Indicators; to help explore whether specific indicators are more helpful in systematically organizing SNS information than what is currently offered by a Global SNS Rating.

In addition, I expanded on Brown and Vaughn's (2011) theory by examining the different ways in which this specific information can be organized, and seeing which way has more potential when it comes to SNS information. There are a number of ways that data can be collected and combined, and this topic has been hotly debated in the clinical psychology

literature for over 80 years (Gough, 1962). The debate was prompted largely due to the work of Meehl (1954), when he explored the implications of clinical (judgmental) versus statistical prediction. For instance, one way of collecting data is to use *judgmental data collection methods*. For judgmental data collection methods, a rater evaluates the different respondents. As an example, a clinician would interview an individual in order to get a better sense of who they are. Another way of collecting data is through *mechanical data collection methods*, whereby prespecified rules are required and subjective judgment is not involved. This can also be referred to as statistical prediction. As an example, an individual could take a general mental ability test to better understand his/her level of intelligence. As part of the test, there are prespecified rules as to what is a correct and an incorrect response. While data can be collected using either judgmental or mechanical methods, they can also be combined using methods that are judgmental or mechanical. *Judgmental combination methods* involve one or more decision makers "developing a composite picture and behavioural prediction of a candidate" (Cascio & Agunis, 2005, p. 315). When judgmental data collection methods are combined with judgmental combination methods, this is referred to as the *pure clinical strategy*. *Mechanical combination methods* involve statistically assigning and combining scores in order to form scales. Researchers have found that the best configuration is to use both judgmental and mechanical data collection methods, and then use mechanical combination methods (Cascio & Aguinis, 2005; Einhorn, 1972; Sawyer, 1966). This is referred to as the *mechanical composite strategy*. There are several plausible reasons for why the mechanical composite strategy is superior to the other strategies, including the pure clinical strategy, in the prediction of candidates' behaviours. It can lead to more appropriate weighting of predictors, leading to greater accuracy in prediction (Bass & Barrett, 1981; Hitt & Barr, 1989). According to Cascio & Aguinis (2005), the appropriate weighting of

predictors is incredibly difficult to judge accurately. The mechanical composite strategy also allows for additional evidence on candidates to be incorporated, which should improve the accuracy of predicting candidates' behaviours (Bass & Barrett, 1981; Hitt & Barr, 1989). Raters relying only on a pure clinical strategy, however, may be unable to incorporate additional evidence because they have reached the limits of their cognitive capacity (Cascio & Aguinis, 2005).

For the current study, judgmental data collection methods as well as mechanical data collection methods were used for the Specific SNS Indicators. As part of the judgmental data collection methods, a relative format was used for some of the Specific SNS Indicators. A relative format is a judgment or rating of an individual with other individuals used as reference points (Goffin & Olson, 2011). The relative format compares the participant to the average Facebook user, which allows for comparative judgment, but the participant is also assigned a percentile for that behaviour, which provides a numerical score. A relative format, specifically, the Relative Percentile Method (RPM; Goffin, Gellatly, Paunonen, Jackson, & Meyer, 1996), was used because research has suggested that the RPM formats may result in incremental predictiveness beyond conventional rating formats (Goffin et al., 1996; Goffin, Jelley, Powell, & Johnston, 2009; Goffin & Olson, 2011). Also, where appropriate, some Specific SNS Indicators consisted of a count format as part of mechanical data collection methods. As an example, a count format item could consist of frequency information such as "number of written posts or comments by the candidate that refer to alcohol or drug use". Using the RPM format as a judgmental data collection method and the count format as a mechanical data collection method, the data were combined using mechanical data combination methods. Again, this was referred to as the mechanical composite strategy. The data were combined mechanically by statistically assigning and combining direct

observations in a structural equation modeling program in order to form the latent variables of the Specific SNS Indicators believed to assess work-related attributes. Additionally, a pure clinical strategy was also employed in the form of the Global SNS Rating. The Global SNS Rating relies on judgmental data collection methods because it is also collected through a RPM rating. It also relies on judgmental data combination methods because the rater has to summarize his or her impressions of the candidate. A major theoretical contribution of the current study was to investigate whether this mechanical composite strategy accounts for better prediction of candidate attributes than does the pure clinical strategy. Essentially, do the Specific SNS Indicators lead to better prediction of candidate attributes than does the Global SNS Rating?

In addition to the theoretical considerations above, there were also potential practical benefits to the current study's focus on Specific SNS Indicators. Using Specific SNS Indicators should help clarify exactly what is being measured, rather than relying on an overall rater judgment of what a rater implicitly deems important when evaluating an applicant's SNS. This clarification is important to help establish content validity, making it easier for hiring managers to demonstrate the job-relatedness of their SNS selection indicators (Goldstein, Zedeck, & Schneider, 1993). Managers will be able to specify which Specific SNS Indicators are related to the different knowledge, skills, abilities, and other characteristics (KSAOs) necessary for specific job tasks in future job analyses. Additionally, managers could benefit from the aforementioned potential to avoid legal ramifications and discrimination through the systematic documentation of the job-relatedness of SNS information that is used to assist in hiring decisions.

2.2.2 Reliability

Before Specific SNS Indicators can be considered for use as a selection tool, they must have the potential to be reliably coded. However, because the current study takes a preliminary look at the potential of Specific SNS Indicators, and different types of Specific SNS Indicators, I did not want to reject items that had the potential to become more reliable items once rewritten/refined as part of future investigations. Thus, I was concerned that the criteria for retaining Specific SNS Indicators for further study not be overly stringent. Therefore, a statistically significant interrater reliability (above .24 in the current study) was considered the minimum cutoff for retaining an item, because nonsignificant interrater reliabilities suggest that the judgments of different raters could be completely unrelated (Gourraud, Le Gall, Puzenat, Aubin, Ortonne, & Paul, 2012).

To measure interrater reliability, intraclass correlation coefficients (ICCs) were used. ICC(2,k) was the focus for the current study. The "2" signifies that the raters were chosen from a random sample of raters, and that there is a generalization to a larger population of raters (LeBreton & Senter, 2007; Shrout & Fleiss, 1979). The "k" signifies that there was more than one rater making the ratings (LeBreton & Senter, 2007; Shrout & Fleiss, 1979). Currently, the ICC(2,k) reflected the interrater reliability if the ratings given by the three SNS Expert Raters were averaged together. Accordingly, the following hypothesis is put forth:

H_1. For the Specific SNS Indicators rated by more than one rater, the interrater reliabilities [ICC(2,k)] will be significantly different from zero.

2.2.3 Structural Validity

The assessment of the SNS instrument's structural validity is an important validation step. Structural validity is defined as the degree of fit between the empirical factor structure

of an instrument and the theoretical organization of the construct that is supposed to be measured by the instrument (Loevinger, 1967). For more on the assessment of structural validity, see Goffin and Jackson (1988). Based on the above definition of structural validity, the following hypothesis is put forth:

H_2: The empirical dimensionality of the Specific SNS Indicators will reflect the theoretical dimensionality as informed by the judgment of expert raters.

2.2.4 Construct Validity

As part of the validation process, a list of major human attributes that have been repeatedly linked to job performance in the personnel selection literature was compiled. These attributes are general mental ability, personality, and integrity.

General mental ability (GMA), also known as general cognitive ability or general intelligence, has been studied for more than 100 years (Grubb, Whetzel, & McDaniel, 2004). It is one of the most widely-cited personnel selection measures, and is deemed the best single predictor of job performance (Murphy, 2002; Schmidt & Hunter, 1998).

Personality has played an important role in selection, and research has demonstrated that it is a useful predictor of job performance (Barrick & Mount, 1991; Hough, Eaton, Dunnette, Kamp, & McCloy, 1990; Ones & Viswesvaran, 1998; Ones, Viswesvaran, & Schmidt, 1993; Rothstein & Goffin, 2006; Salgado, 1997). This research has primarily focused on the FFM (e.g., Conscientiousness, Agreeableness, Openness to Experience, Neuroticism, and Extraversion) of personality. However, Honesty-Humility is another personality variable that has been found to be an important predictor of work performance

measures such as counterproductive work behaviour (Lee, Ashton, & Shin, 2005; Marcus, Lee, & Ashton, 2007; Schneider & Goffin, 2012).

Finally, integrity is an important attribute that has been tied to job performance (Cullen & Sackett, 2004). Integrity tests have often been used to predict criteria such as theft, absence, turnover, job performance, and counterproductive work behaviours (Cullen & Sackett). Integrity tests are thought to encompass characteristics such as trustworthiness, honesty, integrity, conscientiousness, reliability, and dependability (Sackett & Wanek, 1997).

Specific SNS Indicators were created (see the measures section), and experts rated how well they believed each would tap the targeted attributes. These experts will be referred to as Taxonomy Expert Raters throughout this paper. To assess the validity of the Specific SNS Indicators, I examined the correlations of the Specific SNS Indicators with the traditional, well-established scales of the same attributes. This should result in a better understanding of what construct domains the Specific SNS Indicators are tapping. For example, a potential Specific SNS Indicator would be the degree to which an individual posts content about himself/herself using drugs. A higher degree of self-posts with drug content could be an indicator of drug abuse, a behaviour that is subsumed within the attribute of integrity (Cullen & Sackett, 2004). Therefore, if this Specific SNS Indicator were at least weakly correlated with integrity test scores, this might be an indication that the Specific SNS Indicator may be tapping the construct of integrity.

Venting about personal frustrations is another example of a potential Specific SNS Indicator. Ratings of this indicator could be indicative of high Neuroticism, an attribute that is characterized by anxiety, hostility, depression, self-consciousness, impulsiveness, and vulnerability to stress. If ratings of the degree of venting on SNSs were at least weakly

correlated with scores on Neuroticism, then that would demonstrate evidence that the "venting" Specific SNS Indicator may be assessing what it is believed to assess.

I predicted that Specific SNS Indicators that are believed to tap performance-related human attributes, such as those listed in the examples above, should have at least weak, significant correlations with corresponding, established, traditional scales that assess these same attributes. If Specific SNS Indicators are at least significantly, albeit weakly, related to traditional scales, they could be a valuable alternative source of information for a given attribute. According to Cohen (1988) correlations between .10 and .30 are weak, between .30 and .50 are moderate, and greater than .50 are strong. Thus, the following general hypothesis is posited:

H_{3a}: Specific SNS Indicators rated as tapping performance-related human attributes will have significant correlations with traditional selection scales that assess these same attributes, and these correlations will be at least weak ($r > .10$) in magnitude.

Overall judgments rely on what a rater implicitly deems important and they have a large breadth of coverage, whereas Specific SNS Indicators have a more specific focus and narrower content domain. Following from the personality literature on broad versus narrow traits (see Ashton, 1998; Hastings & O'Neill, 2009; Rothstein & Goffin, 2006; Schneider, Hough, & Dunnette, 1996), the Specific SNS Indicators, with their narrower focus, may add unique variance beyond a Global SNS Rating in the prediction of traditional selection scales. Therefore, I predict:

H_{3b}: Specific SNS Indicators will account for significant incremental prediction beyond a Global SNS Rating in the prediction of traditional selection scales.

2.2.5 Testing for Response Distortion

There is a concern that has been debated among some researchers that SNS accounts may not reflect their creators' true identities and characteristics. Some researchers (e.g., Davison, Maraist, & Bing, 2011) have suggested that because of the visibility of SNSs, applicants might "fake good" by distorting their SNS profiles in a bid to impress family, friends, potential mates, and/or employers. However, other researchers (e.g., Back et al., 2010, Brown & Vaughn, 2011; Krämer & Winter, 2008; Lampe, Ellison, & Steinfeld, 2007; Pempek, Yermolayeva, & Calvert, 2009) have suggested that faking good is not likely to occur, because SNS friends provide feedback and accountability on one's profile. It is important to test whether applicants are "faking good" on their SNS information, because if they are, there may be implications for the validity findings of the current study. For example, the relation between self-reported Extraversion and a Specific SNS Indicator that is supposed to assess Extraversion could be weaker than expected if users are quite likely to distort their SNS postings. Thus, for the current study, three different scales of faking were used to test whether applicants were "faking good" on their SNS information.

Scales of impression management, one of the two dimensions of Social Desirability, assess the degree to which one engages in deliberate self-presentation to an audience (Paulhus, 1988). Specific SNS Indicators should be distinct from Impression Management. If users are not managing their impressions, then their Specific SNS Indicators should not be significantly related to Impression Management scores. In addition, the relation between Specific SNS Indicator scales and their respective traditional scales should be stronger than the relations between the Specific SNS Indicator scales and Impression Management. Measuring these relations represents a form of discriminant validity based on the Hubley (2014) definition that theoretically distinct constructs should not be highly correlated with

one another, and that these correlations should be smaller in magnitude than correlations between theoretically similar constructs.

H_{4a}. The Specific SNS Indicator scales will be more strongly correlated with their respective traditional scales than with an Impression Management scale.

It is worth pointing out that there have been alternative interpretations of Impression Management scores, with some researchers (Zettler, Hilbig, Moshagen, & de Vries, 2015) suggesting that they may reflect more virtuous, honest behaviour than dishonest responding. With this in mind, two additional measures of response distortion were included to see if a pattern of dishonest responding emerged across measures of response distortion.

Self-Deception was chosen as the second measure of response distortion. Even if SNS users are not consciously distorting their presentation to others, they may be doing so subconsciously. For instance, individuals who engage in self-deception actually believe their own positively biased self-reports (Paulhus, 1988). Although there is a possibility that users may be unwittingly distorting the information in their SNS profiles, previous research has suggested that SNS profiles are fairly accurate in how they depict their owner's personality (Back et al., 2010). Therefore:

H_{4b}. The Specific SNS Indicator scales will be more strongly correlated with their respective traditional scales than with a Self-Deception scale.

Schneider and Goffin (2012) introduced the Perceived Ability to Deceive scale (PATD), a new scale that may be related to faking. Those who consider themselves better at deception may be more likely to actually fake. Therefore, if users have a high PATD score, they may be more likely to fake their SNS profiles. As in Hypothesis 4a and 4b, it was

predicted that users will not distort their impressions on their SNS profiles. Therefore, their Specific SNS Indicator scales should not be strongly correlated with a PATD scale.

H_{4c}. The Specific SNS Indicator scales will be more strongly correlated with their respective traditional scales than with the PATD scale.

Even if H_{4a}, H_{4b}, and H_{4c} are supported, this does not necessarily prove that faking is not occurring. However, if one or more of these hypotheses are not supported, then that would indicate that any findings supporting or disproving validity evidence for the Specific SNS Indicator scales would have to be interpreted with caution because criterion correlations may be inflated or deflated due to variance from Impression Management, Self-Deception, and/or PATD.

2.3 Method

2.3.1 Participants

The sample consisted of 141 undergraduate students from the participant pool at a large Canadian university. Fifty-seven males and 66 females completed the study, with 18 participants missing gender information. Ages ranged from 19 to 32 (M = 20.44, SD = 1.34). When asked how much time they actively spent on SNSs (excluding chat services) in an average day, six spent 0 to 15 minutes, eight spent 15 to 30 minutes, 29 spent 30 minutes to an hour, 41 spent 1 to 2 hours, 31 spent 2 to 4 hours, six spent 4 to 8 hours, and two spent more than 8 hours. Employment data were only available for 48 participants. Forty-three of these participants had been employed before, and five had never held a job before. Six participants were currently employed, and 42 were not.

Participants earned course credit as part of their participation in this online study. To participate, they were required to have a Facebook account and to "friend" the researcher so that their Facebook information could be rated.

2.3.2 Procedure

When participants first arrived at the study website, they were presented a letter of information explaining the nature of the study, and were told they would be required to accept a "friend" request from the researcher in order to participate. Also, they were informed that the researcher would remove himself from their Facebook account (i.e., 'unfriend them') two weeks after their profiles had been rated.

Demographic information (e.g., age, gender) was collected from participants. As mentioned previously, employment history was available for 48 individuals. These employment data were not collected as part of the current survey, however, that information was collected as part of a larger battery, referred to as "mass testing" within the university. Those 48 respondents completed employment history for the larger battery and were matched to the current study through an anonymous ID.

Participants were then asked to voluntarily complete the following attribute scales: GMA, integrity, and personality. Finally, participants were thanked for their participation, and were shown a debriefing form.

The previously mentioned Taxonomy Expert Raters rated how well each Specific SNS Indicator should tap each performance-related attribute. The ratings were an important step towards testing the theoretical dimensionality of the Specific SNS Indicators. The raters consisted of 10 current and former Industrial/Organizational graduate students (four Ph.D. consultants, four Ph.D. students, and two M.Sc. students) who were familiar with personnel

selection as part of their graduate training. See the Specific SNS Indicator Taxonomy Creation subsection of the results for more on these Taxonomy Expert Raters and how they made their ratings.

Three SNS Expert Raters were charged with rating participants' Facebook profiles. These SNS Expert Raters were separate from the 10 Taxonomy Expert Raters. The three SNS Expert Raters consisted of a MSc. graduate student with experience in selection and an extensive knowledge of social media, an undergraduate student with extensive knowledge of social media, and a human resource professional. They "scored" each participants' Facebook pages using the Specific SNS Indicators and using three items to measure their Global SNS Rating of the profiles. The order of the ratings was counterbalanced such that for even numbered participants, a Global SNS Rating was completed first and Specific SNS Indicators completed second. For odd numbered participants, Specific SNS Indicators were completed first and the Global SNS Rating was completed second. I assumed that hiring managers would want to focus on more recent SNS information because they have limited time to screen candidates' Facebook profiles, so only the SNS information from the past two years was evaluated.

2.3.3 Measures

Demographics. Items asking participants' age, gender, race, year of academic study, and work experience were included.

General mental ability. The Personnel Assessment Form (PAF; Jackson, 2004) was used as a measure of GMA. The verbal subtest consists of 35 multiple-choice items that assess accumulated knowledge of diverse topics, words, and verbal concepts, and assesses

the ability to recognize similarities among objects and concepts. The quantitative subtest consists of 21 numerical problems with multiple-choice response options that assess reasoning and problem solving abilities. When summed together, these subtests assess general mental ability. An internal consistency reliability of .89 has been reported for the verbal score, .87 for the quantitative score, and .94 for the general mental ability score (Jackson, 2004). The PAF has demonstrated strong convergent validity (r = .82; Jackson, 2004) with another measure of general mental ability, the Multidimensional Aptitude Battery-II (MAB-II; Jackson, 1998), upon which the PAF is based. It also has strong convergent validity (r = .85; Jackson, 2004) with the Wonderlic Personnel Test (Wonderlic, 2002).

Integrity. The English version of the Inventar Berufsbezogener Einstellungen und Selbsteinschätzungen [job-related attitudes and self-evaluations inventory; IBES; Marcus, 2006] served as the integrity test for the current study. The IBES consists of 60 Overt and 55 Personality-Based Integrity items, and "the length, content, and breadth of measurement of this test corresponds closely to those of many commercial integrity tests available in North America" (Marcus, Lee, & Ashton, 2007, p. 12). Participants reported the extent to which they disagree or agree with each statement using a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). The validity of the IBES has been supported in a number of studies (see Marcus, 2007, for a review).

Personality. A 120-item version of the FFM from the International Personality Item Pool (IPIP; Goldberg, 1999) was used to measure personality. This version is described in Johnson (2001) and Hastings and O'Neill (2009). Conscientiousness, Agreeableness, Extraversion, Neuroticism, and Openness to Experience are each measured with 24 items on

a 5-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). These scales were designed to capture the same dimensions as the NEO Personality Inventory (NEO-PI-R; Costa & McCrae, 1992). Goldberg reported internal consistency reliabilities of .85 to .91 for the IPIP scales, and strong evidence of validity. Also, Johnson reported strong convergent correlations between the IPIP scales and the respective NEO-PI-R scales.

The 16 Honesty-Humility items from the 100-item HEXACO Personality Inventory-Revised (Ashton & Lee, 2008) were included as part of the personality scales. The same 5-point Likert scale (1 = strongly disagree, 5 = strongly agree) was used. Ashton and Lee reported an internal consistency of .92 for Honesty-Humility, along with extensive evidence of validity.

Impression management. The Impression Management scale from Paulhus' (1984, 1988) Balanced Inventory of Desirable Responding (BIDR) was included in the current study. The 20-item scale was embedded within the personality scales described above. Participants stated their agreement or disagreement with the statements on a 5-point Likert scale (1 = strongly disagree, 5 = strongly agree). Paulhus (1991) reported internal consistency reliabilities of .75 to .86 for the Impression Management scale. After five weeks, the test-retest reliability was .65 for the Impression Management scale (Paulhus, 1988). The BIDR has also demonstrated convergent and discriminant validity (Paulhus, 1991).

Self-Deception. Self-Deception was also measured using the BIDR (Paulhus, 1984, 1988). It also consists of 20 items and uses the same 5-point Likert scale as the Impression Management scale. Internal consistency reliabilities of .68 to .80 have been

reported (Paulhus, 1991), as well as a test-retest reliability of .69 after five weeks (Paulhus, 1988) and evidence of validity (Paulhus, 1988; Paulhus 1991).

Perceived Ability to Deceive. The PATD (Schneider & Goffin, 2012) was used, and consisted of six items on a 5-point Likert scale from Strongly Disagree to Strongly Agree. The PATD has demonstrated evidence of reliability (Cronbach's alpha = .81) and discriminant validity (Schneider & Goffin, 2012). It has also been shown to increase incremental prediction in counterproductive work behaviour beyond Conscientiousness, Agreeableness, Neuroticism, and Honesty-Humility (Schneider & Goffin, 2012).

Careless responding checks. At the end of the study, participants responded to a self-reported careless responding item which was used to identify and remove careless responders from the data. This item was based on a careless responding item recommended by Meade and Craig (2012). For the current study, participants responded "Use my responses" or "Do NOT use my responses" to the question:

> "If, for any reason (e.g., not carefully responding to the items) you do not think your responses should be included in the current study, please click 'Do NOT use my responses'. Otherwise, please click 'Use my responses'. Your response will not affect your credit for this study".

Specific SNS Indicators. The Specific SNS Indicators consisted of items rated by the SNS Expert Raters using three different formats: the RPM format, the count format, and the categorical format. Thirty-five items utilized the RPM format (labelled as "0-100" in Item Scaling column of Table 1; see Figure 1 for an example of the response scale). Using this format, the SNS Expert Raters rated participants on a percentile from 0 to 100, with 50

representing the average for Facebook users on that indicator. For example, a rating of 60 would indicate that the ratee is judged to be higher than 60% of Facebook users on the given indicator.

In addition, 14 indicators were rated using only percentiles from 50-100 (labelled as "50-100" in Item Scaling column of Table 1; see Figure 1), because in certain instances, it would not make sense to make a below-average rating. For example, for the Specific SNS Indicator "vents about personal frustrations at home", the average person would not engage in this behaviour; they would not vent. If the ratee does vent, then the rater could assign them a percentile from 51-100, which, using the RPM format, indicates a level of venting that is above the average of 50. A true or false Specific SNS Indicator was also included (labelled as "True/False" in Item Scaling column of Table 1). In addition, a count format was also used (labelled as "Count" in Item Scaling column of Table 1; see Figure 1) for 25 indicators. The count format Specific SNS Indicators mainly consisted of frequency information, such as "number of written posts or comments by the candidate that refer to alcohol or drug use".

There were 27 Specific SNS Indicators that were less subjective than the other Specific SNS Indicators, and thus were deemed to only require ratings from one SNS Expert Rater. These Specific SNS Indicators are referred to as objective data in the current study, and are listed in the first column of Table 1 with "Ob" titles. The objective data makes use of the same count format previously described for 22 of the 27 objective Specific SNS Indicators. The other 5 objective Specific SNS Indicators relied on a categorical format (labelled as a numbered range in Item Scaling column of Table 1; see Figure 1). For example, one of the categorical Specific SNS Indicators was whether they had their birth date on their

"About" page. A "0" was given if they did not give any birth date information, a "1" if they had the date but not the year, and a "2" was given if they gave the date and the year.

Global SNS Rating. The same RPM format described above was also used for the three SNS ratings used to create the Global SNS Rating. The three ratings consisted of: "how attractive is this applicant as a potential employee of an organization", "how likely would you be to offer this person a job?", "how qualified is this person for the job?".

2.4 Results

2.4.1 Specific SNS Indicator Taxonomy Creation

A search of both peer-reviewed and non-peer-reviewed (e.g., news, online commentaries, etc.) sources that refer to specific Facebook behaviours was conducted in order to identify potential Specific SNS Indicators. Within these sources, the authors either made mention of behaviours that they thought would be related to selection, or they listed behaviours that I thought had the potential to be related to selection by tapping job-relevant human attributes. This initial list consisted of 426 Specific SNS Indicators. Any redundancies were then removed from the list and very similar indicators were consolidated wherever possible, resulting in a refined list of 157 Specific SNS Indicators. It was important to develop a large number of Specific SNS Indicators because, as part of any psychometric scale construction, one should anticipate retaining only half of the items in the final scales (Hinkin, 1998)

There were also some additional Specific SNS Indicators that needed to be created. Facebook is continuously changing and updating, and at one point Facebook had removed a feature that easily showed users' "likes", along with a number of other changes. As a result,

five new indicators were created to stay current with the updates that Facebook had made to its programming, and that refined list of 157 Specific SNS Indicators became 162.

2.4.2 Preliminary Evaluation of the Specific SNS Indicators

Evaluating the Specific SNS Indicator wording quality. The Taxonomy Expert Raters, mentioned previously, evaluated the wording quality of the Specific SNS Indicators to aid in item retention decisions. To help evaluate wording quality, they were asked to consider if the item was clear and easy to understand, whether colloquialisms or technical jargon made it more difficult to understand, and whether double-barrelled questions were used. Wording quality was indicated on a scale from -2 (very poor quality) to +2 (very good quality). This served as a tool for making item retention decisions, as indicators with low quality ratings (average rating less than the absolute value of 1.00) either had their wording changed or were dropped. An average rating of 1.00 was chosen as the cutoff, because only well-worded indicators were desired. Before removing each of these indicators, however, I ensured that there were ample numbers of indicators left in order to measure each attribute.

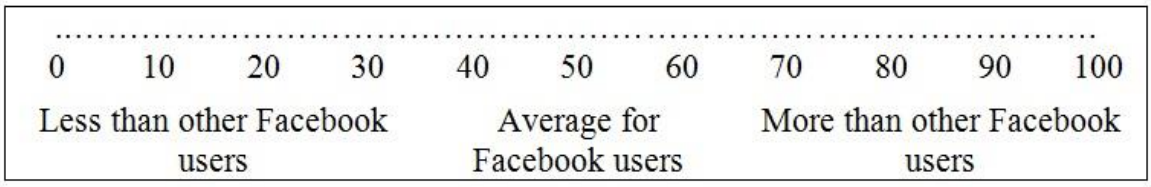
Evaluating the relevance of the Specific SNS Indicators to human attributes. After reading definitions of human attributes commonly assessed by traditional selection scales (Ashton & Lee, 2008; Costa & McCrae, 1992; Jackson, 2004; Marcus, 2006; Paulhus, 1984, 1988; Sackett & Wanek, 1997), the same set of Taxonomy Expert Raters articulated a taxonomy of Specific SNS Indicators by rating the relevance of the Specific SNS Indicators to the human attributes. The attributes consisted of general mental ability, personality, and integrity, as well as their subdimensions listed in the measures section (see Table 2 for all dimensions and subdimensions that were considered). To complete the

taxonomy, raters completed a matrix in Excel, rating the relevance of each Specific SNS Indicator (on each row) to the human attributes listed in each column. Both positive and negative ratings were acceptable, as a Specific SNS Indicator might be associated with the negative pole of a given attribute. For example, when a user posts a considerable amount of negative content about another user on his/her own SNS, there may be a relation to the negative pole of Agreeableness. The following scale from O'Neill, Lewis, and Carswell (2011) was used by the raters: -2 (strong negative relation with the attribute), -1 (slight negative relation with the attribute), 0 (unrelated to the attribute), +1 (slight positive relation with the attribute), and +2 (strong positive relation with the attribute). Because the Specific SNS Indicators may be multidimensional, it was acceptable for a Specific SNS Indicator to have relevance ratings for more than one attribute. Only Specific SNS Indicators with an average absolute rating of 1 (i.e., either a slight positive or negative relation with a human attribute) or above were retained (see X1-X88 and Ob1-Ob27 in Table 1). Average absolute ratings below a 1 were considered too low to be a part of the conceptual, judgment-based approach. Therefore, after removing the aforementioned Specific SNS Indicators, the list of 162 indicators was reduced to 115.

The ICC(2,k) interrater reliabilities for the taxonomy ratings ranged from .65 to .94, with an average of .85. In addition, there was an average ICC(2,k) of .77 for the five additional items (see Specific SNS Indicator Taxonomy Creation section) that were rated by four Taxonomy Expert Raters. Given that .70 has often been used to distinguish between high and low interrater agreement (Lance et al., 2006; LeBreton et al., 2003), all of these ICCs represent good interrater reliability, with the exception of the ratings for Quantitative Ability (.65). All of these interrater reliabilities, however, were statistically significant.

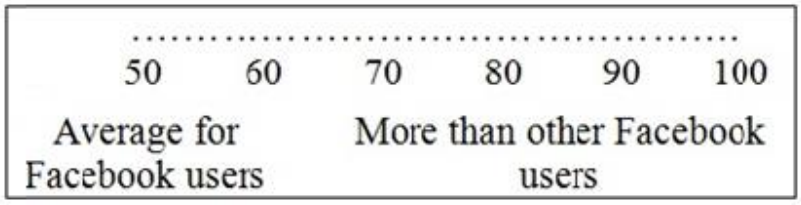
Relative Format

0-100 scale:



Engage in the behaviour more than ____% of all Facebook users.

50-100 scale :



Engage in the behaviour more than ____% of all Facebook users.

Count Format:

List the number of times that the candidate is drinking alcohol: ____ times.

Categorical Format:

Has birth date on the About page (0 if no, 1 if birthday but no year given, 2 if year given): ___

Figure 1 (Study 1): Specific SNS Indicator Response Scales

Table 1 (Study 1): Number of Respondents, Interrater Reliabilities, Means, and Standard Deviations for SNS Indicators

Item	*N*	*ICC(2,k)*	*(2,1)*	*Item Scaling*	*Mean*	*SD*	*SNS Key*
X1	10	*.30**	*.14**	*Count*	*6.34*	*11.05*	*HH, Int*
X2	8	*.38***	*.17***	*0-100*	*53.00*	*13.18*	
X3	10	*.27**	*.11**	*Count*	*6.33*	*11.83*	*Int*
X4	8	*.51***	*.26***	*0-100*	*54.84*	*15.11*	
X5	11	*.32**	*.14**	*Count*	*0.83*	*1.73*	*Int*
X6	11	*.55***	*.29***	*0-100*	*30.04*	*17.18*	*Int*
X7	10	*.73***	*.48***	*0-100*	*37.80*	*19.13*	
X8	10	*.74***	*.48***	*0-100*	*36.20*	*19.42*	
X9	11	*.77***	*.53***	*0-100*	*58.51*	*15.06*	
X10	11	*.69***	*.42***	*0-100*	*53.67*	*15.04*	*IM*
X11	11	*-.07*	*-.02*	*Count*	*0.83*	*3.51*	
X12	11	*.61***	*.34***	*Count*	*0.22*	*0.43*	
X13	11	*.59***	*.33***	*Count*	*2.76*	*3.38*	*Ag*
X14	11	*.16*	*.06*	*Count*	*0.29*	*0.56*	
X15	11	*.80***	*.57***	*Count*	*0.72*	*1.34*	*Int*
X16	11	*.92***	*.80***	*Count*	*1.16*	*3.35*	
X17	11	*.10*	*.03*	*Count*	*0.19*	*0.60*	
X18	7	*.88***	*.71***	*0-100*	*55.41*	*18.55*	
X19	11	*.18*	*.07*	*Count*	*0.08*	*0.24*	
X20	1	*.36***	*.16***	*0-100*	*44.29*	*28.46*	
X21	11	*.37***	*.16***	*Count*	*0.28*	*0.54*	
X22	6	*.44***	*.21***	*0-100*	*59.58*	*21.30*	
X23	11	*.38***	*.17***	*Count*	*0.04*	*0.13*	
X24	1	*Error*	*Error*	*0-100*	*47.00*	*21.61*	
X25	11	*.84***	*.64***	*Count*	*2.34*	*3.65*	
X26	11	*.46***	*.22***	*Count*	*0.28*	*0.70*	*HH, Int*
X27	6	*.86***	*.67***	*0-100*	*52.24*	*18.53*	
X28	11	*.10*	*.04*	*Count*	*0.27*	*0.60*	
X29	2	*.17*	*.06*	*0-100*	*54.88*	*14.56*	
X30	11	*.07*	*.03*	*Count*	*0.42*	*0.83*	
X31	9	*.15*	*.05*	*Count*	*0.05*	*0.18*	
X32	11	*.18*	*.07*	*Count*	*0.08*	*0.24*	

(continued)

Table 1 (continued): Number of Respondents, Interrater Reliabilities, Means, and Standard Deviations for SNS Indicators

Item	*N*	*ICC(2,k)*	*ICC(2,1)*	*Item Scaling*	*Mean*	*SD*	*SNS Key*
X33	11	*.89***	*.72***	*Count*	*1.16*	*2.02*	
X34	11	*.11*	*.04*	*Count*	*4.58*	*12.35*	
X35	11	*.04*	*.01*	*Count*	*0.02*	*0.07*	
X36	11	*.85***	*.66***	*Count*	*0.13*	*1.41*	
X37	11	*.85***	*.65***	*0-100*	*51.01*	*16.73*	*Ext*
X38	11	*.18*	*.07*	*Count*	*1.42*	*2.28*	
X39	10	*.67***	*.40***	*0-100*	*44.21*	*12.17*	*Op*
X40	11	*.73***	*.47***	*0-100*	*55.80*	*13.75*	
X41	11	*.53***	*.28***	*0-100*	*47.47*	*7.98*	*VA*
X42	11	*.71***	*.45***	*0-100*	*55.59*	*11.37*	*Con*
X43	11	*.70***	*.44***	*0-100*	*56.88*	*10.98*	*Con*
X44	11	*.75***	*.50***	*0-100*	*56.12*	*13.63*	*Con*
X45	11	*.62***	*.35***	*0-100*	*47.82*	*8.39*	*VA*
X46	11	*.72***	*.46***	*0-100*	*50.50*	*15.22*	*IM*
X47	11	*.42***	*.19***	*0-100*	*47.88*	*9.11*	*Con, VA*
X48	11	*.78***	*.56***	*50-100*	*56.45*	*7.21*	
X49	11	*.76***	*.54***	*50-100*	*54.43*	*6.59*	
X50	11	*.65***	*.38***	*0-100*	*55.31*	*10.99*	*Con*
X51	11	*.61***	*.34***	*0-100*	*50.06*	*11.71*	
X52	11	*.52***	*.26***	*0-100*	*47.09*	*10.16*	
X53	10	*.64***	*.37***	*0-100*	*51.12*	*10.29*	
X54	11	*.33**	*.14**	*50-100*	*51.40*	*3.14*	
X55	11	*.57***	*.30***	*50-100*	*56.53*	*6.08*	
X56	11	*.56***	*.30***	*50-100*	*51.76*	*5.13*	
X57	11	*.58***	*.31***	*0-100*	*61.59*	*8.70*	*Ag*
X58	11	*.76***	*.52***	*50-100*	*55.89*	*6.62*	
X59	11	-	-	*True/ False*	*0.00*	*0.04*	
X60	11	*-.04*	*-.01*	*50-100*	*50.41*	*1.93*	
X61	11	*.50***	*.25***	*0-100*	*61.19*	*7.83*	
X62	10	*.66***	*.39***	*0-100*	*52.92*	*14.15*	
X63	11	*.83***	*.62***	*0-100*	*51.92*	*15.44*	
X64	11	*.66***	*.39***	*0-100*	*41.42*	*12.67*	
X65	11	*.60***	*.34***	*0-100*	*45.09*	*10.90*	

(continued)

Table 1 (continued): Number of Respondents, Interrater Reliabilities, Means, and Standard Deviations for SNS Indicators

Item	*N*	*ICC(2,k)*	*ICC(2,1)*	*Item Scaling*	*Mean*	*SD*	*SNS Key*
X66	141	*.44***	*.21***	*50-100*	*50.52*	*1.81*	
X67	141	*.75***	*.50***	*0-100*	*55.80*	*14.90*	*Ex*
X68	141	*.77***	*.53***	*0-100*	*51.44*	*13.01*	
X69	141	*.60***	*.34***	*0-100*	*57.71*	*9.98*	
X70	141	*.73***	*.47***	*0-100*	*56.32*	*13.01*	
X71	141	*.67***	*.41***	*0-100*	*52.74*	*13.32*	
X72	141	*.64***	*.38***	*0-100*	*52.67*	*11.57*	*Ex*
X73	141	*.71***	*.45***	*0-100*	*49.68*	*13.08*	*Op*
X74	141	*.74***	*.49***	*0-100*	*57.81*	*12.96*	*Op, Ex*
X75	141	*.70***	*.43***	*0-100*	*52.74*	*12.96*	*IM*
X76	134	*.51***	*.26***	*0-100*	*49.53*	*4.13*	
X77	141	*.79***	*.55***	*50-100*	*67.03*	*8.67*	
X78	140	*.69***	*.43***	*0-100*	*51.04*	*12.37*	*IM*
X79	141	*.67***	*.40***	*0-100*	*52.62*	*10.89*	*IM*
X80	138	*.35***	*.15***	*50-100*	*56.00*	*7.25*	*Con, IM*
X81	138	*.47***	*.23***	*0-100*	*40.53*	*11.02*	
X82	138	*.46***	*.22***	*0-100*	*50.18*	*11.86*	
X83	140	*-.06*	*-.02*	*Count*	*2.66*	*4.60*	
X84	140	*.79***	*.55***	*50-100*	*50.83*	*3.04*	
X85	140	*.49***	*.24***	*50-100*	*58.54*	*8.17*	
X86	140	*.58***	*.32***	*50-100*	*56.12*	*7.30*	
X87	141	*.24**	*.10**	*0-100*	*0.33*	*0.30*	
X88	134	*.24**	*.10***	*50-100*	*53.30*	*5.42*	
Ob1	134	-	-	*0-2*	*1.30*	*.55*	
Ob2	134	-	-	*0-1*	*0.81*	*.40*	
Ob3	134	-	-	*0-1*	*0.61*	*.87*	
Ob4	134	-	-	*0-3*	*2.78*	*.72*	
Ob5	134	-	-	*0-3*	*0.23*	*.79*	
Ob6	130	-	-	*Count*	*19.68*	*21.82*	
Ob7	129	-	-	*Count*	*2.81*	*13.84*	
Ob8	130	-	-	*Count*	*2.12*	*3.39*	
Ob9	129	-	-	*Count*	*0.63*	*1.21*	
Ob10	130	-	-	*Count*	*10.32*	*12.65*	

(continued)

Table 1 (continued): Number of Respondents, Interrater Reliabilities, Means, and Standard Deviations for SNS Indicators

Item	*N*	*ICC(2,k)*	*ICC(2,1)*	*Item Scaling*	*Mean*	*SD*	*SNS Key*
Ob11	125	-	-	*Count*	*3.21*	*5.36*	
Ob12	124	-	-	*Count*	*2.19*	*3.81*	
Ob13	124	-	-	*Count*	*0.09*	*.36*	
Ob14	124	-	-	*Count*	*1.19*	*5.98*	
Ob15	124	-	-	*Count*	*0.79*	*2.02*	
Ob16	124	-	-	*Count*	*16.57*	*34.38*	
Ob17	124	-	-	*Count*	*5.12*	*10.39*	
Ob18	111	-	-	*Count*	*52.25*	*36.78*	*Ex*
Ob19	111	-	-	*Count*	*1.53*	*4.06*	
Ob20	111	-	-	*Count*	*15.35*	*17.16*	
Ob21	111	-	-	*Count*	*9.27*	*7.56*	
Ob22	111	-	-	*Count*	*1.57*	*4.22*	
Ob23	125	-	-	*Count*	*10.48*	*23.63*	
Ob24	120	-	-	*Count*	*9.99*	*14.08*	
Ob25	111	-	-	*Count*	*1.10*	*3.79*	
Ob26	123	-	-	*Count*	*2007.60*	*1.38*	
Ob27	120	-	-	*Count*	*747.43*	*378.29*	*Ex*
Ov1	111	*.64***	*.37***	*0-100*	*53.19*	*9.97*	
Ov2	111	*.71***	*.45***	*0-100*	*54.80*	*11.83*	
Ov3	111	*.64***	*.38***	*0-100*	*52.41*	*14.55*	

Note. Bolded items indicate those items which were retained to create the Specific SNS Indicator scales. Items with an "Ob" refer to less subjective count and categorical items that were rated by one SNS Expert Rater. Ob26 is the year the participant first signed up for Facebook, and Ob27 is number of Facebook friends s/he has. Items with an "Ov" refer to the overall ratings of Facebook profiles. The 'SNS Key' column refers to the attribute coverage for the indicators that were retained for the Specific SNS Indicator scales. ICC(2,k) represents an estimate of interrater reliability if multiple raters were doing the ratings. ICC(2,1) represents an estimate of interrater reliability if a single rater were doing the ratings. ICCs for X59 could not be calculated because a categorical true/false format was used; however, the Fleiss Kappa (Fleiss, 1971), a measure of agreement used for categorical formats, was .99. Ag = Agreeableness; Ex = Extraversion; Con = Conscientiousness; Op = Openness; HH = Honesty-Humility; IM = Impression Management; VA = Verbal Ability; Int = Integrity. *p < .01, **p < .001

Interrater reliabilities for all retained Specific SNS Indicators. After the taxonomy was completed, a different set of expert raters, three SNS Expert Raters (see the procedures section), used this list of 115 Specific SNS Indicators to rate the participants. Overall, the average ICC(2,*k*) interrater reliability in this group was .53, with a range of -.07 to .92. This interrater reliability analysis considered the indicators as single-item scales. Paunonen (1984) suggested that single-item scales for psychological tests typically fall below .30. Thus, an average ICC(2,*k*) interrater reliability of .53 is not surprising. However, these reliabilities are still fairly low. The average did increase slightly once only the Specific SNS Indicators with significant ICC(2,*k*) interrater reliabilities were retained (see the following paragraph). It is also worth mentioning that the ICC(2,*k*) interrater reliabilities were considerably higher, on average, for the indicators using the RPM format (average = .60) than they were for those with the count format (average = .35). For the Global SNS Rating indicators, the average interrater reliability was .66. This will be discussed in further detail in the discussion section. However, overall, the ICC reliabilities were fairly low for the Specific SNS Indicators and the Global SNS Rating indicators.

H_1 was partially supported, as 72 of the 88 Specific SNS Indicators rated by more than one rater had ICC(2,k) interrater reliabilities that were significantly different from zero (see those with one or two asterisks in Table 1; reliabilities above .24 were significant). Upon removal of Specific SNS Indicators with non-significant ICC(2,*k*) interrater reliabilities, the average interrater reliability was now .61, with a range of .24 to .92. The discarded 16 indicators did not concentrate on any one attribute area, so their removal did not result in a loss of coverage for any particular attribute. The three SNS Experts' ratings were averaged together to create single scores for each of the 72 Specific SNS Indicators.

Table 2 (Study 1): Dimensions and Subdimensions of Attributes Used

Dimensions	Subdimensions
General Mental Ability	Verbal Ability
	Quantitative Ability
Personality	Agreeableness
	Extraversion
	Conscientiousness
	Openness
	Neuroticism
	Honesty-Humility
	Impression Management
Integrity	Overt Integrity
	Personality-Based Integrity

ICC(2,1) interrater reliabilities were also included in Table 1 to provide the reader with an estimate of what the interrater reliability would be if only one rater were doing the ratings. The average for the ICC(2,1) interrater reliabilities was .38, with a range of .10 to .80. Objective Specific SNS Indicators were not included as part of hypothesis H_1 because they were only assessed by a single rater. Estimates of interrater reliability require multiple raters.

In addition to the indicators that were removed because they had nonsignificant ICC(2,k), interrater reliabilities, there were two "follow-up" indicators that had to be removed. These follow-up indicators were used to get more information from a previous indicator (e.g., after answering X17 ("number of positive statements about former/current employer(s), work colleagues, clients, supervisors, classmates, or teachers/professors"), the rater was asked to answer X18 ("how positive were these statements?"). Because X17 and X19 had poor interrater reliability (see Table 1), their follow-up indicators (X18 and X20) were not included as part of the Specific SNS Indicator scales created in subsequent analyses (see below).

The three items measuring the Global SNS Ratings of participants' Facebook profiles had ICC(2,k) interrater reliabilities ranging from .64 to .71 (see Table 1). The three SNS Expert Raters' Global SNS Ratings were averaged in order to create a single score for each of the three Global SNS Ratings described above.

2.4.3 Evaluation of Structural Validity

In order to evaluate the structure of the Specific SNS Indicators, a series of models had to be tested for each attribute. The goal was to create individual Confirmatory Factor

Analysis (CFA) models in order to assess the unidimensionality of each Specific SNS Indicator scale. Each Specific SNS Indicator scale was tested individually because of the small sample size in the current study. Testing all of the Specific SNS Indicator scales in a single model would have required a substantial sample size.

According to Kline (2011), one would be well served to start off with less complex measurement models before investigating more complex models. This way any problems with the individual measurement models could be more easily diagnosed before proceeding to the more complex models that include all respective constructs. Based upon this recommendation, what follows is a general overview of the steps that were taken. The following steps are described in more detail in subsequent sections.

First, measurement models were tested using CFAs for each of the traditional scales (Agreeableness, Extraversion, Conscientiousness, Openness to Experience, Impression Management, Neuroticism, Honesty-Humility, Verbal Ability, Quantitative Ability, and Integrity).

Second, using the final list of Specific SNS Indicators that were retained after the preceding sections, measurement models were created using Exploratory Factor Analyses (EFAs) for each of the Specific SNS Indicator Scales (Agreeableness, Extraversion, Conscientiousness, Openness to Experience, Impression Management, Neuroticism, Honesty-Humility, Verbal Ability, Quantitative Ability, and Integrity). EFAs were used for the Specific SNS Indicator measurement models as opposed to CFAs, because unlike the established, traditional scales, their factor structure required more exploration, which is better served by EFA over CFA (Asparouhov & Muthen, 2009).

Finally, a set of Combined CFA Models were tested. They consisted of the Specific SNS Indicators and the traditional selection scales as correlated latent variables with the Global SNS Rating as a correlated, directly observed variable. These models were tested separately for each attribute (Agreeableness, Extraversion, Conscientiousness, Openness to Experience, Impression Management, Neuroticism, Honesty-Humility, Verbal Ability, Quantitative Ability, and Integrity) in order to assess the structural validity of the Specific SNS Indicators and the Global SNS Rating of participants' SNS profiles. For an example of one of these Combined CFA Models, please see Figure 2 (see Appendix A for the remaining Combined CFA Model diagrams).

For the CFAs and EFAs, M*plus*7 (Muthén & Muthén, 2012) was used for the analyses, with maximum likelihood used as the estimator. For all of the above models, the goodness-of-fit indicators described below were used to assess the fit of each model. I have also included general rules of thumb for what are considered by several researchers (i.e., Browne & Cudeck, 1993; Hu & Bentler, 1999; Jörgen & Sörbom, 1993) to be indications of good model fit. These rules of thumbs are only provided to help the reader better understand the nature of the fit indices. Strict rules of thumbs or cut off scores were not used in the current study, because a number of researchers (e.g., Goffin, 2007; Marsh, Hau, Kit-Tai, & Wen, 2004) believe that they may be too conservative, and could lead to the rejection of genuinely acceptable models.

The comparative fit index (CFI) examines the relative improvement in fit between a hypothesized model and a statistical baseline model. This baseline model is, in most cases and in the case of the current study, an independence (null) model, which assumes zero covariances for observed variables (Kline, 2011). CFIs of .90 or greater are generally

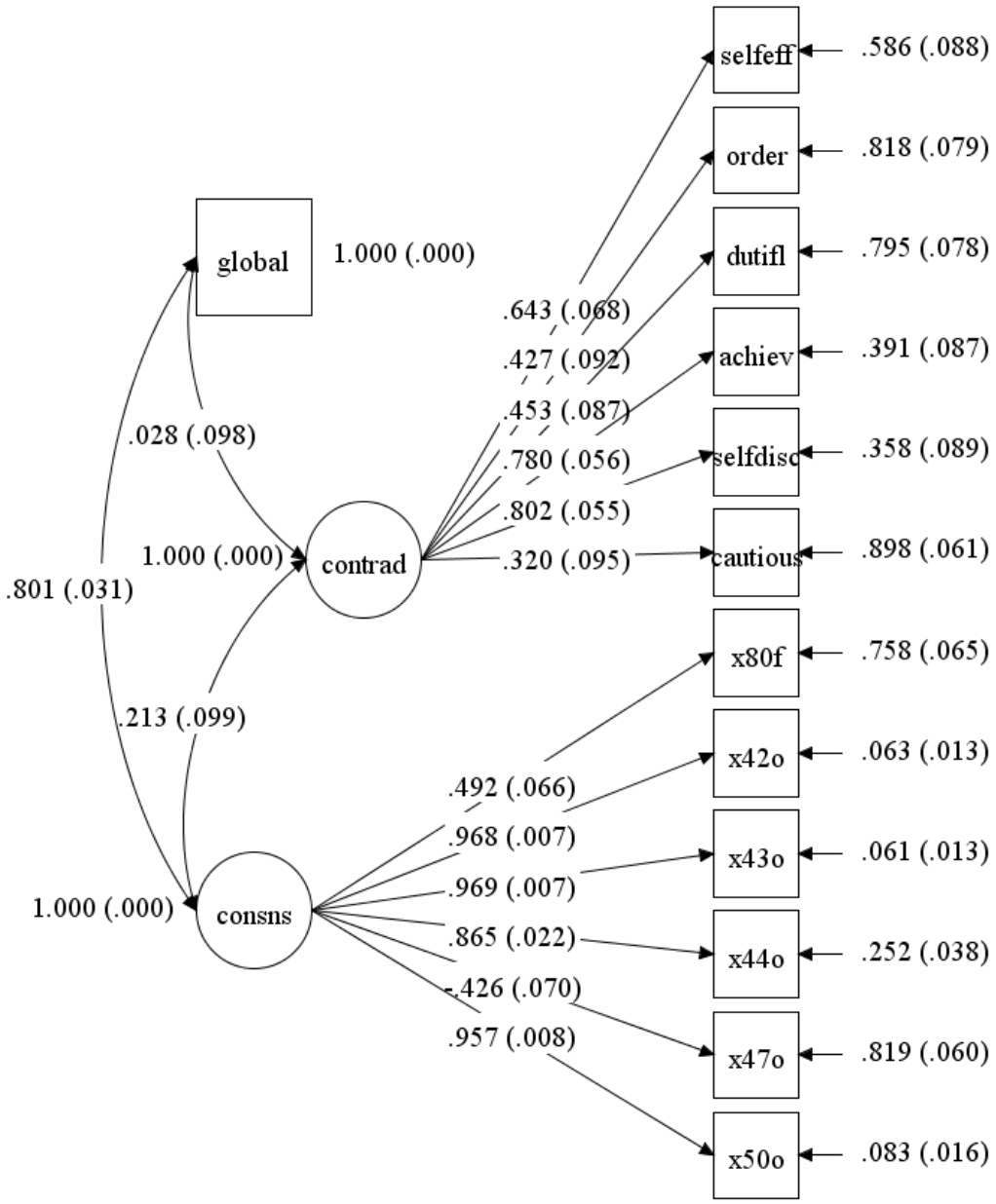


Figure 2. Confirmatory factor analysis for Conscientiousness. contrad = traditional scale of Conscientiousness; consns = Specific SNS Indicator scale of Conscientiousness; global = Global SNS Rating; selfeff = self-efficacy; order = orderliness; dutifl = dutifulness; achiev = achievement; selfdisc = self-discipline; cautious = cautiousness. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.

Figure 2 (Study 1): Combined Confirmatory Factor Analysis for Conscientiousness

considered to be indicative of adequate model fit, while values above .95 are preferred (Hu & Bentler, 1999).

The root mean square error of approximation (RMSEA) represents the differences between the hypothesized model (with optimally chosen parameter estimates) and the population covariance matrix (Hooper, Coughlan, & Mullen, 2008). The RMSEA ranges from 0 to 1, with lower values presumed to indicate a better fit. RMSEAs less than .08 represent reasonable errors of approximation, but RMSEAs less than .05 represent "close fit" (Browne & Cudeck, 1993, Jörgen & Sörbom, 1993).

The standardized root mean square residual (SRMR) is the square root of the difference between the sample covariance matrix and the model covariance matrix (Hooper, Coughlan, & Mullen, 2008). It ranges from 0 to 1, with lower values indicating better model fit. Models with SRMRs less than .08 are presumed to have a good fit according to Hu and Bentler (1999).

At this point, it is worth mentioning that because of the smaller sample size and the low degrees of freedom that could occur for some of the models, the resulting power could be low (Goffin, 2007; MacCallum, Browne, & Sugawara, 1996). Low power can be a concern because it means that the particular model might be supported when it should not be (Fornell, 1983). Therefore, I calculated the sample sizes required to reach a power value of .80 for a CFI of .95, a CFI of .90, a RMSEA of .05, and a RMSEA of .08. It was anticipated that the Verbal Ability Combined CFA model, which had four degrees of freedom, would likely be the model with the least amount of parameters estimated, and the least amount of power. Because the traditional scale for Verbal Ability, the verbal component of the PAF,

was only a raw score, it had to be included in the model as a directly observed variable, reducing the number of parameters estimated. Therefore, the Verbal Ability Combined CFA model was chosen to form a baseline power calculation, as all other models would have more power.

According to the calculations in Kim (2005), a sample size of at least 100 would be required to attain a power of .80 for a model with 4 degrees of freedom and a CFI of .95. A sample size of 50 would be required to attain a power of .80 with 4 degrees of freedom and a CFI of. 90. Given the sample size of 141, the power was considered adequate to interpret the CFIs of each model. It also must be pointed out that models with smaller degrees of freedom and low sample sizes can lead to artificially large RMSEA values (Kenny, Kaniskan, & McCoach, 2014). According to the calculations in Kim, a sample size of 1195 would be required to attain a power of .80 for a model with 4 degrees of freedom and a RMSEA less than .05. For a RMSEA less than .08, a sample size of 467 would be required to attain a power of .80 for a model with 4 degrees of freedom. Therefore, it is recommended not to over rely on the RMSEA values for the interpretation of model fit in the current study.

Measurement models for traditional scales. Confirmatory Factor Analysis models were tested for each of the traditional scales (Agreeableness, Extraversion, Conscientiousness, Openness to Experience, Impression Management, Neuroticism, Honesty-Humility, Verbal Ability, Quantitative Ability, and Integrity). For each of the traditional CFA measurement scales, each facet (i.e., subcomponent) of the given attribute dimension served as an indicator that loaded on the latent variable for the respective traditional attribute. For the traditional personality scales, the facets of each attribute were based on the NEO-PI-R model (Costa & McCrae, 1992), and there were four items per facet.

The Agreeableness' facets were: Trust, Morality, Altruism, Cooperation, Modesty, and Sympathy. For Extraversion, the facets were Friendliness, Gregariousness, Assertiveness, Activity Level, Excitement-Seeking. For Conscientiousness, the facets were: Cheerfulness. Self-Efficacy, Order, Dutifulness, Achievement Striving, Self-Discipline, and Cautiousness. For Openness to Experience, the facets were: Imagination, Artistic, Emotionality, Adventurousness, Intellect, and Liberalism. For Neuroticism, the facets were: Anxiety, Anger, Depression, Self-Consciousness, Immoderation, and Vulnerability. Finally, the Honesty-Humility facets from the HEXACO Personality Inventory-Revised (Ashton & Lee, 2008) were: Modesty, Greed-Avoidance, Sincerity, and Fairness. Each facet for Honesty-Humility also had four items.

The other traditional scale models required different procedures. First, for Impression Management, the scale was not designed to be broken down into facets, and there were a relatively large number of items (20). Including all of the items as separate indicators for the respective latent variable would greatly increase the required sample size (Landis, Beal, & Tesluk, 2000), therefore, parcels were created to reduce the number of estimated parameters in the model (Landis et al., 2000). Four parcels were created with 4-5 items per parcel.

Second, some of the common thinking about the structure of integrity is that it is hierarchical, with Overt and Personality-Based Integrity variables loading on a single higher-order factor of integrity (Ones, Schmidt, & Viswesvara, 1994a; 1994b). Therefore, this hierarchical model was tested in the current study. Integrity served as the superfactor, and both Overt and Personality-Based Integrity served as the second-order factors. The integrity scale used in the current study, the IBES (Marcus, 2006), consists of both an Overt and a Personality-Based Integrity scale. There are four facets for Overt Integrity and five facets for

Personality-Based Integrity, and each facet served as an indicator for its respective scale. The Overt Integrity facets were: Trust, Perceived Norms, [Non]Rationalization, and Intentions/Fantasies. The Personality-Based Integrity facets were: Reliability, Stimulus Seeking [Low], Manipulativeness [Low], Self-Concept, and Trouble Avoidance. However, the model failed to converge, replicating the non-convergence and relatively poor fit found in Marcus, Hoft, and Riediger (2006) when they tested the same hierarchical model. Therefore, I decided to investigate Overt and Personality-Based Integrity as separate, non-hierarchical models. Overt Integrity was tested with its facets serving as indicators in one model, and Personality-Based Integrity was tested with its facets serving as indicators in another model.

Both of the aforementioned models converged. However, for Personality-Based Integrity, the model had extremely poor fit, $\chi^2(2) = 11.138$, $p < .01$, CFI = .932, RMSEA = .201, SRMR = .057, largely as a result of the lack of a significant factor loading for the Self-Concept facet/indicator. It seems as though 8 out of 19 of the items that were used to create this facet reflected anger more than one's positive self-concept. For instance, one item involved rating how much something would anger the respondent, and another involved rating how much the respondent would like to assault someone who was irritating him/her. Therefore, the Self-Concept facet was dropped as an indicator of Personality-Based Integrity. This left Reliability, Stimulus Seeking, Manipulativeness, Self-Concept, and Trouble Avoidance as the remaining indicators in the model. The model had a better fit, albeit with low degrees of freedom, $\chi^2(2) = .229$, *n.s.*, CFI = 1.000, RMSEA = .000, SRMR = .010.

A model could not be generated for Verbal Ability or Quantitative Ability because only the overall score for each dimension of the PAF was provided by the test publisher. Therefore, the Verbal Ability raw score would have to be treated as a directly observed

variable in any subsequent models using the Verbal Ability tradition scale. The same would apply for Quantitative Ability.

The goodness-of-fit indicators of the confirmatory factor analysis measurement models for the traditional scales, with the above changes included, can be found in Table 3. The unstandardized and standardized factor loadings for these models can be found in Table 4.

The internal consistency reliabilities were acceptable for the traditional scales, with a range from .76 to .90. The descriptives and Cronbach's alphas for the traditional scales are presented later on in Table 7 for comparison with those from the Specific SNS Indicator scales.

Measurement models for Specific SNS Indicator scales.

EFA measurement models were created for the Specific SNS Indicators believed to represent Agreeableness, Extraversion, Conscientiousness, Openness to Experience, Impression Management, Neuroticism, Honesty-Humility, Verbal Ability, Quantitative Ability, and Integrity. The indicators for these models consisted of Specific SNS Indicators that the Taxonomy Expert Raters rated as being relevant to their respective human attributes, as described in the 'evaluating the relevance of the Specific SNS Indicators to human attributes' section above. However, as mentioned previously, only indicators with significant ICC(2,k) reliabilities were included (i.e., items with asterisks in Table 1).

Parallel Analysis (O'Connor, 2000) was used to determine the number of factors in the model. It is arguably a superior method to other procedures (Wood, Tataryn, & Gorsuch, 1996), and focuses on the number of factors accounting for more variance than the variance

Table 3 (Study 1): Goodness-of-Fit Indicators for Traditional Scale CFA Measurement Models

Model	CFI	χ^2	*df*	χ^2 p-value	RMSEA	SRMR
Agreeableness	.881	22.120	9	.009	.114	.058
Extraversion	.910	26.670	9	.002	.132	.064
Conscientiousness	.867	30.796	9	.000	.146	.068
Openness	.941	16.092	9	.065	.084	.049
Neuroticism	.986	12.134	9	.206	.056	.036
Honesty-Humility	.994	2.189	2	.335	.029	.028
Impression Management	.970	4.944	2	.084	.114	.036
Verbal Ability	-	-	-	-	-	-
Quantitative Ability	-	-	-	-	-	-
Overt Integrity	.932	11.138	2	.004	.201	.057
Personality-Based Integrity	1.000	.229	2	.892	.000	.010

Note. N = 141. CFI = comparative fit index, χ^2 = chi-squared, *df* = degrees of freedom, χ^2 p-value = significance value of chi-squared, RMSEA = Root Mean Square Approximation, SRMR = standardized root mean square residual. CFA models for Verbal Ability and Quantitative Ability could not be calculated because only the overall scores for each were provided by the test publisher.

Table 4 (Study 1): Factor Loadings for Traditional Scale CFA Measurement Models

Item	Factor Loadings	
	Unstandardized	Standardized
Agreeableness		
Trust	.23* (.07)	.33
Morality	.41* (.05)	.74
Altruism	.33* (.05)	.70
Cooperation	.44* (.08)	.58
Modesty	.17* (.06)	.29
Sympathy	.24* (.05)	.49
Extraversion		
Friendliness	.62* (.06)	.90
Gregariousness	.56* (.07)	.69
Assertiveness	.39* (.07)	.56
Activity Level	.17* (.07)	.25
Excitement-Seeking	.25* (.06)	.41
Cheerfulness	.53* (.06)	.74
Conscientiousness		
Self-Efficacy	.35* (.05)	.64
Order	.36* (.09)	.43
Dutifulness	.22* (.05)	.44
Achievement Striving	.59* (.07)	.77
Self-Discipline	.58* (.06)	.82
Cautiousness	.25* (.08)	.32
Openness		
Imagination	.40* (.07)	.60
Artistic Interests	.49* (.07)	.65
Emotionality	.30* (.05)	.57
Adventurousness	.36* (.07)	.56
Intellect	.49* (.08)	.64
Liberalism	.25* (.07)	.40
Honesty-Humility		
Modesty	.46* (.09)	.68

(continued)

Table 4 (continued): Factor Loadings for Traditional Scale CFA Measurement Models

Item	Factor Loadings	
	Unstandardized	Standardized
Greed Avoidance	.37* (.09)	.50
Sincerity	.21 (.08)	.32
Fairness	.40* (.10)	.49
Impression Management		
Parcel1	1.86* (.25)	.74
Parcel2	2.14* (.28)	.75
Parcel3	1.56* (.29)	.55
Parcel4	1.57* (.26)	.60
Verbal Ability		
Verbal Ability	-	-
Overt Integrity		
Trust	.22* (.03)	.60
Perceived Norms	.19* (.04)	.44
(Non)Rationalization	.38* (.04)	.89
Intentions/Fantasies	.39* (.05)	.76
Personality-Based Integrity		
Reliability	.11* (.04)	.29
Stimulus Seeking (Low)	.23* (.06)	.44
Manipulativeness (Low)	.42* (.08)	.86
Trouble Avoidance	.20* (.06)	.41

Note. $N = 141$. Standard errors are in parentheses.
*$p < .05$

derived from random data (O'Connor, 2000). The EFAs were run in Mplus using a maximum likelihood robust estimator. Geomin served as the rotation method. Geomin was specifically chosen because it has provided good results in various situations, such as for complex factor patterns (Browne, 2001). Also, Geomin is an oblique rotation that allows the factors to be correlated (Muthén & Muthén, 2012). It was important to allow the factors to correlate as the Taxonomy Expert Raters already rated that indicators may tap multiple constructs (i.e., they are multidimensional).

Fit statistics for the models were interpreted at various stages to serve as a barometer for the effect any changes had on the fit of the models. Such changes included indicator elimination from the model, or reductions in number of factors.

Finally, specific SNS Indicators had to have significant factor loadings in order to be included as part of the final measurement model. A factor loading cutoff was used to eliminate Specific SNS Indicators with small loadings on their factors. Tabachnick and Fidell (2007) have recommended that factor loadings should be greater than .320 (considered a poor loading), because at .320, a factor loading would account for 10% of the explained variance in the factor.

The steps taken to conduct each EFA for the Specific SNS Indicator scales are detailed below, along with the final set of Specific SNS Indicators that were used to represent each model in the Combined CFA Models. The goodness-of-fit indicators of these measurement models can be found in Table 5.

Agreeableness. The initial indicators used for the Agreeableness EFA consisted of X13, X21, X22, and X57. The parallel analysis pointed to a one-factor EFA solution,

which had good fit, $\chi^2(6) = 44.795$, $p < .01$, CFI = .977, RMSEA = .056, SRMR = .075. However, X21 did not have a significant factor loading (-.106), and was removed. As a result, X22 had to be removed as well, because it was a follow-up indicator to X21. X13 and X57 were the only remaining Specific SNS Indicators. X13 and X57 were moderately correlated ($r = .483$, $p < .05$), which was used as justification for aggregating them into a single, directly observed variable in subsequent models. The two Agreeableness Specific SNS Indicators were standardized using z-scores, and then combined to form the composite scale. The use of z-scores was to give equal weighting to both items, because they each had different means and standard deviations (X13 was a count item, and X57 was a 0-100 RPM format item). This method was used to prevent the scaling of one item from overly influencing the weighting of the composite scale.

Because there were not enough Specific SNS Indicators for Agreeableness in order to assess the fit of the model, reliability was investigated as a test of unidimensionality. Spearman-Brown reliability has been deemed a better match for two-item scales than Cronbach's alpha reliability (Eisinga, Frotenhuis, & Pelzer, 2013). Therefore, the Spearman-Brown estimate of reliability was calculated, and it was .65. According to some researchers, reliabilities above .60 are minimally acceptable (see Clark & Watson, 1995).

Extraversion. The Extraversion EFA consisted of X7, X8, X37, X67, X72, X74, Ob11, Ob18, Ob24, and Ob27. Again, the Ob variables refer to the objective data that were deemed to only require ratings from one SNS Expert Rater. The parallel analysis pointed to a two-factor solution, which had relatively good fit, $\chi^2(26) = 90.649$, $p < .01$, CFI = .927, RMSEA = .133, SRMR = .074. X7 and X8 seemed to load very highly (1.019 and .968, respectively) on the first factor. These indicators seemed to be tapping another, somewhat-

related construct (e.g., integrity) because they dealt with more criminal and inappropriate behaviours than did the other indicators in the first factor. Therefore, X7 and X8 were removed, and the EFA was conducted again as a one-factor solution. The fit of the one-factor solution was an improvement, $\chi^2(20) = 100.457$, $p < .01$, CFI = .766, RMSEA = .169, SRMR = .104, yet Ob24 did not have a significant factor loading. Therefore, it was removed as well. Finally, Ob11 was removed because it had a loading (.242) below .320. Therefore, the remaining Specific SNS Indicators were: X37, X67, X72, X74, Ob18, and Ob27. The Extraversion EFA measurement model was good, $\chi^2(9) = 57.907$, $p < .01$, CFI = .840, RMSEA = .196, (90% CI = .150–.246; probability less than .05 = .000), SRMR = .088 (see Table 5). The Geomin rotated factor loadings can be found in Table 6. The Extraversion SNS scale had a Cronbach's alpha reliability of .73, adding support to its unidimensionality.

Conscientiousness. X1, X2, X3, X4, X6, X7, X9, X12, X26, X27, X42, X43, X44, X47, X50, X62, and X80 were included in the EFA for Conscientiousness. The parallel analysis pointed to a three-factor EFA solution, however it failed to converge. The two-factor EFA solution, however, did converge, $\chi^2(103) = 375.00$, $p < .01$, CFI = .837, RMSEA = .137, SRMR = .115. X2, X4, X12, and X27 were dropped, as they did not load on either factor. X1, X3, X6, X7, X9, X26, and X62 loaded on the first factor. X9, X42, X43, X47, X50, X62, and X80 loaded on the second factor. The correlation between the two factors was small, $r = -.097$, $p < .05$. Initially I thought that the two factors were Achievement Striving and Methodicalness, because these are the two factors that underlie most measures of Conscientiousness (Ashton, Jackson, Paunonen, Helmes, & Rothstein, 1995). However, the content of the respective Specific SNS Indicators was not consistent with that breakdown. The first factor seemed to be tapping integrity, because the indicator content dealt more with

criminal and inappropriate behaviour than it did with Conscientiousness. Indeed, as discussed later, some of these Specific SNS Indicators were used to predict the Specific SNS Integrity scale. Therefore, the Specific SNS Indicators from this factor were dropped. Finally, X9 was removed because the factor loading (.298) was lower than .320. This left X42, X43, X44, X47, X50, and X80 remaining, resulting in good fit $\chi^2(9) = 16.897$, $p < .01$, CFI = .991, RMSEA = .079, SRMR = .026 (see Table 5). The Geomin rotated factor loadings can be found in Table 6. The unidimensionality of the Conscientiousness SNS scale was additionally supported because its Cronbach's alpha was .81.

Openness to Experience. The indicators for the Openness to Experience EFA consisted of X39, X53, X73, X74, Ob7, and Ob23. The parallel analysis pointed to a one-factor solution. The one-factor EFA approached adequate fit, $\chi^2(9) = 27.855$, $p < .01$, CFI = .810, RMSEA = .122, SRMR = .074, but Ob7 did not have a significant factor loading (.063). Therefore, Ob7 was removed. Ob23 and X53 had loadings below .320 (.234 and .306, respectively), and were also removed, leaving just X39, X73, and X74. The resulting model had 0 degrees of freedom and a resulting perfect fit, so its EFA results could not be interpreted. However, the internal consistency reliability ($\alpha = .67$) was minimally acceptable (Clark & Watson, 1995), adding some support for the unidimensionality of this model.

Neuroticism. The Taxonomy Expert Raters only identified two Specific SNS Indicators that may tap Neuroticism. However, I could not justify the creation of a composite in this case because X1 and X2 were too weakly correlated ($r = .017$, *n.s.*), and the resulting reliability would be unacceptably low (Spearman-Brown reliability = .03). Therefore, this model was dropped from further analysis.

Honesty-Humility. The items for Honesty-Humility consisted of X1, X2, X26, and X27. With only four indicators available, the model would not be identified if testing two or more factors. Therefore, a one-factor solution was tested. That one-factor EFA solution resulted in a good fit, $\chi^2(2) = 1.730$, p = .421, CFI = 1.000, RMSEA = .000, SRMR = .066. However, X2 and X27 did not have significant factor loadings (.030 and .230, respectively), and were thus removed. X1 and X26 were moderately correlated ($r = .424$, $p < .05$), therefore, as with Agreeableness, a composite was created for these two Specific SNS Indicators. The Spearman-Brown reliability was calculated for Honesty-Humility, and it was .60. When rounded, the reliability would reach the cutoff for minimal acceptability (Clark & Watson, 1995).

Impression Management. There were 23 potential Specific SNS Indicators for Impression Management: X1, X2, X3, X4, X5, X6, X7, X8, X10, X12, X15, X21, X26, X27, X46, X56, X62, X75, X78, X79, X80, X81, and X88. The parallel analysis pointed to a three-factor solution, and this solution had poor fit, $\chi^2(187) = 813.889$, p < .01, CFI = .726, RMSEA = .154, SRMR = .088. The first factor was the most conceptually related (based on indicator content) to Impression Management, and consisted of X10, X12, X27, X46, X75, X78, X79, and X80. The Specific SNS Indicators X1, X3, X5, X6, X7, X8, X26, X15, X56, X88, and X62 loaded on to a second factor. Much as was the case with the Conscientiousness EFA, however, these indicators seemed to be tapping another, somewhat-related construct (e.g., integrity) because they dealt with more criminal and inappropriate behaviours than did the first factor. Also, the second factor was only weakly related to the first factor ($r = .075$, $p < .05$, between these two factors in the EFA). Therefore, X1, X3, X5, X6, X7, X8, X26, X15, X56, X88, and X62 were removed. Because X28 was a follow-up to X27, it had to be

Table 5 (Study 1): Goodness-of-Fit Indicators for Specific SNS Indicator EFA Measurement Models

Model	CFI	χ^2	*df*	χ^2 p-value	RMSEA	SRMR
Agreeableness	-	-	-	-	-	-
Extraversion	.840	57.907	9	.000	.196	.088
Conscientiousness	.991	16.897	9	.050	.079	.026
Openness	1.000	.000	0	.000	.000	.000
Neuroticism	-	-	-	-	-	-
Honesty-Humility	-	-	-	-	-	-
Impression Management	.902	79.674	9	.000	.236	.037
Verbal Ability	1.000	.000	0	.000	.000	.000
Quantitative Ability	-	-	-	-	-	-
Integrity	.738	167.699	9	.000	.354	.145

Note. N = 141. CFI = comparative fit index, χ^2 = chi-squared, *df* = degrees of freedom, χ^2 p-value = significance value of chi-squared, RMSEA = Root Mean Square Approximation, SRMR = standardized root mean square residual. When *df* = 0, the fit is always perfect, so these results cannot be interpreted. Some results are blank (represented by a hyphen) because there were not enough Specific SNS Indicators to create adequately identified EFA measurement models for these attributes.

removed as well. The third factor consisted of X2 and X4, which are the follow-up indicators for X1 and X3 (respectively), which had already been removed. Therefore, they were removed as well. X21 and X81 were also removed: they failed to load significantly on any of the three factors. Finally, X12 was removed because its factor loading (.273) was less than .320. Only X10, X46, X75, X78, X79, and X80 remained, resulting in adequate fit, $\chi^2(9) = 79.674$, $p = .000$, CFI = .902, RMSEA = .236, SRMR = .037 (see Table 5). Geomin rotated factor loadings can be found in Table 6. The Cronbach's alpha reliability for the Impression Management SNS scale was quite high at .92, supporting its unidimensionality.

Verbal Ability. The indicators for the Verbal Ability EFA were X41, X45, and X47. However, because there were only three indicators, the model was just-identified, so fit could not be assessed. However, the EFA parallel analysis pointed to a one-factor solution, and all three factor loadings were significant and greater than .32. Also, the high internal consistency reliability ($\alpha = .89$) supported its unidimensionality. Therefore, all three items were retained.

Quantitative Ability. The Taxonomy Expert Raters were not able to identify any Specific SNS Indicators that might be linked to Quantitative Ability. Therefore, no model could be created for Quantitative Ability.

Integrity. It was theoretically difficult to justify whether each integrity Specific SNS Indicator would better measure Overt Integrity or Personality-Based Integrity. Therefore, I decided to only create a single SNS Integrity scale. Later, in the Combined CFA model section, this single SNS Integrity scale is correlated with the traditional Overt Integrity and Personality-Based Integrity scales.

The SNS Integrity scale consisted of X1, X2, X3, X4, X5, X6, X15, X21, X26, X27, and X81. The EFA parallel analysis pointed to a three-factor solution. However, the only model to converge was the one-factor solution, albeit with poor fit, $\chi^2(55) = 789.639$, $p < .01$, CFI = .608, RMSEA = .216, SRMR = .153. The indicators X2, X4, X21, X27, and X81 failed to have significant factor loadings, and were removed, leaving X1, X3, X5, X6, X15, and X26. The resulting model had poor fit and should be interpreted with caution, $\chi^2(9) = 167.699$, p = .000, CFI = .738, RMSEA = .354, SRMR = .145 (see Table 5). Geomin rotated factor loadings can be found in Table 6. The Integrity SNS scale had an internal consistency reliability of .81, supporting its unidimensionality.

Measurement model for Global SNS Rating. Finally, the three items that served as the Global SNS Rating of participants' SNSs were very strongly correlated (Ov1 and Ov2 correlated .870, $p < .05$; Ov2 and Ov3 correlated .897, $p < .05$; Ov1 and Ov3 correlated .734, $p < .05$). In addition, the internal consistency reliability was very high (.93) for all three items, therefore, the three ratings were aggregated into a single item, referred to as the Global SNS Rating.

Confirmatory factor analyses: Combined CFA Models. The first major goal of the current study was to assess the structural validity of the Specific SNS Indicators and the Global SNS Rating of participants' SNS profiles. I wanted to examine whether the Specific SNS Indicators were empirically representing the construct/attribute domains in the way that the Taxonomy Expert Raters had predicted, how the Global SNS Rating related to the Specific SNS Indicators, and the unidimensionality of the Specific SNS Indicators. In order to do this, Combined Model CFAs were create to see whether the models had adequate fit; as initial support of their empirical representation of the construct/attribute domains. The

Combined Model CFAs consisted of: Specific SNS Indicators loading on a Specific SNS Indicator scale latent variable, facets of traditional scales loading on a traditional scale latent variable, and a Global SNS Rating input as a directly observed variable. The two latent variables and one directly observed variable were all correlated with one another. The Combined Model CFAs were conducted separately for each attribute/construct because the sample size was not large enough to test the models for all construct domains at once. As an example, see Figure 2 for a diagram of the Combined CFA model results for Conscientiousness (see Appendix A for remaining diagrams).

H_2 predicted that the empirical dimensionality of the Specific SNS Indicators would reflect the theoretical dimensionality as informed by the judgment of expert raters. This was tested through the creation of the Combined CFA Models and whether these models demonstrated adequate fit. H_2 was partially supported, as the final fit statistics for the Combined CFA models had adequate fit or approached adequate fit, with two exceptions (see Table 8). For the Integrity model that consisted of the single SNS Integrity scale correlated with the traditional Overt Integrity scale and the Global SNS Rating, the model had a worse fit than the other models (see Overt Integrity Model, Table 8), $\chi^2(42) = 209.530$, $p < .01$, CFI = .773, RMSEA = .168, SRMR = .101. Also, the Integrity model with the single SNS Integrity scale correlated with the traditional Personality-Based Integrity scale and the Global SNS Rating had lower fit indices as well, $\chi^2(42) = 194.969$, $p < .01$, CFI = .762, RMSEA = .161, SRMR = .097.

2.4.4 The Evaluation of Construct Validity

The second major goal of the current study was to examine the Specific SNS Indicator scales' construct validity, including both convergent and discriminant validity. To

assess the convergent validity, the data was inserted into SPSS (Version 20.0) and the correlations between the Specific SNS Indicator scales and the traditional scales were investigated. In addition, the Global SNS Rating was included to examine its correlations with traditional scales and the Specific SNS Indicators. Finally, the construct validity was further tested in another step by examining the unique variance that the Specific SNS Indicator scales had beyond the Global SNS Rating. The Specific SNS Indicator scales' discriminant validity was assessed by examining whether the Specific SNS Indicator scales appeared more related to their respective traditional scales than to measures of response distortion.

Assessing Convergent Validity. To test H_{3a}, the correlations between the Specific SNS Indicator scales and the traditional selection scales were examined (see Table 9). It was predicted that Specific SNS Indicator scales rated as tapping performance-related human attributes would have significant correlations with their respective traditional selection scales, and these correlations would be at least weak ($r > .10$) in magnitude. As partial support for H_{3a}, three of the Specific SNS Indicator scales in Table 9 had significant, small to moderate relations with their equivalent traditional selection scales (Extraversion, $r = .376$, $p < .05$; Openness to Experience, $r = .196$, $p < .05$; Verbal Ability, $r = .223$, $p < .05$). The exceptions were: Agreeableness ($r = .131$, *n.s.*), Conscientiousness, ($r = .084$, *n.s.*), Honesty-Humility ($r = .031$, *n.s.*), and Impression Management ($r = -.113$, *n.s.*). Also, the single SNS Integrity scale was not significantly correlated with traditional scales of Overt Integrity ($r = -.028$, *n.s.*), and Personality-Based Integrity ($r = .074$, *n.s.*).

Zero-order correlations were also computed for the relations between the traditional scales and the Global SNS Rating (see Table 9). Only Extraversion ($r = .229$, $p < .05$) and

Table 6 (Study 1): Factor Loadings for SNS Indicator EFA Measurement Models

Item	Geomin Rotated Factor Loadings
Agreeableness	
X13 and X57 Composite	-
Extraversion	
X37	.96* (.03)
X67	.79* (.04)
X72	.68* (.05)
X74	.40* (.08)
Ob18	.32* (.08)
Ob27	.44* (.08)
Conscientiousness	
X80	.49* (.07)
X42	.97* (.01)
X43	.97* (.01)
X44	.86* (.02)
X47	-.41* (.07)
X50	.96* (.01)
Openness	
X39	.84* (.10)
X73	.62* (.09)
X74	.47* (.09)
Honesty-Humility	
X1 and X26 Composite	-
Impression Management	
X75	.90* (.02)
X79	.85* (.03)
X10	.85* (.03)
X46	.88* (.03)
X78	.90* (.02)
X80	.48* (.07)
Verbal Ability	
X41	.94* (.03)

(continued)

Table 6 (Study 1): Factor Loadings for SNS Indicator EFA Measurement Models

Item	Geomin Rotated Factor Loadings
X45	.92* (.03)
X47	-.71* (.05)
Integrity	
X1	.96* (.01)
X3	1.00* (.01)
X5	.43* (.07)
X15	.37* (.07)
X26	.44* (.07)
X6	.43* (.07)

Note. $N = 141$. Standard errors are in parentheses.
*$p < .05$

Table 7 (Study 1): Descriptives for Traditional Scales and Specific SNS Indicator Scales

Scale	Min	Max	*M*	SD	α	ICC (2,k)
Traditional Scales						
Agreeableness	2.67	4.58	3.62	.37	.82	-
Extraversion	2.13	4.63	3.49	.48	.88	-
Conscientiousness	2.17	4.46	3.34	.46	.87	-
Openness	2.33	4.63	3.32	.44	.85	-
Neuroticism	1.58	4.63	2.88	.57	.90	-
Honesty-Humility	1.88	4.50	3.09	.48	.76	-
Impression Management	1.47	3.89	2.77	.43	.76	-
Verbal Ability	0.00	.80	.41	.16	-	-
Quantitative Ability	2.00	.86	.36	.14	-	-
Overt Integrity	2.52	4.10	3.32	.35	.90	
Personality-Based Integrity	2.53	3.80	3.10	.25	.77	-
Specific SNS Indicator Scales						
Agreeableness	18.14	97.38	43.25	10.92	.65	.59
Extraversion	17.54	75.12	50.01	12.13	.73	.75
Conscientiousness	35.83	76.00	54.50	7.72	.81	.67
Openness	20.00	73.22	50.60	9.86	.67	.71
Neuroticism	-	-	-	-	-	.34
Honesty-Humility	9.33	100.00	92.78	12.33	.60	.38
Impression Management	26.39	75.83	52.64	10.74	.92	.64
Verbal Ability	26.67	63.33	47.72	3.88	.89	.52
Quantitative Ability	-	-	-	-	-	-
Integrity	32.45	100.00	88.38	11.40	.81	.45

Note. N = 141. Min = Minimum, Max = Maximum, M = Mean, SD = Standard Deviation, α = Cronbach's Alpha or Spearman-Brown reliability (for Agreeableness and Honesty-Humility Specific SNS Indicator scales), ICC(2,k) = average ICC(2,k) interrater reliability for the final scale. Only Specific SNS Indicator scales were rated, which is why the traditional scales do not have values for ICC(2,k). Neuroticism and Quantitative Ability Specific SNS Indicator scales were dropped from the study, which is why their data is not represented in some cases.

Openness (r = .187, p < .05) were significantly correlated with the Global SNS Rating. Agreeableness (r =.048 , *n.s.*), Conscientiousness (r = -.024 , *n.s.*), Neuroticism (r = -.097 , *n.s.*), Honesty-Humility (r = -.034 , *n.s.*), Impression Management (r = -.122 , *n.s.*), Verbal Ability (r = .019 , *n.s.*), Quantitative Ability (r = -.126 , *n.s.*), Overt Integrity (r = .029 , *n.s.*), and Personality-Based Integrity (r = -.046, *n.s.*) were not significantly correlated with the Global SNS Rating.

Finally, the following Specific SNS Indicator scales were all significantly correlated with the Global SNS Rating: Agreeableness (r = .333, p < .05), Extraversion (r = .362, p < .05), Conscientiousness (r = .746, p < .05), Openness (r = .472, p < .05), Impression Management (r = .772, p < .05), and Verbal Ability (r = .371, p < .05; see Table 9).

For H_{3b}, the relations between Specific SNS Indicator scales and their respective traditional scales were further examined by investigating the unique variance that the Specific SNS Indicator scales predicted beyond the Global SNS Rating. The results can be found in Table 10. Specific SNS Indicator scales demonstrated incremental prediction beyond the Global SNS Rating for two traits: Extraversion (ΔR^2 = .097, p < .05), and Verbal Ability (ΔR^2 = .054, p < .05).

To see whether the opposite were true, whether the Global SNS Rating accounted for incremental prediction beyond the Specific SNS Indicator scales, a second set of hierarchical regressions was also conducted in which the order of variables was reversed (see Table 11). The Global SNS Rating did not account for statistically significant incremental prediction beyond any Specific SNS Indicator scales.

It is worth pointing out that, on its own, the Global SNS Rating was not a good

Table 8 (Study 1): Goodness-of-Fit Indicators for the Combined CFA Models

Model	CFI	χ^2	*df*	χ^2 p-value	RMSEA	SRMR
Agreeableness	.943	26.106	19	.128	.052	.051
Extraversion	.860	138.877	63	.000	.092	.088
Conscientiousness	.940	135.805	63	.000	.091	.070
Openness	.884	62.644	33	.001	.080	.063
Neuroticism	-	-	-	-	-	-
Honesty-Humility	.802	16.043	8	.042	.084	.058
Impression Management	.891	147.814	42	.000	.134	.057
Verbal Ability	.972	14.336	4	.006	.135	.027
Quantitative Ability	-	-	-	-	-	-
Overt Integrity	.773	209.530	42	.000	.168	.101
Personality-Based Integrity	.762	194.969	42	.000	.161	.097

Note. N = 141. CFI = comparative fit index, χ^2 = chi-squared, *df* = degrees of freedom, χ^2 p-value = significance value of chi-squared, RMSEA = Root Mean Square Approximation, SRMR = standardized root mean square residual. Neuroticism and Quantitative Ability Specific SNS Indicator scales were dropped from the study, which is why their data is not represented. The Overt Integrity is referred to as such because it had the SNS Integrity scale correlated with the traditional Overt Integrity scale and the Global SNS Rating. The Personality-Based Integrity is referred to as such because it had the SNS Integrity scale correlated with the traditional Personality-Based Integrity scale and the Global SNS Rating.

predictor of the traditional scales. The Global SNS Rating only significantly predicted 1 out of 9 traditional scales. It significantly predicted the Openness to Experience traditional scale (3.5% of the variance explained; see Table 10). The Specific SNS Indicator scales, however, significantly predicted 3 out of 9 traditional scales. The Specific SNS Indicator scales significantly predicted the traditional scales of Extraversion (14.1% of variance explained), Openness to Experience (3.8% of variance explained), and Verbal Ability (5.0% of variance explained; see Table 11). Although this is an improvement over the predictions by the Global SNS Rating, 3 out of 9 significant predictions is still fairly low.

Assessing Discriminant Validity. None of the Specific SNS Indicator scale correlations with Impression Management were significant, which helps support the discriminant validity of the Specific SNS Indicator scales. However, H_{4a} is a more stringent test of discriminant validity as it assesses whether the Specific SNS Indicator scales will be more strongly correlated with their respective traditional scales than with the Impression Management scale. H_{4a} was almost fully supported because the correlations between the Specific SNS Indicator scales and Impression Management appeared to be weaker than those between the Specific SNS Indicator scales and their respective traditional scales, with two exceptions (see Table 9). The correlation between the Specific SNS Indicator scale for Integrity and its traditional Overt Integrity scale, r = -.028, appeared weaker than the correlation between the Specific SNS Indicator scale for Integrity and Impression Management, r = .064 (see Table 9). The same was the case for the Honesty-Humility Specific SNS Indicator scale, whose correlation with its traditional Honesty-Humility scale appeared weaker, r = .031, than its correlation with Impression Management, r = .075.

Table 9 (Study 1): Zero-Order Relations Between SNS Indicator Scales, Traditional Selection Scales, and Global SNS Rating

Attribute	Correlations								
	SNS with Trad. Scale	Impression Management		Self-Deception		PATD		Global SNS Rating	
		Trad. Scale	SNS Scale	Trad. Scale	SNS Scale	Trad. Scale	SNS Scale	Trad. Scale	SNS Scale
AG	.131	.484*	.089	.099	.039	-.424*	-.175	.048	.333*
EX	.376*	.000	-.094	.357*	.043	.037	.045	.229*	.362*
CON	.084	.433*	-.070	.440*	.142	-.377*	.097	-.024	.746*
OPEN	.196*	.130	-.013	.169	.044	.108	.067	.187*	.472*
NEUR	-	-.046	-	-.706*	-	-.040	-	-.097	-
HH	.031	.627*	.075	.084	.042	-.476*	.029	-.034	.086
IM	-.113	-	-.113	.170	.115	-.377*	.073	-.122	.772*
VA	.223*	.043	-.015	.003	.100	.090	.184	.019	.371*
QA	-	.108	-	.028	-	.061	-	-.126	-
OI	-.028	.513*	.064	.140	.068	-.447*	.032	.029	.085
PBI	.074	.444*	.064	.284*	.068	-.315*	.032	.046	.085
GLB	-	-	-.122	-	.071	-	.169	-	-

Note. N = 141. SNS = Social Networking Sites, AG = Agreeableness, EX = Extraversion, CON = Conscientiousness, OPEN = Openness to Experience, NEUR = Neuroticism, HH = Honesty-Humility, IM = Impression Management, VA = Verbal Ability, QA = Quantitative Ability, OI = Overt Integrity, PBI = Personality- Based Integrity, GLB = Global SNS Rating, PATD = Perceived Ability to Deceive, Trad. Scale = traditional scale. For Overt Integrity and Personality-Based Integrity, these scales were each correlated with the single SNS Integrity scale.

*p <.05

Table 10 (Study 1): Hierarchical Regression with Traditional Scales as Dependent Variable: Global SNS Rating Entered First and Specific SNS Indicator Scales Entered Second

Step	Variable	*b*	*β*	R^2	ΔR^2_{SNS} [a]
1.	Global SNS Rating	.001	.048	.002	
2.	Global SNS Rating	.000	.000	.017	.015
	Agreeableness SNS	.004	.131		
1.	Global SNS Rating	.009	.229	.052	
2.	Global SNS Rating	.004	.099	.150*	.097*
	Extraversion SNS	.013*	.338*		
1.	Global SNS Rating	-.001	-.024	.001	
2.	Global SNS Rating	-.008	-.204	.025	.024
	Conscientiousness SNS	.014	.238*		
1.	Global SNS Rating	.007*	.187*	.035*	
2.	Global SNS Rating	.004	.121	.050	.015
	Openness SNS	.006	.139		
1.	Global SNS Rating	-.001	-.034	.001	
2.	Global SNS Rating	.001	-.037	.002	.001
	Honesty-Humility SNS	.001	.034		
1.	Global SNS Rating	-.004	-.122	.015	
2.	Global SNS Rating	-.003	-.087	.016	.001
	Impression Management SNS	-.002	-.044		
1.	Global SNS Rating	.010	.019	.000	
2.	Global SNS Rating	-.038	-.075	.054*	.054*
	Verbal Ability SNS	.367*	.251*		
1.	Global SNS Rating	.049	.029	.001	
2.	Global SNS Rating	.052	.030	.002	.001
	Overt Integrity	-.054	-.030		
1.	Global SNS Rating	-.052	-.046	.002	
2.	Global SNS Rating	-.056	-.050	.008	.006
	Personality-Based Integrity	.091	.076		

Note. $N = 141$. *b* = unstandardized regression coefficient, *β* = standardized regression coefficient, R^2 = Total variance predicted considering all variables in the equation, ΔR^2 = Change in R^2 value. Overt and Personality-Based Integrity represent the same single SNS Integrity scale, however, they have different titles to signify their different dependent variables (Overt Integrity and Personality-Based Integrity), which is why the title does not have SNS at the end.

[a] Change in R^2 caused by adding the SNS scale to the equation.

*$p < .05$.

Table 11 (Study 1): Hierarchical Regression with Traditional Scales as Dependent Variable: Specific SNS Indicator Scales Entered First and Global SNS Rating Entered Second

Step	Variable	*b*	*β*	R^2	ΔR^2_{Ov} [b]
1.	Agreeableness SNS	.004	.131	.017	
2.	Agreeableness SNS	.004	.131	.017	.000
	Global SNS Rating	.000	.000		
1.	Extraversion SNS	.015*	.376*	.141*	
2.	Extraversion SNS	.013*	.338*	.150*	.000
	Global SNS Rating	.004	.099		
1.	Conscientiousness SNS	.005	.084	.007	
2.	Conscientiousness SNS	.014	.238	.025	.018
	Global SNS Rating	-.008	-.204		
1.	Openness SNS	.009*	.196*	.038*	
2.	Openness SNS	.006	.139	.050	.011
	Global SNS Rating	.004	.121		
1.	Honesty-Humility SNS	.001	.031	.001	
2.	Honesty-Humility SNS	.001	.034	.002	.001
	Global SNS Rating	-.001	-.037		
1.	Impression Management SNS	-.004	-.113	.013	
2.	Impression Management SNS	-.002	-.044	.016	.003
	Global SNS Rating	-.003	-.087		
1.	Verbal Ability SNS	.326*	.223*	.050*	
2.	Verbal Ability SNS	.367	.251	.054	.005
	Global SNS Rating	-.038	-.075		
1.	Overt Integrity	-.051	-.028	.001	
2.	Overt Integrity	-.054	-.030	.002	.001
	Global SNS Rating	.052	.030		
1.	Personality-Based Integrity	.088	.074	.005	
2.	Personality-Based Integrity	.091	.076	.008	.002
	Global SNS Rating	-.056	-.050		

Note. $N = 141$. *b* = unstandardized regression coefficient, *β* = standardized regression coefficient, R^2 = Total variance predicted considering all variables in the equation, ΔR^2 = Change in R^2 value. Overt and Personality-Based Integrity represent the same single SNS Integrity scale, however, they have different titles to signify their different dependent variables (Overt Integrity and Personality-Based Integrity), which is why the title does not have SNS at the end.

[b] Change in R^2 caused by adding the Global SNS Rating to the equation.

*$p < .05$.

The Specific SNS Indicator scales were also not significantly correlated with the Self-Deception scale. In partial support of H_{4b}, the Specific SNS Indicator scales appeared to be more strongly correlated with their respective traditional scales than with a scale of Self-Deception, with four exceptions (see Table 9). The correlation between the Specific SNS Indicator scale for Integrity and the traditional Overt Integrity scale, r = -.028, appeared weaker than the correlation between the Specific SNS Indicator scale for Integrity and Self-Deception, r = .068. The same was the case for the Honesty-Humility Specific SNS Indicator scale, whose correlation with its traditional Honesty-Humility scale appeared weaker, r = .031, than its correlation with Self-Deception, r = .042. The correlation between the Impression Management Specific SNS Indicator scale and the traditional Impression Management scale appeared weaker, r = -.113, than the correlation between the Impression Management Specific SNS Indicator and Self-Deception, r = .115). Finally, the correlation between the Conscientiousness Specific SNS Indicator scale and the traditional Conscientiousness scale appeared weaker, r = .084, than the correlation between the Conscientiousness Specific SNS Indicator scale and Self-Deception, r = .142.

Finally, the Specific SNS Indicator scales were not significantly correlated with the PATD scale. H_{4c} was partially supported because the Specific SNS Indicator scales appeared to be more strongly correlated with their respective traditional scales than with the PATD scale, with three exceptions (see Table 9). The correlation between the Integrity Specific SNS Indicator scale and its traditional Overt Integrity scale, r = -.028, appeared weaker than the correlation between the Integrity Specific SNS Indicator scale and PATD, r = .032. Also, the correlation between the Agreeableness Specific SNS Indicator scale and its traditional Agreeableness scale appeared weaker, r = .131, than the correlation between the

Agreeableness Specific SNS Indicator scale and PATD, *r* = -.175. However, the relation with PATD is negative, which could mean that individuals scoring higher on Agreeableness are less inclined to distort their responses: as supported by their lower scores on PATD. Finally, the correlation between the Conscientiousness Specific SNS Indicator scale and its traditional Conscientiousness scale, *r* = .084, appeared weaker than the correlation between the Conscientiousness Specific SNS Indicator scale and PATD, *r* = .097.

In summary, based on the above findings, there was supporting evidence for the discriminant validity of the Specific SNS Indicators. Nine out of 11 Specific SNS Indicator scales appeared more highly correlated with their respective traditional scales than with Impression Management. Seven out of 11 Specific SNS Indicator scales appeared more highly correlated with their respective traditional scales than with Self-Deception. Eight out of 11 Specific SNS Indicator scales appeared more highly correlated with their respective traditional scales than with PATD.

The Global SNS Rating also demonstrated some evidence of discriminant validity, as it was not significantly correlated with Impression Management, Self-Deception, or PATD (see Table 9). However, the discriminant validity cannot be as fully tested as it was with the Specific SNS Indicators because there was obviously no traditional global rating with which to compare these results.

2.5 Discussion

There were several interesting, and in some cases unexpected, findings in the current study, as well as some problems that were encountered. Interrater reliabilities were substantially higher for Specific SNS Indicators using a RPM format than those using a count format. Only two Objective Specific SNS Indicators were retained as part of the Specific

SNS Indicator scales. The reliabilities for the Specific SNS Indicator scales were fairly low. There was some evidence of structural validity for the scales. There was also some evidence of convergent validity for the scales, although the magnitude of these correlations was small to moderate. There was supporting evidence for the discriminant validity of the Specific SNS Indicator scales. Surprisingly, they appeared to be less related to scales of response distortion than were their respective traditional scales. There may also be legal advantages to using Specific SNS Indicators over a Global SNS Rating, however, both had fairly minimal reliability and validity evidence. One of the problems encountered in the study was that Taxonomy Expert Raters could not identify enough Specific SNS Indicators in order to create two of the Specific SNS Indicator Scales: Neuroticism and Quantitative Ability. Also, only a small percentage of Specific SNS Indicators were retained in order to create the Specific SNS Indicator scales. The exploration of these findings begins with the types of data that were collected in the study.

2.5.1 Data Collection Formats

For the current study, there were three different types of data that were collected. The RPM format asked raters to assign a percentile to a ratee's behaviour, with 50 representing the average for Facebook users. In addition to being used for Specific SNS Indicators, the RPM format was also used for the Global SNS Rating indicators. The count format was also used for the Specific SNS Indicators. It consisted of frequency information, such as how often a candidate would refer to alcohol or drug use. Finally, a categorical format was also used. Raters would assign a number if certain criteria were met. For example, If the ratee had no birth date information on the "About" page, they would be given a "0", and if they had the date but not the year, they would be given a "1", and if they had both the date and year, they would be given a "2". Each of these formats will be discussed in turn.

Although overall the interrater reliabilities for the Specific SNS Indicators were low, it was interesting that the interrater reliabilities were substantially higher (average = .60) for Specific SNS Indicators that relied on the RPM format than those that relied on the count format (average = .35). This was unexpected, as the count format relied on counting the frequency of a given behaviour, whereas the indicators using the RPM format involved judging a given behaviour in comparison to other individuals. It was expected that the count indicators would be more objective in nature, relying on a more mechanical approach to data collection (see Cascio & Aguinis, 2005). However, it appears that rating count data on a SNS may be a more subjective task than previously thought, and may have involved a more judgmental approach to data collection (see Cascio & Aguinis, 2005) than anticipated. For example, one indicator asked "number of negative statements about former/current employer(s), work colleagues, clients, supervisors, classmates, or teachers/professors". The mental cutoff required to decide what constitutes a negative statement could be quite subjective. This was corroborated when speaking with the SNS Expert Raters after they had completed their ratings. The mental cutoff may have been different for the raters, which could have led to a lack of consistency between the raters. For the RPM format indicators, however, a mental cutoff was not required. Rather than deciding what constitutes a negative statement, raters just needed to rate how negative that statement was. For example, an indicator using the RPM format could look like "how negative are the statements that s/he makes about former/current employer(s), work colleagues, clients, supervisors, classmates, or teachers/professors". Presumably, rating this indicator is an easier task, which could explain the higher consistency that was found between raters for RPM format indicators. These findings highlighted the fact that what seems like a mechanical data collection method may in fact not be. The data collection of count format Specific SNS Indicators may in fact be

more judgmental in nature. The findings also illustrated the value of using a RPM format for assessing judgmental data (Goffin et al., 1996; Goffin et al., 2009; Goffin & Olson, 2011).

The interrater reliabilities for the objective Specific SNS Indicators that made use of the categorical format could not be calculated because these indicators were rated by a lone rater. In addition, they did not meet the cutoff criteria for being conceptually related to the human attributes assessed by the traditional selection scales (see Preliminary Evaluation of the Specific SNS Indicators section). Therefore, they were not included in any of the Specific SNS Indicator scales.

These findings point to the potential benefits of using a judgmental format for SNS data collection, such as the one used for the RPM format Specific SNS Indicators, as opposed to other formats (e.g., count and categorical).

2.5.2 Objective Indicators

As mentioned earlier, objective indicators were rated by a sole rater, as they were considered less subjective than the other Specific SNS Indicators. Only two of the 27 objective indicators (only Ob18 and Ob27) had loadings that were acceptable for retention in the CFA models. The lack of objective indicators with acceptable loadings may have been due to criterion contamination. The term criterion contamination is borrowed from the job performance literature. Criterion contamination occurs when aspects of the actual or operational criterion are unrelated to the conceptual criterion (Cascio & Aguinis, 2005). The actual or operational criterion is what is being measured. The conceptual criterion is the construct that one wants to assess in the first place. As an example, consider the following Specific SNS Indicator to be an actual criterion, while Extraversion is the conceptual criterion. The Taxonomy Expert Raters predicted that the objective Specific SNS Indicator

"number of photos or videos of the candidate participating in group activities (including team sports) in photos albums or posts" would be related to Extraversion. However, this Specific SNS Indicator loaded only weakly on the Extraversion SNS latent variable in the EFA measurement model. The Specific SNS Indicator could be unrelated to the conceptual criteria, in this case, Extraversion, which may account for the poor loading. For instance, perhaps the participant has a number of photos at a computer gaming party because s/he is passionate about computer gaming, and not because s/he is higher on Extraversion. The participant could have had the same score as another participant who prefers playing multiplayer computer games and who is higher on Extraversion.

In order to remain objective, the objective indicator was collected through mechanical data collection methods, defined previously in the "A New Way to Assess SNS Information" section. The objective indicator relied on frequency/count information so that it could be easily counted. However, if the indicator were allowed to be more subjective in nature by relying on judgment data collection methods, rather than relying purely on mechanical data collection, this could have allowed for more discrimination between participants. For instance, using the RPM format, an item could have been created such as "participates in group activities (including team sports) in photo albums or posts". RPM format Specific SNS Indicators such as this may be less susceptible to criterion contamination because the candidate's photos and posts can be judged based on more conceptual criterion-related elements. For instance, perhaps a rater considers a participant who attends a computer gaming party to be less involved in group activities than a participant who is involved in soccer. A rater may consider soccer, which requires players to work together, to be more of a group activity. Whereas a computer gaming party may be considered to be less of a group activity if the individual is simply there because s/he is an enthusiast of single-player

computer games. Assuming they both had the same number of photos or posts relating to their respective group activities, if using the count format to rate these two individuals, they would both have the same score. However, if using the RPM format, their scores could be different. In this example, the individual playing soccer would likely be higher on Extraversion than the individual playing the single-player computer games, because soccer can be considered more of a group activity than playing single-player computer games with others. Therefore, both of these individuals would be judged based on more conceptual criterion-related elements, namely Extraversion. Perhaps objective Specific SNS indicators are, in general, less related to their conceptual criterion than are Specific SNS Indicators that use a RPM format.

2.5.3 Evidence of Reliability and Validity

Interrater Reliability. The interrater reliabilities met the bare minimum requirements of being significantly different from zero. The initial average ICC(2,*k*) of .53 across all indicators appears low (see Table 1 for all interrater reliabilities). The indicators were, however, used as part of Specific SNS Indicator scales consisting of multiple indicators, which resulted in slightly higher reliability. Also, the interrater reliability would be higher if only Specific SNS Indicators that relied on the RPM format were used (average for RPM format = .60, average for count format = .35). However, it is worth pointing out that these reliabilities are still fairly low. As for the Global SNS Rating indicators, they had the best interrater reliabilities (average = .66), although these were also still fairly low. Their slightly higher reliabilities may be because global ratings elicit higher interrater reliabilities in general: a number of studies in clinical settings have demonstrated higher interrater reliabilities for global ratings than checklists or other forms of ratings (Hodge, Regehr,

Hanson, & McNaughton, 1997; Malau-Aduli et al., 2012; Martin et al., 1997; Regehr, MacRae, Reznick, & Szalay, 1998).

It is important to remind the reader, however, that the current study used three SNS Expert Raters. ICC(2,k) reflects the interrater reliabilities if the ratings given by the three SNS Expert Raters were averaged together. ICC(2,1) reflects the interrater reliability if the ratings were based on a single rater. Moving from the ICC(2,*k*) column to the ICC(2,1) column in Table 1, one can see how much lower the predicted interrater reliability would be if a single rater were rating an individual's SNS account. The average for the ICC(2,*k*) reliabilities was .53 and the average for the ICC(2,1) reliabilities was .32. The average ICC(2,1) reliability appears to be quite low. If there were only one HR staff member conducting the ratings, this could be a concern. Ratings would be likely to vary depending on who is conducting the ratings. This would not be ideal in a selection setting.

Internal Consistency Reliability. The internal consistency reliabilities for the Specific SNS Indicator scales[1] were acceptable, and ranged from .67 to .92 (see Table 7). This range was fairly similar to that of the more established, traditional scales, which ranged from .76 to .90 (see Table 7). The Spearman-Brown formula was used for Agreeableness (reliability = .65) and Honesty-Humility (reliability = .60), because they only had two Specific SNS Indicators per scale. These scales had minimally acceptable reliability (see Clark & Wilson, 1995).

[1] These were based on ratings of all three SNS Expert Raters at the indicator level. If just one rater were used, the reliabilities are generally lower. This applies to the internal consistency reliabilities as well as the Spearman-Brown reliabilities.

Structural Validity. There was some supporting evidence of the structural validity of the Specific SNS Indicator scales. The Combined CFA models had adequate fit or approached adequate fit for all but the two models using the SNS Integrity scale (see Table 8). For the model with the Specific SNS Integrity scale correlating with the traditional Overt Integrity scale and the Global SNS Rating, the fit was poor, $\chi^2(42) = 209.530$, $p < .01$, CFI = .773, RMSEA = .168, SRMR = .101. Also, the fit was poor for the model with the Specific SNS Integrity scale correlating with the traditional Personality-Based Integrity scale and the Global SNS Rating, $\chi^2(42) = 194.969$, $p < .01$, CFI = .762, RMSEA = .161, SRMR = .097.

The lack of model fit for both models could be due to using a general SNS Integrity construct in both models as opposed to more specific SNS subdimensions such as a Specific SNS Personality-Based Integrity scale and a Specific SNS Overt Integrity scale. As mentioned earlier, unfortunately, these more specific Integrity scales could not be created, because it was theoretically difficult to justify which Specific SNS Indicators belonged to each. Using a more general SNS Integrity construct in both of these models could have resulted in a lack of content overlap (Christiansen & Tett, 2013). Content overlap relies on matching constructs based on thematic linkages (Christiansen & Tett, 2013). In the current example, the content of the general SNS Integrity scale may not be matched adequately with the content of each Integrity subdimension on its own. The thematic content for Integrity would be more encompassing than the thematic content would be for Overt Integrity. The same would apply to Personality-Based Integrity. With more research, perhaps Specific SNS Indicators could be created specifically to generate both an Overt Integrity Specific SNS scale and a Personality-Based Integrity Specific SNS scale. If these two Specific SNS Integrity scales could be created, the resulting content overlap with the respective traditional

Overt and Personality-Based Integrity scales could lead to better model fit and the potential for better structural validity.

Convergent Validity. To test convergent validity, the correlation between the Specific SNS Indicator scales and the traditional scales were examined. The current study found partial support for the convergent validity of the Specific SNS Indicator scales. Support was found for Extraversion, Openness to Experience, and Verbal Ability. Back et al. (2010) found similar convergent validity findings when they studied the accuracy of FFM SNS ratings. They found a moderate, significant correlation between observer SNS ratings of Extraversion and self-rated Extraversion from the NEO Five-Factor Inventory (r = .39, p < .05; Costa & McCrae, 1992). In the current study, the correlation between the Extraversion Specific SNS scale and the traditional Extraversion scale was also significant, and of a very similar magnitude (r = .38, p < .05). Back et al. found a moderate, significant correlation between observer SNS ratings of Openness to Experience and self-rated Openness to Experience (r = .41, p < .05). In the current study, the correlation between the Specific SNS Indicator scale for Openness to Experience and the traditional Openness to Experience scale was also significant (r = .20, p < .05), but smaller than the correlation found in Back et al. Back et al. found a small, significant correlation between observer SNS ratings of Agreeableness and self-rated Agreeableness (r = .22, p < .05). In the current study, there was no significant correlation between the Agreeableness Specific SNS Indicator scale and the traditional Agreeableness scale. Also, Back et al. found a small, significant correlation between observer SNS ratings of Conscientiousness and self-rated Conscientiousness (r = .27, p < .05). However, there was no significant correlation between the Conscientiousness Specific SNS Indicator scale and the traditional Conscientiousness scale in the current study.

No significant correlations were found for observer ratings of Neuroticism in either the current study nor in Back et al. Finally, the other attributes assessed in the current study could not be compared because they were not assessed in Back et al. These attributes were: general mental ability (Verbal and Quantitative Ability), Impression Management, Honesty-Humility, and Integrity. The Verbal Ability Specific SNS Indicator scale was significantly related to the traditional Verbal Ability scale ($r = .223$, $p < .05$). However, the other Specific SNS Indicator scales were not significantly related to their respective traditional scales.

It was not surprising that there was no supporting evidence for Neuroticism, as researchers have suggested that it is difficult to assess the less the rater knows the ratee (Back et al, 2010; Funder, 1999; Kenny, 1994). In the current study, the raters had no previous experience with the ratees. It was, however, surprising that there was no supporting evidence for the convergent validity of Agreeableness and Conscientiousness in the current study. Perhaps the observation of Agreeableness and Conscientiousness traits on one's SNS profile is more difficult than it is for the other FFM traits, such as Extraversion and Openness to Experience. Maybe Extraversion and Openness to Experience are easier to rate because there is simply more Extraversion and Openness to Experience behavioural information available on one's SNS profile.

Gosling, Augustine, Vazire, Holtzman, and Gaddis (2011) offers supporting evidence for this claim. Using a sample of 159 undergraduate studies in one study and 133 in the next, they investigated how frequently participants used Facebook and what user features they used. This information was then correlated with participants' self-rated personality. They reported that those who scored high on Extraversion self-reported a higher frequency of Facebook usage ($r = .18$, $p < .05$) and they engaged more on the site (number of photos, $r =$

.28, $p < .05$; number of photo albums, $r = .20$, $p < .05$; number of wall posts, $r = .26$, $p < .05$; number of groups, $r = .28$, $p < .05$; number of friends in local network, $r = .48$, $p < .05$; total number of friends, $r = .49$, $p < .05$; number of associated networks, $r = .38$, $p < .05$). Also, those who scored high on Openness to Experience were more likely to experience new people and explore new activities (adding photos of other people without yourself, $r = .16$, $p < .05$; number of friends in local network, $r = .24$, $p < .05$; total number of friends, $r = .27$, $p < .05$; number of associated networks, $r = .22$, $p < .05$). Those who scored high on Agreeableness, however, were less likely to connect with others based on a common region or organization (number of associated networks, $r = -.18$, $p < .05$). Also, those who scored high on Conscientiousness were less likely to have spent time on Facebook ($r = -.17$, $p < .05$). Neuroticism was not related to how frequently participants used Facebook nor what user features they used.

It stands to reason that if individuals with high Conscientiousness scores are posting less SNS information, that there are less Conscientiousness behavioural cues available on SNSs than there are for other FFM traits. The higher usage of Facebook and its features by those who scored high on Extraversion and Openness to Experience may explain why they were the only two FFM constructs that evidenced convergent validity in the current study. The more a SNS is used and engaged in by those who scored high on a particular trait (e.g., Extraversion, Openness to Experience), the higher the likelihood that there is behavioural information that can be assessed for that given trait. The reverse applies as well: the less a SNS is used and engaged in by those who score high on a particular trait (e.g., Agreeableness, Conscientiousness, Neuroticism), the lower the likelihood that there is valid behavioural information that can be assessed for that given trait.

There was additional support found for the construct validity of Extraversion and Verbal Ability Specific SNS Indicator scales. These Specific SNS Indicators added unique incremental prediction beyond the Global SNS Rating, as indicated by the results of the hierarchical regressions (see Table 10). The Specific SNS Indicator scale for Extraversion accounted for 9.7% of the variance in its traditional scale beyond the Global SNS Rating, and the specific indicator scale for Verbal Ability accounted for 5.4% of the variance in its traditional scale beyond the Global SNS Rating. However, the opposite was not true, as the Global SNS Rating did not add unique incremental prediction beyond the Specific SNS Indicators (see Table 11). These findings speak to the advantages of using Specific SNS Indicators over a Global SNS Rating, at least for constructs such as Extraversion and Verbal Ability.

On its own, the Global SNS Rating did not do well at predicting the traditional selection scales, only significantly predicting the Openness to Experience traditional scale (accounted for 3.5% of the variance). Specific SNS Indicator scales on their own, however, did do slightly better by significantly predicting the respective traditional scales of Extraversion (accounted for 14.1% of the variance), Openness to Experience (accounted for 3.8% of the variance), and Verbal Ability (accounted for 5.0% of the variance).

Despite the overall weaker-than-expected validity findings for the Specific SNS Indicators, these findings lend some preliminary support to previous theory from the performance appraisal literature. This literature suggests that making the information-processing easier for raters should lead to better validity (Heneman, 1986). The raters relied on mechanical data collection methods for the Specific SNS Indicators that used the count format and judgmental data collection methods for the indicators that used the RPM format.

The data were then combined statistically by assigning and combining scores in order to form the Specific SNS Indicator scales. This is the *mechanical composite* strategy that was discussed in the introduction (Cascio & Aguinis, 2005). The Global SNS Rating, however, relied purely on judgmental data collection methods (the RPM format) and judgmental data combination methods, referred to as the *pure clinical strategy* (Cascio & Aguinis, 2005). In their head, the raters had to organize all of the data that they had observed in order to come up with their ratings and then create their global expert judgments. Plausibly, this task would be more taxing on the rater, because the incorporation of all the necessary evidence is limited by his/her cognitive capacity (Cascio & Aguinis, 2005). The benefits of using the mechanical composite strategy over the pure clinical strategy was evidenced by the slightly better incremental prediction of the Specific SNS Indicator scales over the Global SNS Rating. These findings support previous research that points to the benefits of the mechanical composite strategy over the pure clinical strategy (Einhorn, 1972; Ganzach, Kluger, & Klayman, 2000; Sawyer, 1966).

Discriminant Validity. In addition to the evidence of convergent validity, there was also support for the discriminant validity of the Specific SNS Indicator scales. The specific SNS scales were not significantly related to the response distortion scales: Impression Management, Self-Deception, and PATD (see Table 9). In addition, other than for a relatively small number of exceptions, those correlations were weaker than the correlations between the Specific SNS Indicator scales and their respective traditional scales (see Table 9). Focusing specifically on the Specific SNS Indicators that had evidence of convergent validity (Extraversion, Openness to Experience, and Verbal Ability), all of their correlations with the response distortion scales were non-significant, and were weaker than

the correlations between the Specific SNS Indicator scales and their respective traditional scales (see Table 9).

The Specific SNS Indicator scales' discriminant validity results are particularly interesting in light of the fact that 15 of the 33 correlations between traditional scales and the scales of response distortion were significant (see Table 9). Of particular note, the correlation between the traditional Honesty-Humility scale and the traditional Impression Management scale was significant and strong ($r = .627$, $p < .05$). The correlation between the traditional Overt Integrity scale and the traditional Impression Management scale was also significant and strong ($r = .513$, $p < .05$). As mentioned previously, the exact meaning of Impression Management scores has recently become controversial (Zettler et al., 2015), however these results are interesting nonetheless.

The use of the Specific SNS Indicators may be advantageous because they appear, according to the current findings, to be less related to scales of response distortion than are the traditional scales. The traditional scales were measured using self-report, which could be influenced by an "ego protection motive", whereby individuals deliberately or unconsciously distort their responses to protect their ego (Paunonen & O'Neill, 2010). The Specific SNS Indicators, however, are rated by others, and are less influenced by such a motive. Although the validity evidence for the Specific SNS Indicator scales may be too weak to warrant their use as a stand-alone selection assessment, there may be a benefit to using them to serve as a supplement or benchmark (although not necessarily a substitute) to help verify the accuracy of traditional selection scale self-ratings. For instance, by using Specific SNS Indicators as a criterion measure, a researcher or practitioner could help verify that Extraversion scores on a self-reported personality assessment are indeed more reflective of the true level of

Extraversion for the individual taking the assessment. An additional benefit is that SNS ratings do not require the administration of additional scales, because the SNS information is already there, thus reducing administration time for applicants. However, there may be some additional administration time required on behalf of hiring managers who would have to rate the Specific SNS Indicators.

Additional Validity Considerations. There may be some advantages of Specific SNS Indicator scales over a Global SNS Rating. First, it is unclear exactly what the Global SNS Rating measures. It does appear that the Global SNS Rating could tap aspects of Extraversion ($r = .229$, $p < .05$) and Openness to Experience ($r = .187$, $p < .05$), as it did have significant correlations with the traditional scales for both of these attributes (see Table 9). Additionally, the Global SNS Rating was significantly correlated with the majority of the Specific SNS Indicator scales (Agreeableness, $r = .333$, $p < .05$; Extraversion, $r = .362$, $p < .05$; Conscientiousness, $r = .746$, $p < .05$; Openness to Experience, $r = .472$, $p < .05$; Impression Management, $r = .772$, $p < .05$; and Verbal Ability, $r = .371$, $p < .05$; see Table 9). Therefore, the Global SNS Rating seems to have some overlap with the construct domains of the Specific SNS Indicators. This casts some doubt on the value of using the Global SNS Rating, as its convergent validity evidence is weak: it correlated with a large number of constructs.

Second, the discriminant validity of the Global SNS Rating could not be fully tested with the same level of scrutiny as the Specific SNS Indicators were. There was no traditional global rating with which to compare the results. However, the discriminant validity could be partially tested by examining the relations between the Global SNS Rating and the response distortion scales. The Global SNS Rating was not significantly correlated with the response

distortion scales (see Table 9). However, it was strongly correlated with the Specific SNS Indicator scale for Impression Management ($r = .772$, $p < .05$). This high correlation could mean that the Global SNS Rating may be influenced by the degree to which the rater believes the ratee is managing his/her impressions. These findings do little to alleviate researcher and hiring managers' concerns that non-job-related SNS information may be used for selection (Brown & Vaughn, 2011; Grasz, 2009; SHRM, 2011).

2.5.4 Legal Defensibility of Specific SNS Indicators

Despite the fairly minimal validity evidence for Specific SNS Indicators found in the current study, their use may offer some additional advantages for future researchers over a Global SNS Rating. The Specific SNS Indicators are tied to specific behaviours that can be evaluated by raters and documented, which could be of benefit for legal proceedings (Brown & Vaughn, 2011). For instance, imagine a situation where an applicant was suing an organization on the grounds of discrimination for discovering protected class information on the applicant's SNS account. An organization using Specific SNS Indicators could show the court exactly what was being assessed by the raters, and how those indicators are work-related by tying them to the specific attributes required for the job. With a Global SNS Rating, it would be more difficult to document what influenced the rating, and to which work-related attributes that rating is linked. These ratings rely on general observations and not specific behaviours, which could be more difficult to support in court, much as it has been the case for unstructured interviews (Williamson, Campion, Malos, Roehling, & Campion, 1997). As mentioned previously, a Global SNS Rating involves judgment at both the data collection and data combination stages, which is referred to as the *pure clinical strategy* (Cascio & Aguinis, 2005). A recent meta-analysis of mental health practitioners (Ægisdóttir et al., 2006) supported that the pure clinical strategy was less accurate than

strategies that relied on statistical methods (e.g., the mechanical composite strategy). In the personnel selection literature, the pure clinical strategy has also been found to be a less accurate method of prediction than other methods such as the mechanical composite strategy (Sawyer, 1966), which was the strategy used for the Specific SNS Indicator scales. The pure clinical strategy may be less accurate because raters' needs, response set, and wishes could be influencing how accurately their subjective information is combined (Cascio & Aguinis, 2005). It stands to reason that other elements, such as protected group characteristics, could also influence the combination of the subjective information.

2.5.5 Dropping Neuroticism and Quantitative Ability

One of the problems that was encountered in the current study was that a Neuroticism Specific SNS Indicator scale could not be created. The Taxonomy Expert Raters only identified two Specific SNS Indicators that they believed would tap Neuroticism, and those two indicators were too weakly correlated (r = .017, *n.s.*) and the reliability was too low (Spearman-Brown reliability = .03) to justify the creation of a composite scale.

Also, a Quantitative Ability Specific SNS Indicator scale could not be created for the current study. The Taxonomy Expert Raters could not identify any Specific SNS Indicators that would tap Quantitative Ability. Quantitative Ability has not been assessed in other SNS studies. In the current study, the Taxonomy Expert Raters suggested that SNSs seem to be largely void of content dealing with or related to math or numerical reasoning. Perhaps this is due to the focus on communication within SNSs. Most SNSs were created to promote conversations for individuals and groups (Kietzmann, Hermkens, McCarthy, & Silvestre, 2011). As a result, posts on SNSs are focused on conversations. Assuming that math or numerical reasoning is infrequently used in typical everyday conversations, it seems likely

that it is generally infrequently used on SNSs as well. However, there may be specific samples in which math or numerical reasoning is more frequently used, such as with mathematicians or astrophysicists. Perhaps Specific SNS Indicators could be created for use in such samples in the future. There were also some additional problems that were encountered in the study, such as the large number of Specific SNS Indicators removed during the item refinement stage.

2.5.6 Removal of Specific SNS Indicators

Of the 162 Specific SNS Indicators that were used at the start of the scale validation process, 47 indicators were removed. After the Taxonomy Expert Raters completed their ratings, these Specific SNS Indicators did not meet the cutoff criteria for wording quality, or they did not meet the cutoff criteria for being conceptually related to the human attributes commonly assessed by traditional selection scales (see Preliminary Evaluation of the Specific SNS Indicators section). Of the 115 Specific SNS that remained, 88 were rated by more than one rater. For the Specific SNS Indicators with multiple raters, only those with significant interrater reliabilities were retained. Sixteen Specific SNS Indicators were removed because they had non-significant interrater reliabilities. Also, two follow-up indicators were removed because the indicators that preceded them had been removed. The content of the follow-up indicators relied on the information from their previous indicators, so it did not make sense to retain these follow-up indicators. When the Specific SNS Indicator EFAs were tested, indicators were retained if they had significant factor loadings, and if these factor loadings were greater than .320 (as recommended by Tabachnick and Fidell, 2007). As a result, 21 Specific SNS Indicators were removed from the Specific SNS scales. Also, if an indicator were removed because its factor loading was not significant, its follow-up indicator was removed as well. In total, including the two follow-up indicators previously mentioned, six

follow-up indicators were removed. After these refinements had been conducted, only 29 Specific SNS Indicators remained to create the final Specific SNS Indicator scales. However, as mentioned previously, some of these Specific SNS Indicators were multidimensional in nature, and were used in the creation of more than one Specific SNS Indicator scale. For instance, X1, X26, X47, and X80 each loaded on two Specific SNS Indicator scales (see Table 1).

The percentage of Specific SNS Indicators that were retained (17.90%; 29/162) throughout the scale validation process was well below the 50% that Hinkin (1998) suggested that one can expect as part of any psychometric scale construction. The goal of the current study was to have refined scales with approximately 8 indicators per Specific SNS Indicator scale. Using the Hinkin estimate that 50% of items are typically retained as part of psychometric scale construction, I calculated that approximately 160 Specific SNS Indicators would be required at the beginning of the study. Because there was very little existing peer-reviewed literature to draw on to help with the creation of Specific SNS Indicators, I expected that it would be difficult to find reliable and valid indicators, and thus a large number of indicators would have to be removed. However, I did not expect to retain fewer indicators than the 50% that Hinkin had discussed. When possible, I took a more conservative approach when removing indicators in case future modifications could further improve the indicators. Additionally, a more conservative approach allowed adequate content validity coverage with respect to the major attributes that were assessed. See the right-hand column of Table 1 to see the attribute coverage for the indicators that were retained for the Specific SNS Indicator scales. The large number of indicators that had to be removed should serve as a reminder of the importance of starting with a large number of items before beginning any psychometric scale construction.

2.5.7 Limitations and Future Research Directions

There were several limitations in the current study that I recommend should be addressed as part of future research. As mentioned previously, I did not want to reject Specific SNS Indicators with low interrater reliability because further modifications could allow them to be more reliable indicators in the future. We encourage other researchers to expand on this list and see if additional Specific SNS Indicators with adequate interrater reliability can be created.

When examining the Specific SNS Indicators' factor loadings, I also took a more conservative approach for retaining indicators. Specific SNS Indicators were retained based on the recommended cut off of .320 in Tabachnick and Fidell (2007), rather than the higher cut offs other researchers have recommended (e.g., Hair et al., 2006). The higher cut offs could have led to the rejection of even more Specific SNS Indicators. Regardless of which cut offs are chosen as part of a CFA, it is important to remember that any model modification can capitalize on sample-specific fluctuations in the data, potentially overestimating how well a model would fit in the population (Goffin, 2007). Therefore, as with any validity study, I strongly recommend that these models be cross-validated in an independent sample to get a better sense of how well they really fit in the population.

In addition, there are some other topics worthy of further study. The current study was a preliminary investigation into the construct validity of SNS scales and their potential for use in selection. As a result, I strongly recommend that additional studies be conducted before any definitive conclusions are made on the topic. First, future researchers should collect criterion-related validity information such as correlations with performance data. Criterion-related validity would be valuable for those who aim to use Facebook profiles for

selection, as it is recommended under the Society for Industrial Organizational Psychology's Principles for the Validation and Use of Personnel Selection Procedures (2003). I sought to collect criterion-related validity information in the current study by contacting participants' previous supervisors. Unfortunately, only six participants gave the name of their supervisor, and none of them gave any contact information for their supervisors. This was likely due to a large number of participants not being currently employed. Only 6 out of 48 participants with employment data were currently employed. Therefore, I recommend future researchers use a concurrent sample of existing job incumbents or find a sample with more job experience in order to investigate criterion-related validity.

Although the Specific SNS Indicators had better incremental prediction in the prediction of their equivalent traditional measures than did the Global SNS Rating, a better test would be to focus on job performance as the dependent variable. Given that the items used to create the Global SNS Rating were not focused on personality, but rather an individual's potential suitability for a job, using job performance would serve as a more fair comparison.

The issue of discrimination has been mentioned at several points, and is often discussed in other articles within the social networking literature (e.g., Brown & Vaughn, 2011; Ruggs, Speights, & Walker, 2013; Willey, White, Domagalski, & Ford, 2012). Brown and Vaughn (2011) called for an investigation into discrimination associated with individual bias. In light of the results of the current study, I also recommend that both experimental and applied studies should investigate under which rating conditions discrimination or biases are most likely to occur when rating SNS accounts, and to what extent they do occur. For instance, does relying on Specific SNS Indicators rather than a Global SNS Rating actually

lead to less discrimination, and if so, is it limited to certain situations? Answers to these issues will help guide researchers and practitioners as they look to SNSs as a potential source for selection information. Future research may also help discover a more valid way of assessing SNS information for selection purposes. In the meantime, based on the results of the current study, it is recommended that practitioners avoid using SNS information for job selection.

2.6 Summary

- The initial average interrater reliability across all Specific SNS Indicators was .53, which is fairly low
- RPM format Specific SNS Indicators benefited from higher interrater reliabilities (average = .60) than did count format items (average = .35), but Global SNS Ratings were highest (average = .66)
- Interrater reliabilities were based on three Expert SNS Raters; however, interrater reliabilities would be lower if only one HR staff were conducting the ratings
- The structural validity of the Specific SNS Indicators was partially supported with two major exceptions: the two models using the Integrity SNS scale
- Convergent validity support was only found for Extraversion, Openness to Experience, and Verbal Ability SNS scales, and the effect sizes were only small to moderate
- The Extraversion and Verbal Ability Specific SNS Indicator scales led to better incremental prediction over the Global SNS Rating in the prediction of their respective traditional scales
- The Specific SNS Indicator scales may be less subject to response distortion than traditional scales

- The Global SNS Rating was not significantly correlated with the self-reported measures of response distortion, however, it was strongly correlated with the Specific SNS Indicator scale for Impression Management
- It is recommended that hiring managers avoid using SNS information as part of their selection systems until more research can be conducted on the validity of using this information

2.7 References

Ægisdóttir, S., White, M. J., Spengler, P. M., Maugherman, A. S., Anderson, L. A., Cook, R. S., & Rush, J. D. (2006). The meta-analysis of clinical judgment project: Fifty-six years of accumulated research on clinical versus statistical prediction. *The Counseling Psychologist, 34*(3), 341-382.

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behaviour*, *19*(3), 289-303.

Ashton, M. C., Jackson, D. N., Paunonen, S. V., Helmes, E., & Rothstein, M. G. (1995). The criterion validity of broad factor scales versus specific facet scales. *Journal of Research in Personality*, *29*(4), 432-442.

Ashton, M. C., & Lee, K. (2008). The prediction of Honesty–Humility-related criteria by the HEXACO and Five-Factor Models of personality. *Journal of Research in Personality*, *42*(5), 1216-1228.

Asparouhov, T. & Muthén, B. (2009). Exploratory structural equation modeling. *Structural Equation Modeling, 16,* 397-438.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). *Psychological Science*, *21*, 372-374.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology*, *44*, 1-26.

Bass, B. M., & Barrett, G. V. (1981). *People, work, and organizations: An introduction to industrial and organizational psychology*. Boston: Allyn and Bacon.

Black, S. L., & Johnson, A. F. (2012). Employers' Use of Social Networking Sites in the Selection Process. *The Journal of Social Media in Society, 1*, 7-28.

Bohnert, D., & Ross, W. H. (2010). The influence of social networking web sites on the evaluation of job candidates. *Cyberpsychology, Behaviour, and Social Networking, 13*, 341-347.

Brown, V. R., & Vaughn, E. D. (2011). The writing on the (Facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology, 26*, 219-225.

Browne, M. W. (2001). An overview of analytic rotation in exploratory factor analysis. *Multivariate Behavioural Research, 36*, 111-150.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136-162). Newbury Park, CA: Sage.

Cascio, W.F. & Aguinis, H. (2005). *Applied psychology in human resource management* (6th ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Catano, V. M., Wiesner, W. H., & Hackett, R. D. (2012). *Recruitment and Selection in Canada* (5th ed.). Scarborough, Nelson.

Christiansen, N., & Tett, R. (2013). *Handbook of personality at work*. New York, NY: Routledge.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences (2nd ed.). Hillsdale, NJ: Erlbaum.

Cook, R. L., & Stewart, T. R. (1975). A comparison of seven methods for obtaining subjective descriptions of judgmental policy. *Organizational Behaviour and Human Performance*, *13*(1), 31-45.

Costa, P. T. Jr., & McCrae, R. R. (1992). *NEO Personality Inventory-Revised (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) professional manual*. Odessa, FL: Psychological Assessment Resources.

Cullen, M. J., & Sackett, P. R. (2004). Integrity testing in the workplace. In C. Thomas (Ed.), *Comprehensive handbook of psychological assessment: Industrial and organizational assessment* (Vol. 4, pp. 149-165). Hoboken, NJ: Wiley.

Davison, H. K., Maraist, C., & Bing, M. N. (2011). Friend or foe? The promise and pitfalls of using social networking sites for HR decisions. *Journal of Business and Psychology*, *26*(2), 153-159.

Dillon, W. R., Kumar, A., & Mulani, N. (1987). Offending estimates in covariance structure analysis: Comments on the causes of and solutions to Heywood cases. *Psychological Bulletin*, *101*(1), 126.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational behaviour and human performance*, *7*(1), 86-106.

Eisinga, R., Te Grotenhuis, M., & Pelzer, B. (2013). The reliability of a two-item scale: Pearson, Cronbach, or Spearman-Brown?. *International Journal of Public Health*, *58*, 637-642.

Fleiss (1971) Measuring nominal scale agreement among many raters. *Psychological Bulletin, 76,* 378-382

Fornell, C. (1983). Issues in the application of covariance structure analysis: A comment. *Journal of Consumer Research, 9,* 443-448.

Funder, D. C. (1999). *Personality judgment: A realistic approach to person perception*. San Diego, CA: Academic Press.

Ganzach, Y., Kluger, A. N., & Klayman, N. (2000). Making decisions from an interview: Expert measurement and mechanical combination. *Personnel Psychology*, *53*(1), 1-20.

Gatignon, H. (2010). Confirmatory Factor Analysis in Statistical analysis of management data. DOI: 10.1007/978-1-4419-1270- 1_4.

Goffin, R. D. (2007). Assessing the adequacy of structural equation models: Golden rules and editorial policies. *Personality and Individual Differences*, *42*(5), 831-839.

Goffin, R. D., Gellatly, I. R., Paunonen, S. V., Jackson, D. N., & Meyer, J. P. (1996). Criterion validation of two approaches to performance appraisal: The behavioural observation scale and the relative percentile method. *Journal of Business and Psychology*, *11*(1), 23-33.

Goffin, R. D., & Jackson, D. N. (1988). The structure validity of the index of organizational reactions. *Multivariate Behavioural Research*, *23*(3), 327-347.

Goffin, R. D., Jelley, R. B., Powell, D. M., & Johnston, N. G. (2009). Taking advantage of social comparisons in performance appraisal: The relative percentile method. *Human Resource Management*, *48*(2), 251-268.

Goffin, R. D., & Olson, J. M. (2011). Is it all relative? Comparative judgments and the possible improvement of self-ratings and ratings of others. Perspectives on Psychological Science, 6, 48-60.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Goldstein, I. L., Zedeck, S., & Schneider, B. (1993). An exploration of the job analysis-content validity process. *Personnel selection in organizations*, *3*, 34.

Gosling, S. D., Augustine, A. A., Vazire, S., Holtzman, N., & Gaddis, S. (2011). Manifestations of personality in online social networks: Self-reported Facebook-related behaviors and observable profile information. *Cyberpsychology, Behavior, and Social Networking*, *14*(9), 483-488.

Gough, H.C. (1962). Clinical versus statistical prediction in psychology. In L. Postman (Ed.). *Psychology in the making.* New York: Knopf.

Gourraud, P. A., Le Gall, C., Puzenat, E., Aubin, F., Ortonne, J. P., & Paul, C. F. (2012). Why statistics matter: limited inter-rater agreement prevents using the psoriasis area and severity index as a unique determinant of therapeutic decision in psoriasis. *Journal of Investigative Dermatology*, *132*(9), 2171-2175.

Grasz, J. (2009). Forty-five percent of employers use social networking sites to research job candidates, CareerBuilder survey finds. *CareerBuilder Press Releases*. Retrieved June, 5, 2012.

Grubb, W.L., Whetzel, D.L., & McDaniel, M.A. (2004). *General mental ability tests in industry*. In J.C. Thomas (Ed.), Comprehensive handbook of psychological assessment: Vol. 4, Industrial and organizational assessment (pp. 7–21). New Jersey: John Wiley & Sons.

Hair, J., Black, W., Babin, B., Anderson, R., & Tatham, R. (2006). *Multivariate Data Analysis, Sixth Ed*. Upper Saddle River, N.J.: Pearson Education Inc.

Hastings, S. E., & O'Neill, T. A. (2009). Predicting workplace deviance using broad versus narrow personality variables. *Personality and Individual Differences*, *47*(4), 289-293.

Heneman, R. L. (1986). The relationship between supervisory ratings and results-oriented measures of performance: A meta-analysis. *Personnel Psychology*, *39*(4), 811-826.

Hinkin, T. R. (1998). A brief tutorial on the development of measures for use in survey questionnaires. *Organizational Research Methods*, *1*, 104-121.

Hitt, M. A., & Barr, S. H. (1989). Managerial selection decision models: Examination of configural cue processing. *Journal of Applied Psychology, 74*(1), 53.

Hodges, B., Regehr, G., Hanson, M., & McNaughton, N. (1997). An objective structured clinical examination for evaluating psychiatric clinical clerks. *Academic Medicine, 72*, 715-721. doi:10.1097/00001888-199708000-00019

Hooper, D., Coughlan, J., & Mullen, M. (2008). Structural equation modeling: Guidelines for determining model fit. *Journal of Business Research Methods, 6,* 53-60.

Hough, L. M., Eaton, N. K., Dunnette, M. D., Kamp, J. D., & McCloy, R. A. (1990). Criterion-related validities of personality constructs and the effect of response distortion on those validities. *Journal of applied psychology*, *75*(5), 581.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55.

Hubley, A. M. (2014). Discriminant validity. *Encyclopedia of quality of life and well-being research*, 1664-1667.

Jackson, D. N. (1998). Multidimensional aptitude battery II. *Sigma Assessment Systems, Port Huron, MI*.

Jackson, D. N. (2004). Personnel assessment form manual. *Sigma Assessment Systems, Port Huron, MI*.

Johnson, J. A. (2001). Screening massively large data sets for non-responsiveness in web-based personality inventories. The Netherlands: Invited talk to the joint Bielefeld-Groningen Personality Research Group, University of Groningen.

Jöreskog, K. G., & Sörbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language*. Chicago, IL: Scientific Software International.

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis*. New York, NY: Guilford Press.

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2014). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research.*

Kenny, D. A. (1994). *Interpersonal perception: A social relations analysis.* New York: Guilford Press.

Keeley, S. M., & Doherty, M. E. (1972). Bayesian and regression modeling of graduate admission policy. *Organizational Behaviour and Human Performance*, *8*, 297-323.

Kietzmann, J. H., Hermkens, K., McCarthy, I. P., & Silvestre, B. S. (2011). Social media? Get serious! Understanding the functional building blocks of social media. *Business horizons*, *54*, 241-251.

Kim, K. H. (2005). The relation among fit indexes, power, and sample size in structural equation modeling. *Structural Equation Modeling, 12*(3), 368-390.

Kline, R. B. (2011). *Principles and practice of structural equation modeling*. Guilford press.

Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: evaluating social networking web sites. *Journal of managerial Psychology*, *24*(6), 567-580.

Kluemper, D. H., Rosen, P. A., & Mossholder, K. W. (2012). Social Networking Websites, Personality Ratings, and the Organizational Context: More Than Meets the Eye?. *Journal of Applied Social Psychology*, *42*(5), 1143-1172.

Krämer, N. C., & Winter, S. (2008). Impression management 2.0: The relationship of self-esteem, extraversion, self-efficacy, and self-presentation within social networking sites. *Journal of Media Psychology: Theories, Methods, and Applications*, *20*(3), 106.

Lampe, C. A., Ellison, N., & Steinfield, C. (2007, April). A familiar face (book): profile elements as signals in an online social network. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 435-444). ACM.

Lance, C. E., Butts, M. M., & Michels, L. C. (2006). The Sources of Four Commonly Reported Cutoff Criteria What Did They Really Say?. *Organizational Research Methods*, *9*(2), 202-220.

Landis, R. S., Beal, D. J., & Tesluk, P. E. (2000). A comparison of approaches to forming composite measures in structural equation models. *Organizational Research Methods*, *4*, 186-207. doi:10.1177/109442810032003

LeBreton, J. M., Burgess, J. R., Kaiser, R. B., Atchley, E. K., & James, L. R. (2003). The restriction of variance hypothesis and interrater reliability and agreement: Are ratings from multiple sources really dissimilar?. *Organizational Research Methods*, *6*(1), 80-128.

LeBreton, J. M., & Senter, J. L. (2007). Answers to 20 questions about interrater reliability and interrater agreement. *Organizational Research Methods, 11*, 815-852.

Lee, K., Ashton, M. C., & Shin, K. H. (2005). Personality Correlates of Workplace Anti-Social Behaviour. *Applied Psychology*, *54*(1), 81-98.

Levinson, M. (2011). Social networks: New hotbed for hiring discrimination claims. *Computer World*. Retrieved from http://www.computerworld.com/s/article/9215907/Social_Networks_New_Hotbed_for_Hiring_Discrimination_Claims.

Loevinger, J. Objective tests as instruments of psychological theory. In D. M. Jackson and S. Messick (Eds.), *Problems in human assessment.* New York: McGraw-Hill, 1967.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods*, *1*(2), 130.

Malau-Aduli, B. S., Mulcahy, S., Warnecke, E., Otahal, P., Teague, P. A., Turner, R., & Van der Vleuten, C. (2012). Inter-rater reliability: Comparison of checklist and global scoring for OSCEs. *Creative Education*, *3*, 937-942.

Marcus, B. (2007). Recent Evidence on the Job-Related Attitudes and Self-Evaluation Inventory (IBES). *Zeitscrhift fur Personalpsychologie, 6(3)*, 129-132.

Marcus B. (2006).Inventar berufsbezogener Einstellungen und Selbsteinschatzungen [Job-Related attitudes and self-evaluations inventory]. Göttingen, Germany: Hogrefe

Marcus, B., Höft, S., & Riediger, M. (2006). Integrity tests and the five-factor model of personality: A review and empirical test of two alternative positions. *International Journal of Selection and Assessment*, *14*(2), 113-130.

Marcus, B., Lee, K., & Ashton, M. C. (2007). Personality dimensions explaining relationships between integrity tests and counterproductive behaviour: Big Five, or one in addition?. *Personnel Psychology*, *60*(1), 1-34.

Marsh, H. W., Hau, Kit-Tai, & Wen, Z. (2004). In search of Golden rules: comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*, 320–341.

Martin, J.A., Regehr, G., Reznick, R., Murnaghan, M.J., Hutchison, C., & Brown, M. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. *British Journal of Surgery, 84*, 273-278.

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data, *Psychological Methods*, *17*(3), 437-455.

Meehl, P.E. (1954). *Clinical versus statistical prediction: A theoretical analysis and review of the evidence*. Minneapolis: University of Minnesota Press.

Murphy, K. R. (2002). Can conflicting perspectives on the role of g in personnel selection be resolved?. *Human Performance*, *15*(1-2), 173-186.

Muthén, L. K., & Muthén, B. O. (2012). Mplus. *The comprehensive modelling program for applied researchers: User's guide*, *5*.

O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and Velicer's MAP test. *Behaviour Research Methods, Instruments, & Computers*, *32*, 396-402.

O'Neill, T. A., Lewis, R. J., & Carswell, J. J. (2011). Employee personality, justice perceptions, and the prediction of workplace deviance. *Personality and individual differences*, *51*(5), 595-600.

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth–fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behaviour*, *17*(6), 609-626.

Ones, D. S., & Viswesvaran, C. (1998). The effects of social desirability and faking on personality and integrity assessment for personnel selection. *Human performance*, *11*(2-3), 245-269.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of applied psychology*, *78*(4), 679.

Ones, D.S., Schmidt, F.L. & Viswesvaran, C. (1994a). *Do broader personality variables predict job performance with higher validity?* Paper presented at the 1994 Conference of the Society for Industrial and Organizational Psychology, Nashville, TN.

Ones, D.S., Schmidt, F.L. & Viswesvaran, C. (1994b). *Correlates of pre-employment integrity tests*. Paper presented at the 23rd International Congress of Applied Psychology, Madrid, Spain.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*, 598-609.

Paulhus, D. L. (1988). Assessing self-deception and impression management in self-reports: The Balanced Inventory of Desirable Responding. (Manual available from the author).

Paulhus, D. L. (1991) Measurement and control of response bias. In J.P. Robinson, P.R. Shaver and L.S. Wrightsman (Eds), *Measures of personality and social psychology attitudes* (pp. 17–59). New York: Academic Press.

Paunonen, S. V. (1984). The reliability of aggregated measurements: Lessons to be learned from psychometric theory. *Journal of Research in Personality,18,* 383-394.

Paunonen, S. V., & O'Neill, T. A. (2010). Self-reports, peer ratings and construct validity. *European Journal of Personality*, *24*(3), 189-206.

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.

Preston, J. (2011). Social media history becomes a new job hurdle. *New York Times*.

Rothstein, M. G., & Goffin, R. D. (2006). The use of personality measures in personnel selection: What does current research support?. *Human Resource Management Review*, *16*(2), 155-180.

Regehr, G., MacRae, H., Reznick, R. K., & Szalay, D. (1998). Comparing the psychometric properties of checklists and global rating scales for assessing performance on an OSCE-format examination. *Academic Medicine, 73*, 993-997. doi:10.1097/00001888-199809000-00020

Ruggs, E. N., Speights, S., & Walker, S. S. (2013). Are You In or Out? Employment Discrimination in Online and Offline Networks. *Industrial and Organizational Psychology, 6*, 457-462.

Sackett, P. R., & Wanek, J. E. (1997). Integrity testing: An overview. *Security Journal, 8*, 11-18.

Salgado, J. F. (1997). The Five Factor Model of personality and job performance in the European Community. *Journal of Applied psychology*, *82*(1), 30.

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological bulletin*, *66*(3), 178.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, *124*(2), 262.

Schneider, T. J., & Goffin, R. D. (2012). Perceived ability to deceive and incremental prediction in pre-employment personality testing. *Personality and Individual Differences*, *52*(7), 806-811.

Schneider, R. J., Hough, L. M., & Dunnette, M. D. (1996). Broadsided by broad traits: How to sink science in five dimensions or less. *Journal of Organizational Behaviour*, *17*(6), 639-655.

Slovensky, R., & Ross, W. H. (2012). Should human resource managers use social media to screen job applicants? Managerial and legal issues in the USA. *Emerald Insight, 14(1)*, 55-69.

Society for Human Resource Management (2011). *Use of social networking websites to recruit potential job candidates in 2008 vs 2011*. Retrieved from http://www.shrm.org/research/surveyfindings/documents/social%20networking%20flyer_staffing%20conference_final1.pdf

Society for Industrial and Organizational Psychology. (2003). *Principles for the validation and use of personnel selection* (4th ed.) Bowling Green, OH: Author (Available on-line).

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. (2007). Multivariate analysis of variance and covariance. *Using multivariate statistics, 3*, 402-407.

Thompson, R. L., Iacovou, C. L., & Shirland, L. E. (2006). A comparison of evaluation techniques for decision analysis involving large attribute sets. *Journal of Applied Business Research, 22*, 25-38.

Van Iddekinge, C. H., Lanivich, S. E., Roth, P. L., & Junco, E. (2013). Social Media for Selection? Validity and adverse impact potential of a Facebook-Based Assessment. *Journal of Management*. DOI: 10.1177/0149206313515524

Vazire, S., & Gosling, S. D. (2004). e-Perceptions: Personality impressions based on personal websites. *Journal of personality and social psychology*, *87*, 123-132.

Willey, L., White, B. J., Domagalski, T., & Ford, J. C. (2012). Candidate-screening, information technology and the law: Social media considerations. *Issues in Information Systems, 13,* 300-309.

Williamson, L. G., Campion, J. E., Malos, S. B., Roehling, M. V., & Campion, M. A. (1997). Employment interview on trial: Linking interview structure with litigation outcomes. *Journal of Applied Psychology*, *82*(6), 900.

Wonderlic, Inc. (2002). Wonderlic personnel test & scholastic level exam user's manual. Libertyville, IL: Wonderlic, Inc.

Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under-and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, *1*, 354-365.

Zettler, I., Hilbig, B. E., Moshagen, M., & de Vries, R. E. (2015). Dishonest responding or true virtue? A behavioral test of impression management. *Personality and Individual Differences, 81*, 107-111.

Zwick, W. R., & Velicer, W. F. (1986). Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*, *99*, 432-442.

Chapter 3

3 “Give Us Your Social Networking Site Passwords”: Implications for Personnel Selection and Personality

In this chapter, I will investigate the practical realities surrounding the use of SNS information for personnel selection decisions, and what effects SNS selection might have on the workforce. The goal is to determine what the positive or negative effects could be if employers were to ask applicants for their SNS password as part of the selection process. This article was published in Personality and Individual Differences in 2015: Schneider, T. J., Goffin, R. D., & Daljeet, K. N. (2015). “Give us your social networking site passwords”: Implications for personnel selection and personality. *Personality and Individual Differences*, *73*, 78-83.

3.1 Abstract

Recently, employers have begun asking applicants for their social networking site (SNS) password to access private information that could be job-relevant. However, the effect that this request can have on an organization’s selection process and its selection of individual applicant traits has not been previously examined. Findings from the current study of 892 employed or previously-employed participants suggested that 57.87% of the sample would refuse the password request, thereby removing themselves from the applicant pool. Such a large reduction in the applicant pool could necessitate a drastic decrease in cutoff scores on subsequent pre-employment tests, which would lower workforce productivity and personnel selection utility. In addition, the SNS password request caused adverse impact for several minority groups, and affected the personality scores of the remaining applicant pool.

Keywords

Social Networking Sites; Selection; Personality; Individual Differences; Adverse Impact

3.2 Introduction

It is becoming increasingly common among employers to look beyond traditional pre-employment evaluations to the internet as a source for collecting relevant pre-hiring information (Karl, Peluchette & Schlaegel, 2010; Stoughton, Thompson, & Meade, 2013; Taylor, 2006). This trend has been increasing (CareerBuilder.com, 2009), to the extent that, as of 2011, up to 65% of organizations are estimated to use SNSs to screen job applicants (Levinson, 2011). In fact, according to the popular press, some employers have begun asking applicants for their personal social networking site (SNS) password in order to access private information that could be job-relevant (McFarland, 2012). Predictably, this practice has raised numerous legal, privacy, and ethical concerns (Levinson, 2011). Although legislation has been created in several U.S. states to ban this practice of requiring a SNS password in order to be considered for a job (Noble & Davis, 2014), it is still legal in many countries (Determann, 2012). In addition, even if only a few employers are engaging in or are considering using this practice, this is an important issue worthy of further study.

I investigated the implications that the SNS password request has regarding the individual differences of the applicants, and I investigated questions relevant to the use of the SNS password request in employee selection. In particular, could asking applicants for their SNS password, and eliminating those who refuse the password request (see Noble & Davis, 2014), substantially reduce the applicant pool? Moreover, could this practice lead to discrimination associated with applicant characteristics (e.g., gender, ethnicity, etc.)? Finally, a number of employers have adopted the practice of removing applicants from the applicant

pool if they refuse the SNS password (Noble & Davis, 2014). Employers may presume that a refusal of the password request is indicative of personalities who would tend to have an abundance of undesirable material displayed in their SNS, and that this does not bode well for their future job performance. Accordingly, they might expect that those who remain in the applicant pool by complying with the password request will generally have personalities that will lead to higher job performance. Thus the question of whether the SNS password request will tend to improve the applicant pool by enhancing, in work-relevant ways, the personality scores of the remaining applicant pool, was of great interest. In particular, I looked at traits of the Five-Factor Model (FFM; McCrae & Costa, 2010), the Dark Triad (Paulhus & Williams, 2002), and Impression Management (IM; Paulhus, 1991). The hypotheses and research questions that were formulated in order to address the above issues are developed below.

If a large number of applicants were removed from the applicant pool because they refused the password request, there would ultimately be fewer candidates to choose from. As a result, a less stringent cutoff score would have to be set on the subsequent selection tests, which would reduce the bottom-line value of testing (Cascio & Aguinis, 2011). In a typical selection setting, such as the one reported by Ryan, Sacco, McFarland, and Kriska (2000), approximately 35% of applicants withdrew from the early stages of selection for a variety of reasons (e.g., negative perceptions of the organization, low commitment to obtain the job, need to relocate, less social support). Using Ryan et al.'s statistic as a baseline, if the proportion of applicants who refuse the SNS password request is significantly greater than .35, there may be grounds for inferring that the password request has caused an inordinate loss of individuals from the applicant pool. I predict that a sizable proportion of the sample will refuse the SNS password request because most individuals are concerned about the

privacy of their personal information on the internet (see Joinson, Reips, Buchanan, & Schofield, 2010). Therefore, I predict:

H_1: The proportion of participants who would refuse the SNS password request will be significantly greater than .35.

Of course, if personality traits are associated with password-request compliance such that those who comply with the request tend to have noticeably superior scores on job-relevant traits, then a reduction in pool size could be beneficial. I explore this, and related implications, in a later section.

3.2.1 Discrimination and SNS Password Requests

In addition to reducing the number of available applicants, the request for a SNS password may also be problematic as a screening device if it results in discrimination. The doctrine of disparate impact holds that if a selection test disproportionately and adversely affects a protected group (e.g., based on age, race, sex, marital status, or religious affiliation), its use must be discontinued unless it can be adequately demonstrated that it is necessary to ensure the success of a business (Catano, Wiesner, Hackett, & Methot, 2012). I speculate that those who are not members of the majority group may be less likely to divulge their SNS password because their SNS information could confirm and even exacerbate their outsider status or certain biases against them. For instance, women might not release their SNS password for fear of not being hired if their SNS information reveals that they are, or wish to be, pregnant.

If requiring job applicants to provide their SNS password leads to a disproportionate number of protected group members being removed from the selection process, adverse impact could result, which could have serious legal implications. Adverse impact occurs

when a protected group is hired at less than 80% of the rate of another group (see the four-fifths rule in Cascio & Aguinis, 2011). Therefore:

RQ_1: Will the SNS password request result in evidence of adverse impact for gender, age, ethnicity, sexual orientation, and/or religious affiliation?

3.2.2 Personality Differences and Complying with or Refusing the Password Request

The Five-Factor Model of Personality. FFM traits are important and useful predictors of job performance (Ones, Dilchert, Viswesvaran, & Judge, 2007) and personality could play a role in complying with or refusing a SNS password request. Thus, it is important to examine whether employers who limit their choice of applicants to those who comply with the password request may be systematically affecting the FFM scores of the remaining applicant pool. Because compliance with the SNS password request could be associated with personality, and because personality scores are related to job performance, organizations could be improving or worsening the personality scores of the subsequent applicant pool when they require one's SNS password as a precondition for hiring.

Junglas, Johnson, and Spitzmüller (2008) found a significant, weak, but positive (standardized path coefficient of .11) relation between Openness to Experience and concern for privacy. Junglas et al. described open individuals as being more aware, and therefore, more sensitive to threatening situations. However, other characteristics of Openness, such as the Actions facet (see McCrae & Costa, 2010), might make high scorers more willing to release their information because they are more open to trying and experiencing new things. Thus, I propose this non-directional hypothesis:

H_2: Individuals who would agree to divulge their SNS password will have a significantly different mean level of Openness to Experience than those who would refuse.

Individuals high in Agreeableness are more likely to trust and less likely to appraise others' actions as harmful (Junglas et al., 2008). Also, agreeable individuals may reveal their SNS password as a way of avoiding conflict and maintaining harmony (see McCrae & Costa, 2010). Moreover, Agreeableness had a weak but significant negative standardized path coefficient (-.22) with concern for privacy in Junglas et al. (2008). Therefore:

H_3: Individuals who would agree to divulge their SNS password will have a significantly higher mean level of Agreeableness than those who would refuse.

On the one hand, Junglas et al. (2008) reported that Conscientiousness had a significant, positive, albeit weak (.12), standardized path coefficient with one's concern for privacy. On the other hand, conscientious individuals are likely to be more careful about the content they allow on their SNS accounts in the first place, meaning that they have less to fear by allowing access to their SNSs. Thus, as with Openness, I propose a non-directional hypothesis:

H_4: Individuals who would agree to divulge their SNS password will have a significantly different mean level of Conscientiousness than those who would refuse.

Based on the definition of Extraversion (McCrae & Costa, 2010), and the fact that Exhibition can be an important component of Extraversion (Jackson, Paunonen, Fraboni, & Goffin, 1996), extraverted individuals should tend to display their social, outgoing nature via their SNS content. Therefore:

H_5: Individuals who would agree to divulge their SNS password will have a significantly higher mean level of Extraversion than those who would refuse.

Neurotic individuals tend to experience more negative affect, including fear and anxiety (McCrae & Costa, 2010). Therefore, they may have heightened concerns about employers accessing their SNS content. Accordingly:

H_6: Individuals who would agree to divulge their SNS password will have a significantly lower mean level of Neuroticism than those who would refuse.

The Dark Triad. The Dark Triad is a commonly-studied set of personality traits that falls outside the FFM, but is linked to job performance and counterproductive work behaviors (O'Boyle, Forsyth, Banks, & McDaniel, 2012). *Machiavellianism* describes individuals who are manipulative and cold (Paulhus & Williams, 2002). *Psychopathy* describes individuals who lack empathy for others, engage in antisocial behaviors, engage in risk-taking, and are callous (Williams, Paulhus, & Hare, 2007). Finally, *Narcissism* describes individuals who are egotistic, self-enhancing, and who believe they are entitled and superior to others (Paulhus & Williams, 2002). I could find no compelling conceptual or empirical linkages between Machiavellianism and the likelihood of complying with the SNS password request, so attention was focused on Psychopathy and Narcissism.

Psychopathy has been linked to rebelling against authority and engaging in illegal activities (O'Boyle et al., 2012). As a result, those who score high on Psychopathy may tend to have evidence of illegal activities and rebellion against authority within their SNS accounts. Also, refusing the SNS password request may be a form of rebellion against authority. Consequently,

H_7: Individuals who would agree to divulge their SNS password will have a significantly lower mean level of Psychopathy than those who would refuse.

The more narcissistic people are, the more willing they are to have others find out all about them because of their strong need for self-affirmation (Morf & Rhodewalt, 2001). Therefore, I predict:

H_8: Individuals who would agree to divulge their SNS password will have a significantly higher mean level of Narcissism than those who would refuse.

Impression Management. Those who score high on Impression Management try to actively manage or distort the perception of their image (Paulhus, 1991). Such individuals should be less concerned about employers accessing their SNS because they should be more likely to censor their SNS content in the first place as part of their self-presentation strategy. Thus:

H_9: Individuals who would agree to divulge their SNS password will have a significantly higher mean level of Impression Management than those who would refuse.

3.2.3 The Present Study

In order to investigate these issues, I decided to present a hypothetical question about a SNS password request by a hiring manager to participants who had work experience but were not actually applying for a job. I did not use actual applicants because if they were asked whether they were no longer in contention for a prospective job as a result of a SNS password request, I presumed they would not readily admit this. Such an admission would imply that they had something to hide, and this runs counter to the pervasive tendency to respond in a socially desirable manner (Jackson, 1999) Consequently, they might be more inclined to identify other factors (e.g., poor fit with the organization, other job prospects, relocation, etc.) as the reasons for them no longer being in contention for the job, in much the same way that employees who are fired often do not report "firing" as the cause of their job loss (Giacalone & Duhon, 1991). Consequently, using actual applicants as participants could result in a drastic underestimate of the true rate of refusing the SNS password request. Also, I avoided asking participants whether they would have complied with the request for their SNS password if it had been part of the selection process for their present job, because I believed that cognitive dissonance (Wicklund & Brehm, 2013) would push them towards rationalizing

their current job tenure by saying that they would have complied with the SNS password request, whether that were true or not. Therefore, I felt that asking a hypothetical question about compliance with the SNS password request to individuals who have work experience, and who will likely be in the job market in the near future, would provide an effective approach to testing our research question and hypotheses.

3.3 Method

3.3.1 Participants

After removing careless responders (see below), the sample consisted of 892 employed or previously-employed individuals who also were enrolled in at least one psychology course at a major Canadian university.

Based on suggestions by Meade and Craig (2012), 227 participants were identified as careless responders and were excluded because they responded incorrectly to at least one of two embedded bogus items , or, they chose the option "do NOT use my responses" at the end of the study.

Participants' ages ranged from 16 to 44 (M = 18.41, SD = 2.13). Further description of the sample is provided in Table 12.

3.3.2 Measures

Demographics. Participants reported their gender, age, ethnicity, sexual orientation, and religious affiliation.

Five-Factor Model Personality Traits. Participants completed the International Personality Item Pool (IPIP) version of the NEO-PI-R (Goldberg, 1999). It uses 50 self-report items to assess Openness to Experience, Agreeableness, Conscientiousness, Extraversion, and Neuroticism (10 items per trait). Each item uses a 5-point scale ranging

from 1 (*strongly disagree*) to 5 (*strongly agree*). These scales have internal consistency ranging from .77 to .86 and have evidence of validity (Goldberg, 1999).

Dark Triad Personality Traits. Participants completed a measure (Paulhus & Jones, 2011) of Psychopathy (9 items) and Narcissism (9 items), which used the same response format as the FFM measure. Paulhus and Jones (2011) reported that the internal consistency of the two scales ranged from .73 to .90, and provided evidence of validity.

Impression Management. Participants completed the Impression Management scale (Paulhus, 1991) using a response scale ranging from 1 (*strongly disagree*) to 7 (*strongly agree*). Due to its sensitive nature, one item was dropped from this scale, leaving 19 items. The internal consistency of this scale ranges from .75 to .86 and there is evidence of its validity (Paulhus, 1991).

SNS Password Request. Participants were asked to respond "yes" or "no" to the following question:

> "*Imagine that you are applying for a highly desirable job, and during the job interview, the hiring manager requests that you give him/her the username/password for your social media account(s). The hiring manager explains that the organization wants to look through your social media account(s) to get better insight into who you really are. Would you give the hiring manager your username/password?"*

Table 12 (Study 2): Number of Participants who Would Comply with or Refuse the Request for SNS Password

	Overall	Comply with Request		Refuse Request	
	N	N	Selection Ratio	N	Adverse Impact Ratio
Entire Sample	883	372	0.42	511	
Gender					
Male	270	110	0.41	160	
Female	613	262	0.43	351	1.05
Age					
Under 40	876	370	0.42	506	
40 and Over	2	1	0.50	1	1.18
Ethnicity					
White/European	617	275	0.45	342	
Black North American/African	11	5	0.45	6	1.02
Asian	144	46	0.32	98	0.72
Hispanic	5	3	0.60	2	1.35
North American Indian	10	5	0.50	5	1.12
East Indian	28	14	0.50	14	1.12
Other	66	24	0.36	42	0.82
Sexual Orientation					
Heterosexual	844	356	0.42	488	
Homosexual	9	3	0.33	6	0.79
Bisexual	13	6	0.46	7	1.09
Other	4	2	0.50	2	1.19

(continued)

Table 12 (Study 2) continued: Number of Participants who Would Comply with or Refuse the Request for SNS Password

	Overall	Comply with Request		Refuse Request	
	N	N	Selection Ratio	N	Adverse Impact Ratio
Religious Affiliation					
Christian	362	167	0.46	195	
Jewish	46	18	0.39	28	0.85
Muslim	21	11	0.52	10	1.14
Hindu	23	9	0.39	14	0.85
Buddhist	14	5	0.36	9	0.77
Atheist	96	37	0.39	59	0.84
Personal Religion (no affiliation with a religious group)	33	16	0.48	17	1.05
No Religion (though am not an atheist)	218	77	0.35	141	0.77
Other	51	25	0.49	26	1.06

Note. N = 883. To create adverse impact ratios, males were the majority group for gender, individuals under 40 were for age (those 40 and over are the protected class according to the U.S. Age Discrimination in Employment Act), Whites were for ethnicity, heterosexuals were for sexual orientation, and Christians were for religious affiliation.

3.3.3 Procedure

Participants participated online in return for course credit. After giving consent, they completed the current measures as part of a larger battery.

3.4 Results and Discussion

3.4.1 Descriptives

The mean item scores (min of 1.0, max of 7.0 for IM; min of 1.0, max of 5.0 for all other personality scales), standard deviations, and internal consistencies of all the measures are reported in Table 13. The zero-order correlations for all of the study variables can be found in Table 14.

3.4.2 SNS Password Request and Reduction of the Applicant Pool

The number of participants who indicated they would comply with or refuse the request for their SNS password is reported in Table 12, along with the selection ratios and adverse impact ratios for each demographic group. The selection ratio is the number of participants in the respective demographic group who indicated they would comply with the password request, divided by the total number of participants in that group. The selection ratio for a given group is divided by the selection ratio for the "highest group" to compute the adverse impact ratio (described in the introduction). Overall, the proportion of participants indicating they would refuse the password request was large (511/883=.58). As explained earlier, this refusal would result in their removal from the applicant pool. In support of H_1, .58 was significantly greater than the baseline .35 ratio ($z = 14.33$, $p < .001$), and the respective Cohen's *d* (.96) was large. The rationale for setting .35 as the baseline was explained in the introduction. Such a large loss of applicants would lead to a large increase in an organization's selection ratio, which would lower the utility of an

Table 13 (Study 2): Means, Standard Deviations, Reliabilities, Z-Test, and Cohen's d

	$M_{Overall}$	$SD_{Overall}$	α	M_{Comply}	M_{Refuse}	Z-Test	Cohen's *d*
Openness to Experience	3.47	.58	.73	3.45	3.49	-1.09	-.07
Agreeableness	3.68	.54	.77	3.75	3.63	3.29*	.22
Conscientiousness	3.47	.59	.80	3.54	3.41	3.04*	.21
Extraversion	3.54	.74	.89	3.56	3.53	.53	.04
Neuroticism	2.65	.76	.86	2.61	2.67	-1.18	-.08
Psychopathy	2.11	.58	.76	2.06	2.15	-2.32*	-.16
Narcissism	3.08	.50	.66	3.10	3.06	1.07	.07
Impression Management	3.62	.73	.79	3.74	3.52	4.40*	.30

Note. N = 892. α = Cronbach's alpha reliability. $M_{Overall}$ = Mean item score for the overall sample. $SD_{Overall}$ = Standard Deviation of item score for the overall sample. M_{Comply} = Mean item score for those who would comply with the SNS password request. M_{Refuse} = Mean item score for those who would refuse the request. Negative z-tests and Cohen's *d* values indicate that the mean was higher for those who would refuse than those who would comply with the request.

*$p < .05$, two-tailed.

Table 14 (Study 2): Zero-Order Correlations

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1. Gender	-													
2. Age	-0.07	-												
3. ETH	-0.04	-0.01	-											
4. SO	-0.02	0.02	0.03	-										
5. RA	-0.12*	0.05	0.19*	0.05	-									
6. OPEN	0.00	0.11*	0.08*	0.08*	0.14*	-								
7. AG	0.18*	0.03	-0.04	-0.06	-0.06	0.12*	-							
8. CON	0.10*	0.06	-0.15*	-0.04	-0.12*	0.07*	0.30*	-						
9. EXT	0.09*	-0.08*	-0.14*	-0.14*	-0.08*	0.17*	0.12*	0.19*	-					
10. NEUR	0.21*	-0.03	0.05	0.13*	-0.01	0.01	-0.38*	-0.28*	-0.39*	-				
11. MACH	-0.21*	-0.04	0.11*	0.01	0.10*	-0.12*	-0.52*	-0.16*	-0.06	0.18*	-			
12. PSYC	-0.20*	0.01	0.04	0.01	0.09*	-0.12*	-0.57*	-0.29*	0.11*	0.17*	0.48*	-		
13. NARC	-0.10*	-0.11*	-0.06	-0.06	-0.02	0.10*	-0.05	0.18*	0.60*	-0.24*	0.21*	0.26*	-	
14. IM	0.07	0.05	0.02	-0.05	-0.04	0.06	0.45*	0.36*	-0.09*	-0.21*	-0.33*	-0.50*	-0.10*	-

Note. N = 892. ETH = Ethnicity. SO = Sexual Orientation. RA = Religious Affiliation. OPEN = Openness to Experience. AG = Agreeableness. CON = Conscientiousness. EXT = Extraversion. NEUR = Neuroticism. MACH = Machiavellianism. PSYC = Psychopathy. NARC = Narcissism. IM = Impression Management. The following variables were recoded: ETH (0 = Whites, 1 = other groups), SO (0 = Heterosexual, 1 = other groups), and RA (0 = Christian, 1 = other groups).
*$p < .05$, two-tailed.

organization's selection procedures (Cascio & Aguinis, 2011). For example, assuming that some form of standardized testing with top-down selection occurred after the SNS password request, an organization would have to drastically lower their cutoff score on the test as a result of the loss of applicants. Assuming there is a positive, mostly linear, relationship between test scores and job performance, the resulting drop in productivity from hiring individuals with lower test scores could be substantial. Moreover, in tight job markets, such a large loss of applicants might mean that many positions would be left vacant, which could result in large decreases in productivity.

3.4.3 Discrimination

RQ_1 explored whether the adverse impact ratios relating to gender, age, ethnicity, sexual orientation, or religious affiliation would be less than .80, thereby indicating adverse impact, if only those who complied with the password request were hired. There was no evidence of adverse impact for gender and age (Table 12). However, in violation of the adverse impact rule, Asians faced an adverse impact ratio of .72, homosexuals faced an adverse impact ratio of .79, and Buddhists and those without a formal religion (but not atheist) both had adverse impact ratios of .77[2]. Therefore, according to the findings from the current study, before traditional selection measures would even be used, employers who

---

[2] We considered the four-fifth rule for analyzing adverse impact because it is the most frequently used method of assessing adverse impact (Bobko & Roth, 2004; Morris & Lobsenz, 2000; Tippins, 2000). However, I also considered a more statistically-grounded alternative that has been suggested in the literature: the Morris and Lobsenz (2000) approach. Following their approach, only the adverse impact ratio for Asians and those without a formal religion (but not atheist) demonstrated *statistically significant* adverse impact. However, there may still be real differences for both of these groups, because Morris and Lonsenz's approach has low power for small groups (Collins & Morris, 2008; Morris & Lobsenz, 2000; Roth, Bobko, & Switzer, 2006), such as the homosexual and Buddhist groups. Given the importance of the four-fifths rule in the practice of personnel selection, I recommend that the results from the more widely-used approach, the four-fifths rule, should be the main focus rather than Morris and Lonsenz's approach.

screen out applicants by requesting SNS passwords may have already created adverse impact for a number of subgroups within the population.

3.4.4 Personality Differences

Based on previous findings and theory, I investigated whether individuals who would agree to divulge their SNS password would have significantly different means on certain personality traits than those who would refuse. As shown in Table 13, those who were willing to divulge their SNS password had a significantly higher mean level of Agreeableness than did those who were unwilling, and had a significantly higher mean level of Conscientiousness than did those who would refuse the password request. Thus, H_3, and H_4, respectively, were supported. In support of H_7, those willing to divulge their password had a significantly lower mean on Psychopathy than did those who were unwilling. Additionally, in support of H_9, those who would agree to divulge their password had a significantly higher mean level of Impression Management than did those who would refuse.

As shown in Table 13, H_2 was not supported because those who would agree to divulge their SNS password did not have a significantly different mean level of Openness to Experience. H_5 and H_8 were also not supported -- those who were willing to release their SNS password did not have a significantly higher mean level of Extraversion or Narcissism. Finally, participants who would comply with the password request did not have a significantly lower mean level of Neuroticism than did those who would refuse, therefore, H_6 was not supported.

Several hypotheses were supported by significant differences on the personality variables (e.g., H_3, H_4, H_7, H_9) between those who would comply with a SNS password request and those who would not. Nonetheless, the effect sizes were generally small (Table 13). Moreover, not all significant personality differences between those who would comply

with versus refuse the password request would necessarily contribute to higher job performance. This is because the link between personality traits and job performance is known to vary depending on the nature of the job (O'Neill, Goffin, & Rothstein, 2013). For example, high Agreeableness scores may be linked to higher performance for customer service jobs and jobs involving teamwork (Hurtz & Donovan, 2000; Mount, Barrick, & Stewart, 1998), but lower performance for managerial jobs high in autonomy (Barrick & Mount, 1993). Thus, on balance, these results suggest that the SNS password request would not likely contribute to any important improvements in the personality scores of the remaining applicant pool.

3.4.5 Limitations and Future Research Directions

I explained the logic behind our hypothetical SNS password request in *The Present Study*. Nonetheless, it may be informative if future researchers consider other approaches and see whether their findings converge with those from the current study. Accordingly, one could survey actual job applicants or obtain retrospective accounts from incumbents, bearing in mind the potential limitations that were outlined earlier for such designs. Researchers conducting these studies could also investigate whether differences emerge as a result different occupational groups' personality subtraits (see O'Neill et al., 2013).

3.5 Conclusions

Overall, our findings demonstrated that asking applicants for their SNS password, and purging the applicants who refuse, may do more harm than good, as more than half of the applicant pool would be lost, adverse impact may occur, and any potential gains in personality scores may have minimal impact on job performance.

3.6 References

Age Discrimination in Employment Act of 1967 (as amended in 1978 and 1986), 29 U.S.C. § 621.

Barrick, M. R., & Mount, M. K. (1993). Autonomy as a moderator of the relationships between the big five personality dimensions and job performance. *Journal of Applied Psychology, 78*, 111-118.

Barrick, M. R., & Mount, M. K. (1996). Effects of impression management and self-deception on the predictive validity of personality constructs. *Journal of Applied Psychology, 81*, 261-272. DOI: 10.107/0021-9010.81.3.261.

Bobko, P. & Roth, P. (2004). The four-fifths rule for assessing adverse impact: An arithmetic, intuitive, and logical analysis of the rule and implications for future research and practice. In J.J. Martocchio (Ed.), *Research in personnel and human resources management* (pp. 177-198). Amsterdam: Elsevier.

CareerBuilder.com. (2009). Forty-five percent of employers use networking sites to research job candidates, CareerBuilder survey finds. *CareerBuilder.com*. Retrieved from http://www.careerbuilder.com/share/aboutus/pressreleasesdetail.aspx?id=pr519&sd=8/19/2009&ed=12/31/2009&siteid=cbpr&sc_cmp1=cb_pr519_&cbRecursionCnt=3&cbsid=0c13dc78541544338f04998d7c0dd165-317523116-JD-5.

Cascio, W.F. & Aguinis, H. (2011). *Applied Psychology in Human Resource Management* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Catano, V. M., Wiesner, W. H., & Hackett, R. D. (2012). *Recruitment and Selection in Canada* (5th ed.). Scarborough, Nelson.

Determann, L. (2012). Social media privacy: A dozen myths and facts. *Stanford Technology Law Review, 7*, 1-14.

Giacalone, R. A. & Duhon, D. (1991). Assessing intended employee behavior in exit interviews. *Journal of Psychology: Interdisciplinary and Applied, 125*, 83-90. DOI: 10.1080/00223980.1991.10543273.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality Psychology in Europe*, Vol. 7 (pp. 7-28). Tilburg, Netherlands: Tilburg University Press.

Hurtz, G. M. & Donovan, J. J. (2000). Personality and job performance: The big five revisited. *Journal of Applied Psychology, 85*, 869-879.

Jackson, D.N. (1999). Personality Research Form Manual. Port Huron, MI: Sigma Assessment Systems.

Jackson, D. N., Paunonen, S. V., Fraboni, M., & Goffin, R. D. (1996). A five-factor versus six-factor model of personality structure. *Personality and Individual Differences, 20*, 33-45.
DOI: 10.1016/0191-8869(95)00143-T.

Jackson, D. N., Wroblewski, V. R., & Ashton, M. C. (2000). The impact of faking on employment tests: Does forced choice offer a solution? *Human Performance, 13*, 371-388. DOI: 10.1207/S15327043HUP1304_3.

Joinson, A. N., Reips, U. D., Buchanan, T., & Schofield, C. B. P. (2010). Privacy, trust, and self-disclosure online. *Human–Computer Interaction*, *25*(1), 1-24. DOI: 10.1080/07370020903586662.

Junglas, I. A., Johnson, N. A., & Spitzmüller, C. (2008). Personality traits and concern for privacy: An empirical study in the context of location-based services. *European Journal of Information Systems*, *17*(4), 387-402. DOI: 10.1057/ejis.2008.29.

Karl, K., Peluchette, J., & Schlaegel, C. (2010). Who's posting Facebook faux pas? A cross-cultural examination of personality differences. *International Journal of Selection and Assessment*, *18*(2), 174-186. DOI: 10.1111/j.1468-2389.2010.00499.x.

Levinson, M. (2011). Social networks: New hotbed for hiring discrimination claims. *Computer World*. Retrieved from http://www.computerworld.com/s/article/9215907/Social_Networks_New_Hotbed_for_Hiring_Discrimination_Claims.

McCrae, R. R., & Costa, P. T., Jr., (2010). *NEO Inventories: Professional manual*. Lutz, FL: Psychological Assessment Resources, Inc.

McFarland, S. (2012, March 21). Job seekers getting asked for Facebook passwords. Retrieved from http://usatoday30.usatoday.com/tech/news/story/2012-03-20/job-applicants-facebook/53665606/1

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological methods*, *17*(3), 437-455. DOI: 10.1037/a0028085.

Morf, C. C., & Rhodewalt, F. (2001). Unraveling the paradoxes of narcissism: A dynamic self-regulatory processing model. *Psychological Inquiry, 12,* 177-196.

Morris, S. B., & Lobsenz, R. E. (2000). Significance tests and confidence intervals for the adverse impact ratio. *Personnel Psychology*, *53*(1), 89-111.

Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance, 11*, 145-165.

Noble, L. & Davis, D. (2014, January 2). Social media and employment law. Retrieved from http://laborandemploymentlaw.ncbar.org/newsletters/leljanuary2014/socialmedia

O'Boyle Jr, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, *97*(3), 557.-579. DOI: 10.1037/a0025679.

O'Neill, T. A., Goffin, R. D., & Rothstein, M. (2013). *Personality and the need for personality-oriented work analysis*. In R. P. Tett & N. Christiansen (Eds.), Handbook of Personality at Work (pp 226-252). Routledge: Taylor & Francis.

Ones, D. S., Dilchert, S., Viswesvaran, C., & Judge, T. A. (2007). In support of personality assessment in organizational settings. *Personnel Psychology*, *60*, 995-1027.

Paulhus, D. L. (1991). *Measurement and control of response bias*. In J. P. Robinson, P. R. Shaver, & L. S. Wrightsman (Eds.), Measures of personality and social psychology attitudes (pp. 17–59). New York: Academic Press.

Paulhus, D.L. & Jones, D.N. (2011, January). *Introducing a short measure of the Dark Triad*. Presented at the meeting of the Society for Personality and Social Psychology, San Antonio.

Paulhus, D. L. & Williams, K. M. (2002). The dark triad of personality: Narcissism, machiavellianism, and psychopathy. *Journal of Research in Personality*, *20*, 556-563. DOI: 10.1016/S0092-6566(02)00505-6.

Roth, P. L., Bobko, P., & Switzer III, F. S. (2006). Modeling the behavior of the 4/5ths rule for determining adverse impact: reasons for caution. *Journal of Applied Psychology*, *91*(3), 507.

Ryan, A. M., Sacco, J. M., McFarland, L. A., & Kriska, S. D. (2000). Applicant self-selection: Correlates of withdrawal from a multiple hurdle process. *Journal of Applied Psychology*, *85*(2), 163-179. DOI: 10.1037/0021-9010.85.2.163.

Stoughton, J. W., Thompson, L. F., & Meade, A. W. (2013). Examining applicant reactions to the use of social networking websites in pre-employment screening. *Journal of Business Psychology.* DOI: 10.1007/s10869-013-9333-6.

Taylor, S. (2006). Seeking secrets in cyberspace. Staffing Management, 2. Retrieved December 1, 2012 from http://www.shrm.org/ema/sm/articles/2006/julysept06cover.asp.

Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703-742. DOI: 10.1111/j.1744-6570.1991.tb00696.x.

Tippins, N. T. (2010). *Adverse impact in employee selection procedures from the perspective of an organizational consultant*. In J.L. Outtz (Ed.), Adverse impact: Implications for organizational staffing and high stakes selection (pp. 201-225). New York: Routledge.

Wicklund, R. A., & Brehm, J. W. (2013). *Perspectives on cognitive dissonance*. Psychology Press.

Williams, K. M., Paulhus, D. L., & Hare, R. D. (2007). Capturing the four-factor structure of psychopathy in college students via self-report. *Journal of Personality Assessment, 88*(2), 205-219. DOI: 10.1080/00223890701268074.

Winter, S. J., Stylianou, A. C. & Giacalone, R. A. (2004). Individual differences in the acceptability of unethical information technology practices: The case of Machiavellianism and ethical ideology. *Journal of Business Ethics, 54*, 279-301. DOI: 10.1007/s10551-004-1772-6.

Chapter 4

4 Summary and Implications

The purpose of this research programme was to increase our understanding of the use of SNSs for selection. To date, very little peer-reviewed research has been conducted on the validity or practical consequences of using SNSs in selection. Addressing these two topics were the primary goals of the current dissertation, with Chapter 2 dedicated to examining the validity of SNSs as instruments for selection, and Chapter 3 dedicated towards addressing some of the practical legal consequences of using SNSs in selection.

Chapter 2 was created to start a preliminary investigation into the construct validity of using SNSs to assess job-relevant information. It was also created to address researchers' and hiring managers' concern that non-job-related SNS information may be used for selection (Brown & Vaughn, 2011; Grasz, 2009; SHRM, 2011). For instance, employers could purposely or accidentally access protected group status information through SNSs, influencing their global rating of an applicant's SNS profile. Therefore, Specific SNS Indicators were created to serve as a way of systematically organizing information that could lead to less biased collection of SNS information and help document information in case of future litigation. They also had the potential to assess SNS information beyond more traditional, global ratings of SNS profiles.

Reliability was investigated first, because a lack of reliability can limit the validity of a selection instrument (Kramer, Bernstein, & Phares, 2014). It was predicted that the Specific SNS Indicators would demonstrate acceptable interrater reliability, which was not the case: most only met the bare minimum requirements for statistical significance. The initial average

ICC (2,*k*) reliabilities for each Specific SNS Indicator was only .53. The average interrater reliability for the Global SNS Rating indicators was only .66.

Other than for the two models that used SNS Integrity scale, the Specific SNS Indicator models had evidence of adequate fit. These findings supported the structural validity of the scales as the empirical dimensionality of the scales reflected the conceptual dimensionality derived from the Taxonomy Expert Raters, and they also supported the unidimensionality of the scales.

There was some support found for the convergent validity of the Specific SNS Indicators. Specific SNS Indicator scales of Extraversion, $r = .376$, $p < .05$, Openness to Experience, $r = .196$, $p < .05$, and Verbal Ability, $r = .223$, $p < .05$, had significant correlations with their equivalent traditional scales. However, these correlations only had small to moderate effect sizes. Also, Agreeableness, Honesty-Humility, Impression Management, and Integrity were not significantly related to their traditional scales. Therefore, based on the current approach of measuring SNS indicators, the constructs of Extraversion, Openness to Experience, and Verbal Ability were the only ones that evidenced the potential for future use as part of the selection process.

The Specific SNS Indicator scales also predicted their traditional scales beyond the Global SNS Rating for Extraversion ($\Delta R^2 = .097$, $p < .05$), and Verbal Ability ($\Delta R^2 = .054$, $p < .05$). The reverse was not the case: the Global SNS Rating did not predict the traditional scales beyond any of the Specific SNS Indicators. As stated in the introduction, the Global SNS Rating was deemed to have a large breadth of coverage, and the Specific SNS Indicators were deemed to have a more specific focus and narrower content domain. Thus, it makes intuitive sense that the Global SNS Rating, with its larger breadth, would have weaker

relations to specific traditional selection scales. Nonetheless, these results offered some preliminary support that methods used to make the information-processing easier for raters could lead to better validity (Heneman, 1986).

The creation of the Specific SNS Indicator scales relied on a mechanical composite strategy, which had better validity evidence than did the Global SNS Rating which relied on the pure clinical strategy. This adds support to previous research which has suggested that mechanical composite strategies are more beneficial than pure clinical strategies because they are less cognitively taxing on the rater (Einhorn, 1972; Ganzach, Kluger, & Klayman, 2000; Sawyer, 1966). Interestingly, as mentioned previously, the interrater reliabilities were higher for the Global SNS Rating than the Specific SNS Indicators. However, reliability does not imply validity, and just because the Global SNS Rating can be measured more consistently between raters does not imply that it is measuring what it is believed to measure (Murphy & Davidshofer, 2005). In addition, the interrater reliabilities were still too low for both the Global SNS Rating indicators and the Specific SNS Indicators to be considered acceptable.

Discriminant validity was also evidenced, because for 24 out of 33 times, the correlations between the Specific SNS Indicator scales and their respective traditional scales were higher than the correlations with the Specific SNS Indicator scales and each of the three response distortion scales (Impression Management, Self-Deception, and PATD). Interestingly, for the traditional scales themselves, only 18 out of 33 were *not* significantly correlated with each of the three response distortion scales. Thus, it appears that Specific SNS Indicator scales may be less subject to response distortion than are traditional scales.

The Global SNS Rating itself was not significantly related to the three response distortion scales, but it was strongly correlated with the Specific SNS Indicator scale for Impression Management ($r = .772$, $p < .05$). This could indicate that the Global SNS Rating may be unduly influenced by how well individuals are believed to manage their impressions. In addition, the Global SNS Rating's convergent validity evidence was problematic. It correlated with two traditional scales, Extraversion ($r = .229$, $p < .05$) and Openness to Experience ($r = .187$, $p < .05$). It also correlated significantly with six Specific SNS Indicator scales (Agreeableness, $r = .333$, $p < .05$; Extraversion, $r = .362$, $p < .05$; Conscientiousness, $r = .746$, $p < .05$; Openness to Experience, $r = .472$, $p < .05$; Impression Management, $r = .772$, $p < .05$; and Verbal Ability, $r = .371$, $p < .05$). These findings cast some doubt on the Global SNS Rating's usefulness as a selection instrument because it is unclear exactly what construct domain it is assessing and it may be affected by the degree to which the rater believes the ratee is managing his/her impressions. Given researchers' and hiring managers' concern about non job-related SNS information being used for selection (Brown & Vaughn, 2011; Grasz, 2009), this lack of clarity around the construct domain of a Global SNS Rating could be even more of a cause for concern. Therefore, given the accumulation of findings in Chapter 2, the Specific SNS Indicator scales may hold more promise for assessing SNS information for job selection purposes in the future. However, at this point in time, neither the Specific SNS Indicators nor the Global SNS Rating can be considered viable options for selecting applicants.

Chapter 3 served as a reminder that even though there may be some evidence of validity for the use of SNSs in job selection, using SNS information for selection could be problematic. The goal of Chapter 3 was to move past validity and consider the practical realities and impacts that SNS selection could have on the workforce. The findings from

Chapter 3 revealed that there could be ramifications for employers who ask their applicants to give up their SNS password. The potential ramifications ranged from drastically reducing the sample pool to discriminating against protected groups.

The first prediction in Chapter 3 was that the proportion of participants who would refuse a potential employer's request for their SNS password would be significantly greater than .35. As predicted, the proportion was significantly higher (.58). That proportion was much larger than expected, however, and the effect size (Cohen's *d* of .96) was quite large. This large proportion seemed to support the Joinson, Reips, Buchanan, and Schofield (2010) claim that most individuals are concerned about their personal information privacy on the internet. Such a large loss of applicants would greatly increase an organization's selection ratio and lower the utility of the selection procedures (Cascio & Aguinis, 2011). In a tight job market, a number of positions would be left vacant, potentially affecting an organization's bottom line.

A research question was posed to explore whether adverse impact occurred based on gender, ethnicity, sexual orientation, and/or religious affiliation. Although there was no adverse impact for gender and age (impact ratios were above .80), Asians (impact ratio of .72), homosexuals (impact ratio of .79), Buddhists (impact ratio of .77), and those without a formal religion (but not atheist; impact ratio of .77) had all been adversely impacted. If employers screened out applicants by requesting their SNS passwords, they could be contributing to discrimination.

Although some of the above differences may be considered small from a scientific viewpoint, it is important to acknowledge that from a legal and practical point of view, failure to meet the .80 adverse impact cutoff could lead to further investigation by the U.S.

Equal Employment Opportunity Commission (in the U.S.) or the Canadian Human Rights Tribunal (in Canada). As a result, an organization could eventually be open to legal action on behalf of the applicant (see Catano, Wiesner, & Hackett, 2012).

Finally, if the reduction in the applicant pool led to only those with superior job-relevant personality traits remaining, then that would help justify such a large proportion of applicants withdrawing from the pool. Those who were willing to divulge their SNS password did have significantly higher scores than those who were unwilling for Agreeableness (3.75 versus 3.65, respectively), for Conscientiousness (3.54 versus 3.41, respectively), for Impression Management (3.74 versus 3.52, respectively), and significantly lower scores for Psychopathy (2.06 versus 2.15, respectively). However, those who were willing to divulge their SNS password were not significantly different from those who were unwilling for Openness to Experience, Extraversion, or Narcissism. Also, the benefits of having a larger proportion of applicants with high Agreeableness, high Conscientiousness, and low Psychopathy were limited for two primary reasons. 1) The effect sizes were small. The highest Cohen's *d* for these three traits was .22. 2) The link between personality traits and job performance varies based on the nature of the job (O'Neill, Goffin, & Rothstein, 2013). For example, having employees with high Agreeableness may not be beneficial if their job requires them to work in isolation. Additionally, individuals with higher Impression Management scores remained in the applicant pool. High levels of Impression Management have been linked to low levels of Honesty-Humility, which can have negative consequences for organizations, such as increased counterproductive work behaviours (Ashton & Lee, 2007; Bourdage, Wiltshire, & Kibeom Lee, 2015). Therefore, given these limitations, the SNS password request is unlikely to improve the personality scores for the remaining applicant pool. There may be other, superior ways of collecting SNS information for

selection, such as asking applicants to volunteer that information or friending applicants. However, the effect of these methods needs to be researched (see the Future Research Directions section).

In summary, the findings in these two chapters demonstrated that the validity evidence for SNS selection information is quite limited at this time. The convergent validity for the Global SNS Rating was not well supported, as it seemed to tap a variety of constructs. As for the Specific SNS Indicators, although there was some support for their use over a Global SNS Rating, they only had convergent validity evidence for Extraversion, Openness to Experience, and Verbal Ability. Evidence of convergent validity for more human attributes would need to accumulate if the Specific SNS Indicators' usefulness as selection device were to be supported. In addition, stronger convergent validity effect sizes would need to be reported. Also, the findings that certain ways of collecting SNS information can lead to negative ramifications is alarming. Asking applicants for their SNS password led to a drastic reduction of the applicant pool, which can affect an organization's bottom line. It also could lead to discrimination through adverse impact, potentially leading to legal action against an organization.

4.1 Future Research Directions

I would suggest that hiring managers refrain from using SNSs as part of job selection until more validation research can be conducted on SNSs and until the impact of gathering SNS information can be more fully evaluated. First, future researchers should investigate the criterion-related validity information for a Global SNS Rating and the Specific SNS Indicator scales. Second, the models and findings from Chapter 2 should be cross-validated with an independent sample to provide a better understanding of how they generalize across other

samples. Third, the prediction that Specific SNS Indicators may lead to less discrimination than a Global SNS Rating remains to be tested. Fourth, the effect of an employer's request for SNS passwords should be investigated with actual job applicants and job incumbents to build upon Chapter 3's findings. In fact, researchers and practitioners should explore the impact that other methods of collecting SNS information can have on applicants and employers. For instance, if the password request were voluntary, rather than a requirement, would that lead to different results or outcomes? If applicants volunteered their SNS information, perhaps they would believe that their privacy had not been breached. Also, what would result from having an employer friend applicants on Facebook? Does this action change applicants' and employees' perceptions of the organization? Fifth, although Chapter 3 could generalize to any SNS platform, Chapter 2 was specific to Facebook. Future researchers should see if reliable and valid Specific SNS Indicators can be created for other SNS platforms (e.g., Twitter, LinkedIN, Instagram).

4.2 Conclusion

As mentioned in the introduction, 65% or more of companies are using SNSs to screen job applications (Levinson, 2011; Preston, 2011). Therefore, a large number of employers are using SNSs in selection and will likely continue to do so. As discussed previously, there are a number of reasons why this might not be an ideal situation, most notably the use of SNS information that is not job-related (Brown & Vaughn, 2011; Grasz, 2009; SHRM, 2011) and the concern about breaching employment discrimination laws (SHRM, 2011). Despite considerable effort to create a method for assessing SNS information in a way that is more likely to capture job-related information, the validity evidence for the Specific SNS Indicators was still fairly limited. It is recommended that practitioners avoid

the use of SNS information in selection decisions until further validity evidence can be accumulated, and stronger validity evidence reported.

4.3 References

Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review, 11*, 150–166.

Bourdage, J. S., Wiltshire, J., & Lee, K. (2015). Personality and workplace impression management: Correlates and implications. *Journal of Applied Psychology*, *100,* 537-546.

Brown, V. R., & Vaughn, E. D. (2011). The writing on the (facebook) wall: The use of social networking sites in hiring decisions. *Journal of Business and Psychology, 26,* 219-225.

Cascio, W.F. & Aguinis, H. (2011). *Applied Psychology in Human Resource Management* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Catano, V.M., Wiesner, W.H., Hackett, R.D., & Methot, L.L. (2010). *Recruitment and Selection in Canada* (4th Ed.). Toronto, ON: Thomson Nelson.

Murphy, K.R., & Davidshofer, C. O. (2005). *Psychological testing: Principles and applications* (6th ed.). Upper Saddle River, N.J.: Pearson/Prentice Hall.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. *Organizational behaviour and human performance*, *7*(1), 86-106.

Ganzach, Y., Kluger, A. N., & Klayman, N. (2000). Making decisions from an interview: Expert measurement and mechanical combination. *Personnel Psychology*, *53*(1), 1-20.

Grasz, J. (2009, August 19). Forty-five percent of employers use social networking sites to research job candidates, career builder survey finds. *CareerBuilder.com.* Retrieved from http://www.careerbuilder.com/Article/CB-1337-Interview-Tips-More-Employers-Screening-Candidates-via-Social-Networking-Sites/

Joinson, A. N., Reips, U. D., Buchanan, T., & Schofield, C. B. P. (2010). Privacy, trust, and self-disclosure online. *Human–Computer Interaction*, *25*(1), 1-24.

Kramer, G. P., Bernstein, D. A., & Phares, V. (2014). *Introduction to clinical psychology*. Upper Saddle River, NJ: Prentice-Hall.

Levinson, M. (2011). Social networks: New hotbed for hiring discrimination claims. *Computer World*. Retrieved from http://www.computerworld.com/s/article/9215907/Social_Networks_New_Hotbed_for_Hiring_Discrimination_Claims

O'Neill, T. A., Goffin, R. D., & Rothstein, M. (2013). Personality and the need for personality-oriented work analysis. *Handbook of personality at work, Taylor & Francis, Routledge*, 226-252.

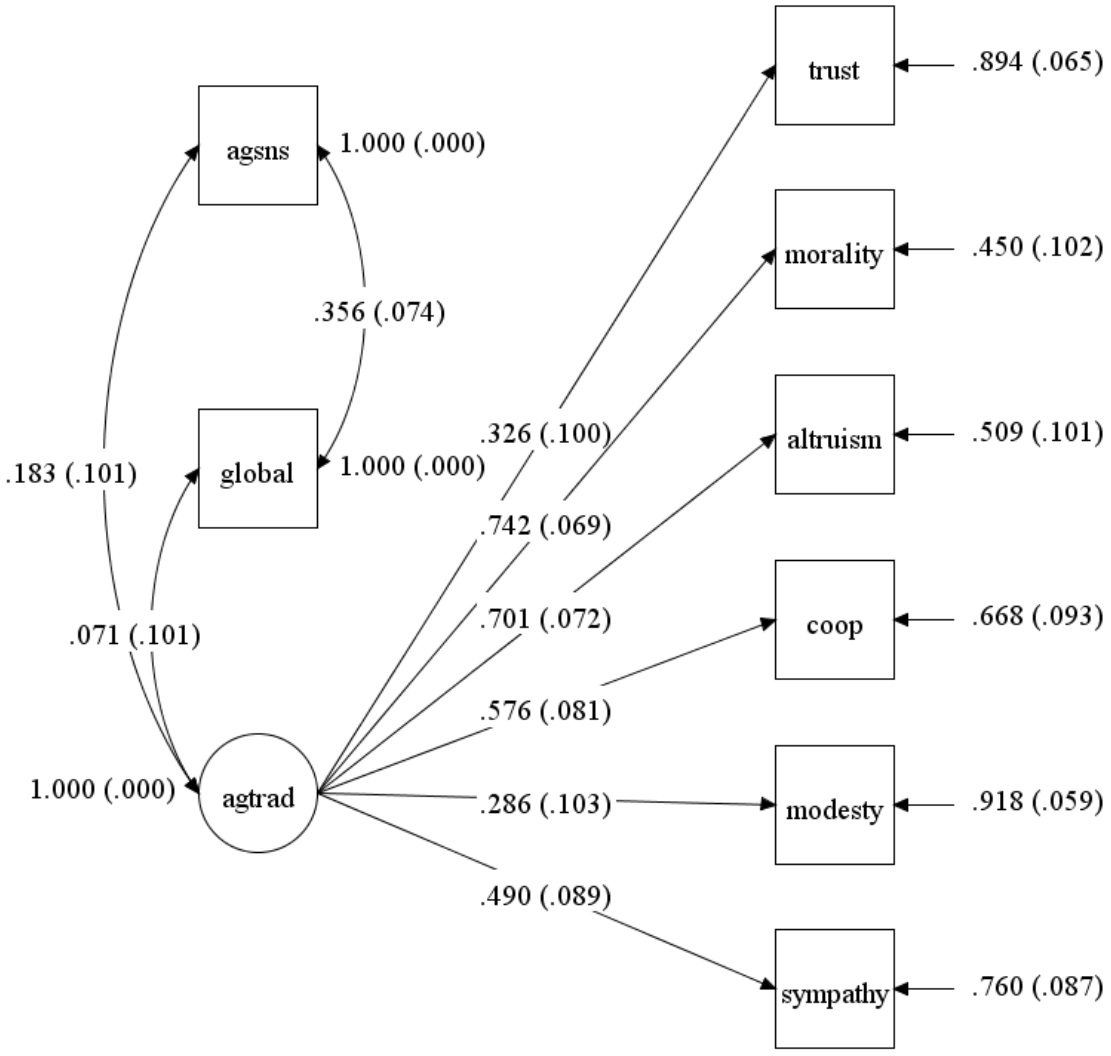
Preston, J. (2011). Social media history becomes a new job hurdle. *New York Times*.

Sawyer, J. (1966). Measurement and prediction, clinical and statistical. *Psychological bulletin*, *66*(3), 178.

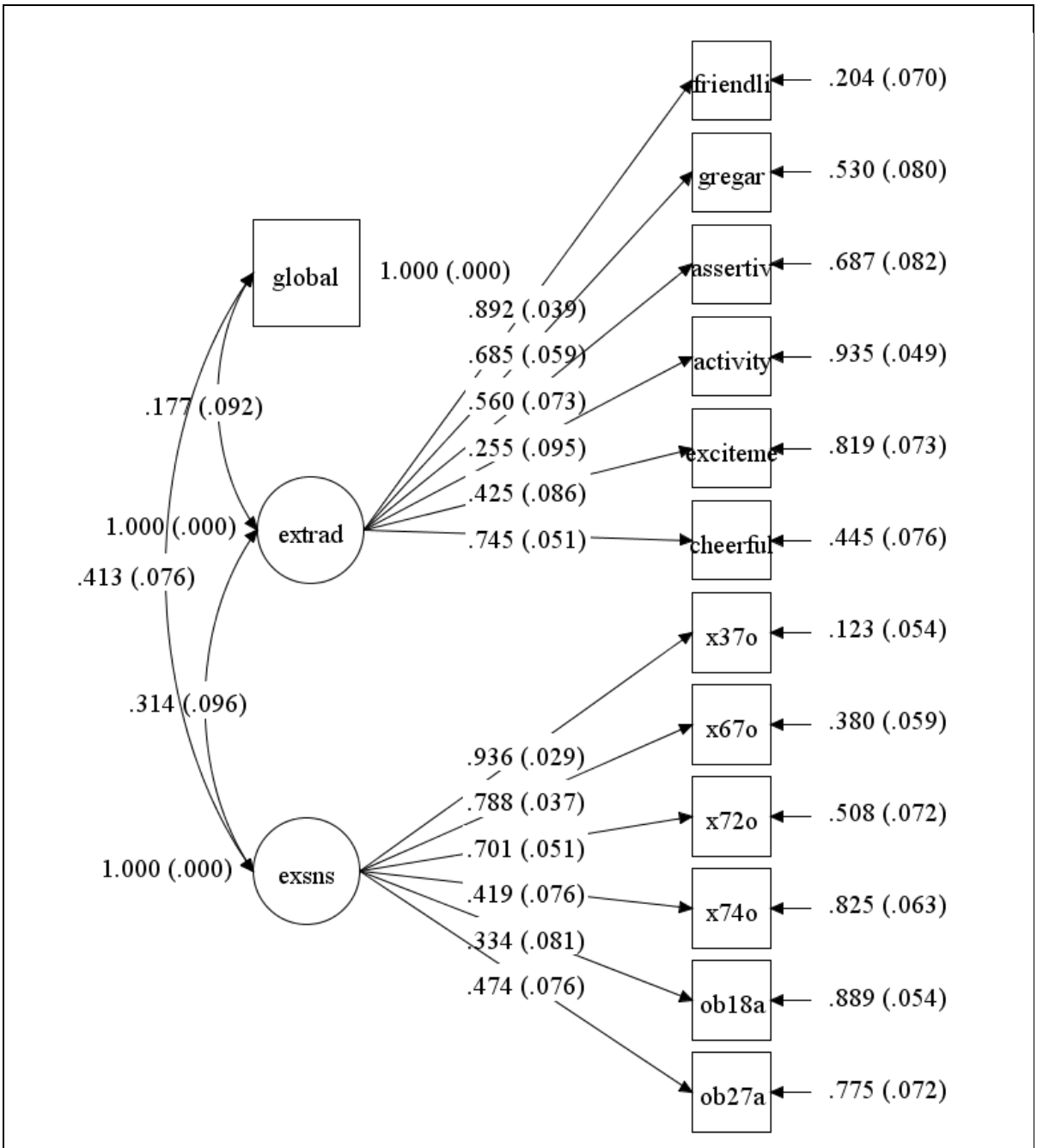
Society for Human Resource Management (2011, October). *SHRM Research Spotlight: Social Networking Websites and Online Search Engines for Screening Job Candidates.* Retrieved from http://www.shrm.org/Research/SurveyFindings/Articles/Documents/PART%202_Social%20Media%20Flier_FINAL.pdf

Appendices

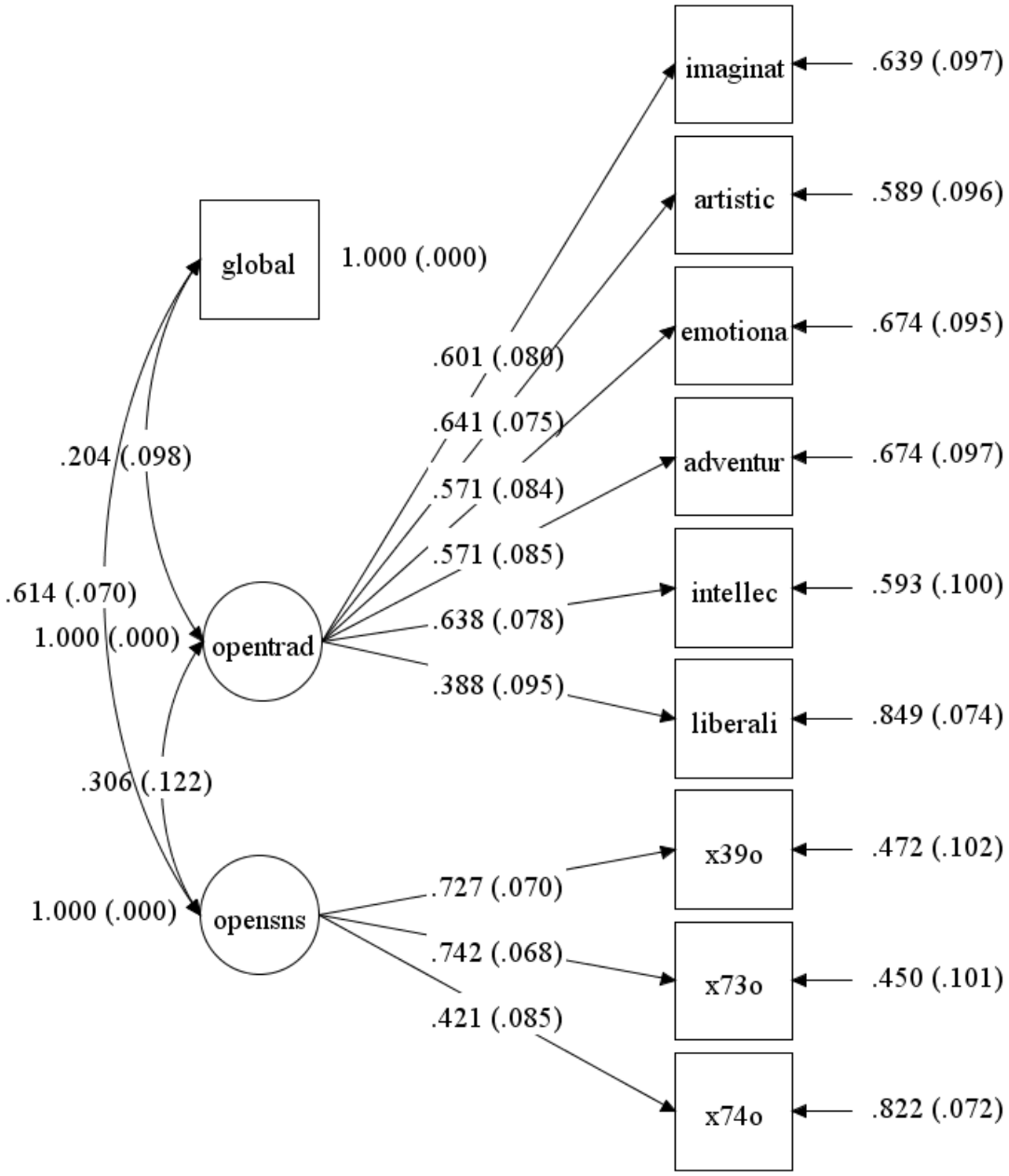
Appendix A: Additional Confirmatory Factor Analysis Figures



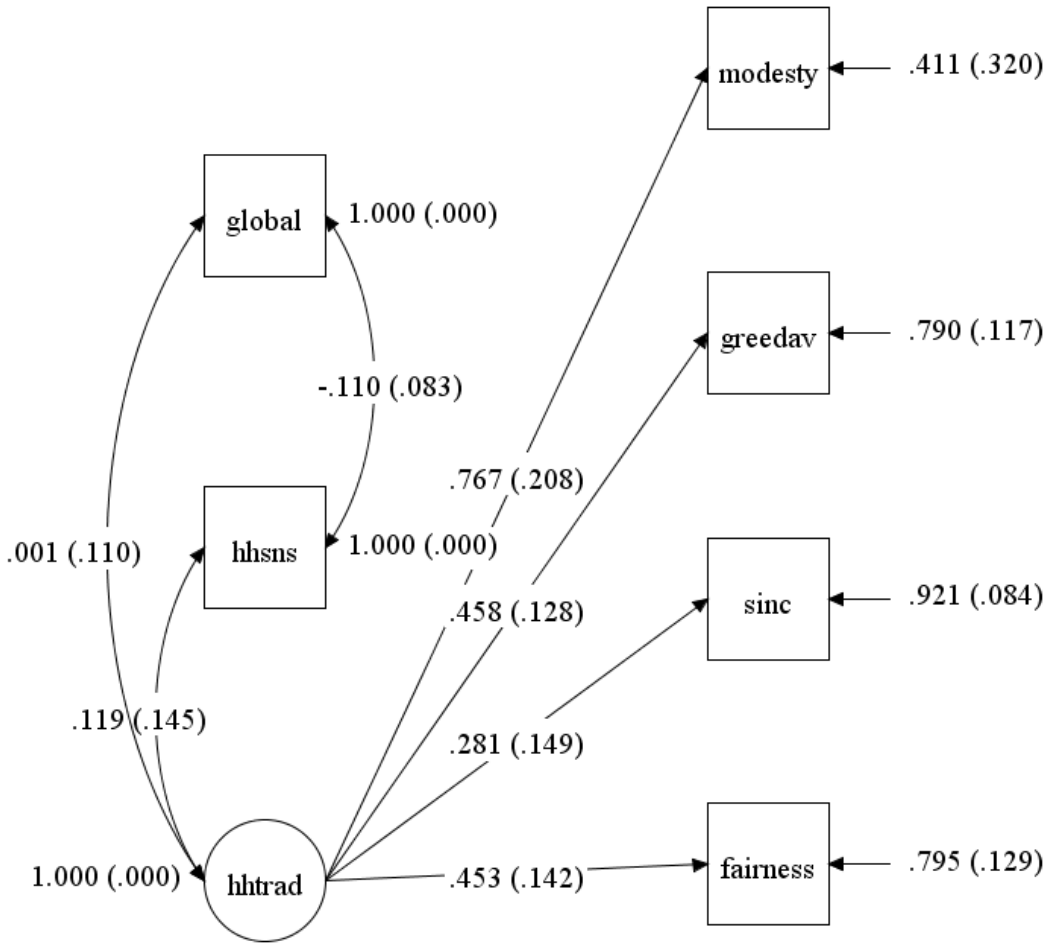
Confirmatory factor analysis for Agreeableness. agtrad = traditional scale of Agreeableness; agsns = Specific SNS Indicator scale of Agreeableness; global = Global SNS Rating; coop = cooperation; The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



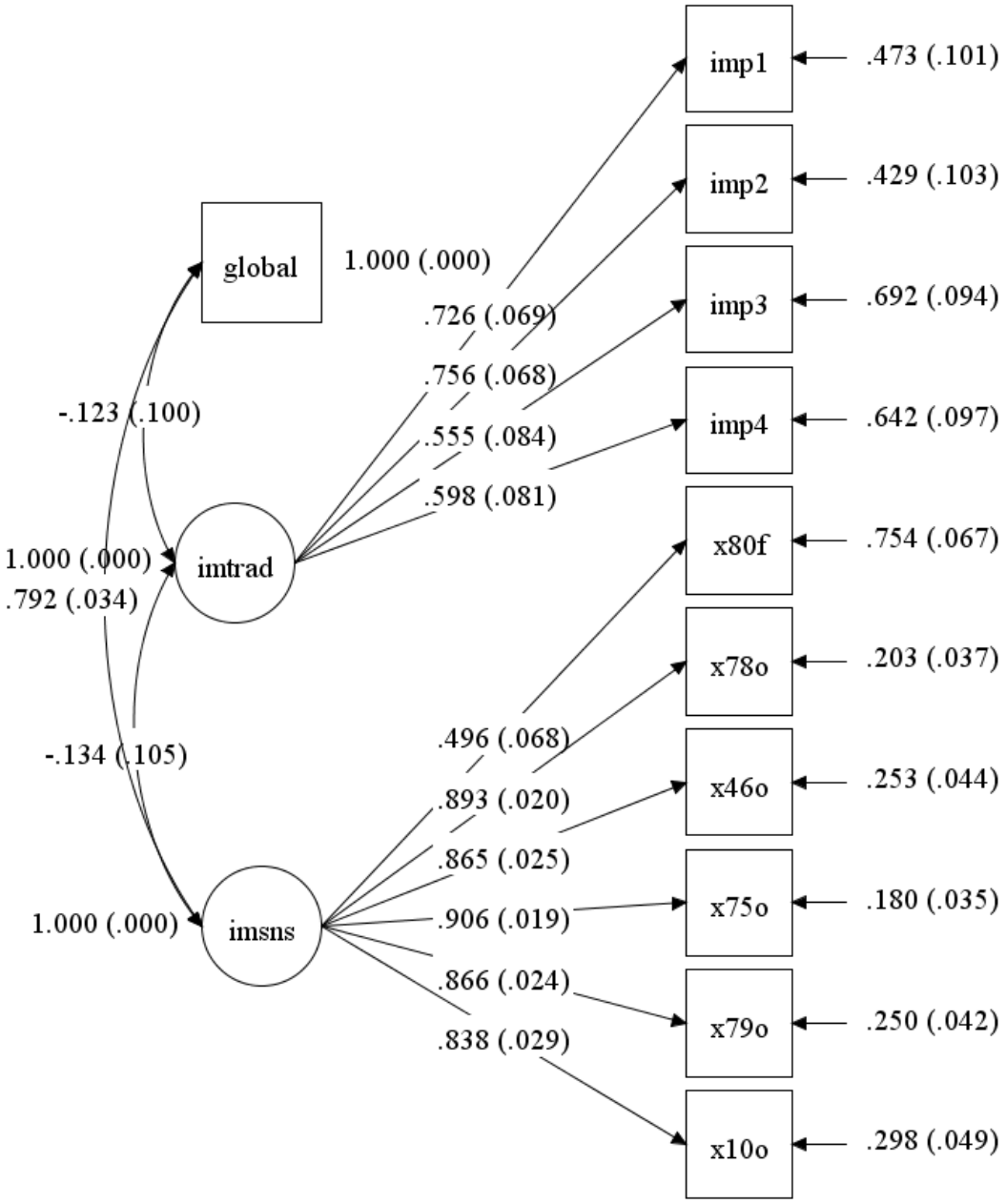
Confirmatory factor analysis for Extraversion. extrad = traditional scale of Extraversion; exsns = Specific SNS Indicator scale of Extraversion; global = Global SNS Rating; friendli = friendliness; gregar = gregariousness; assertiv = assertiveness; activity = activity level; exciteme = excitement-seeking; cheerful = cheerfulness. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



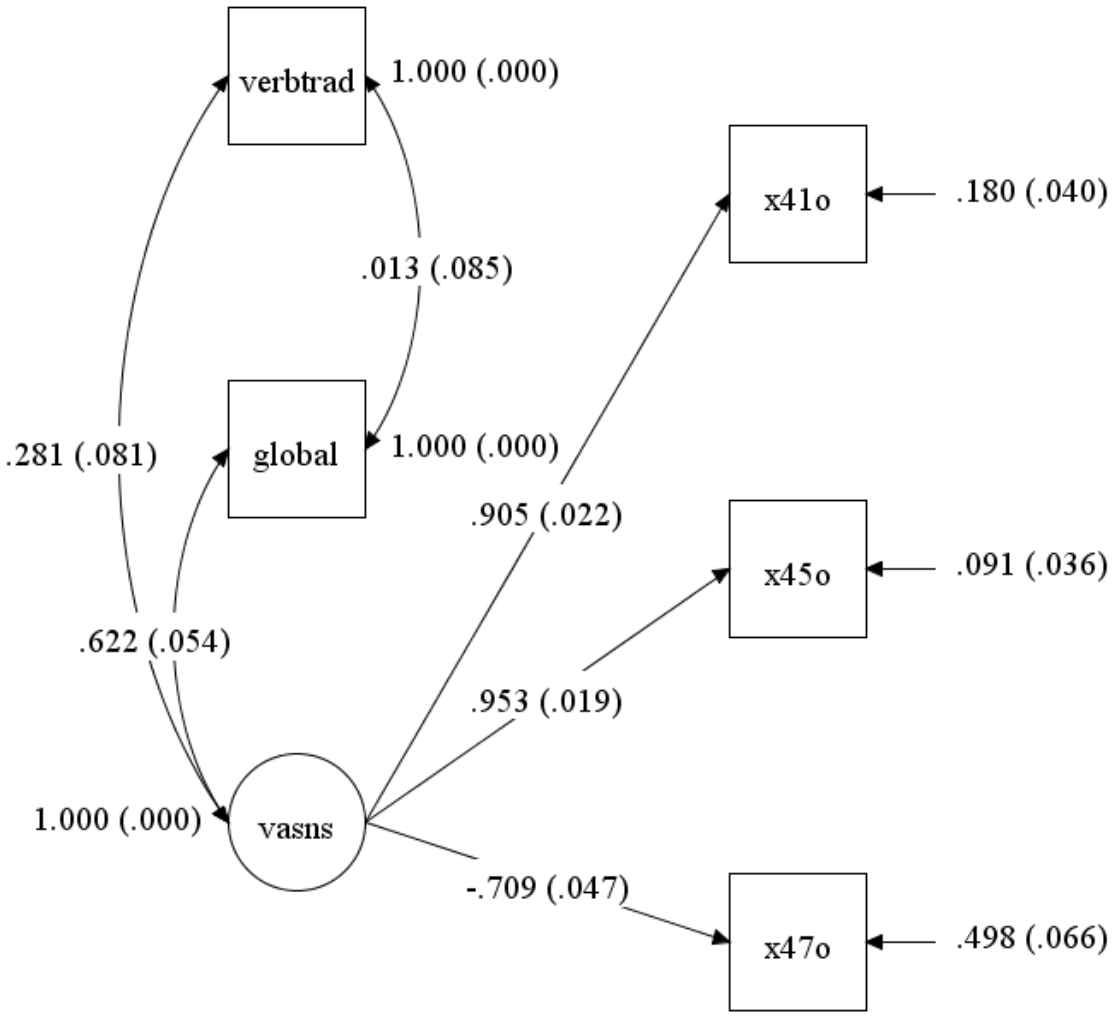
Confirmatory factor analysis for Openness to Experience. opentrad = traditional scale of Openness to Experience; opensns = Specific SNS Indicator scale of Openness to Experience; global = Global SNS Rating; imaginat = imagination; emotiona = emotionality; adventur = adventurousness; intellec = intellect; liberali = liberalism. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



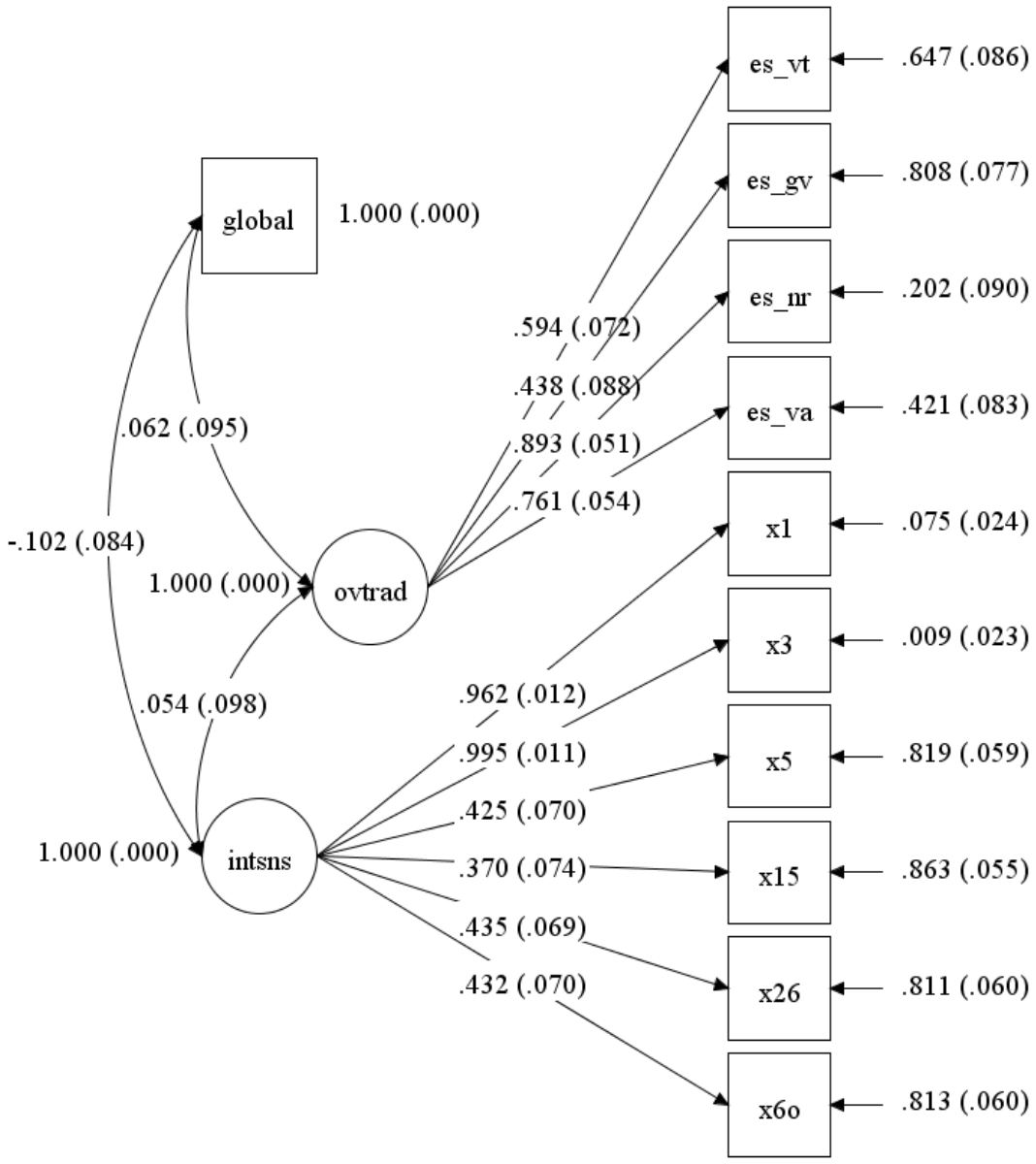
Confirmatory factor analysis for Honest-Humility. hhtrad = traditional scale of Honesty-Humility; hhsns = Specific SNS Indicator scale of Honesty-Humility; global = Global SNS Rating; greedav = greed-avoidance; sinc = sincerity. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



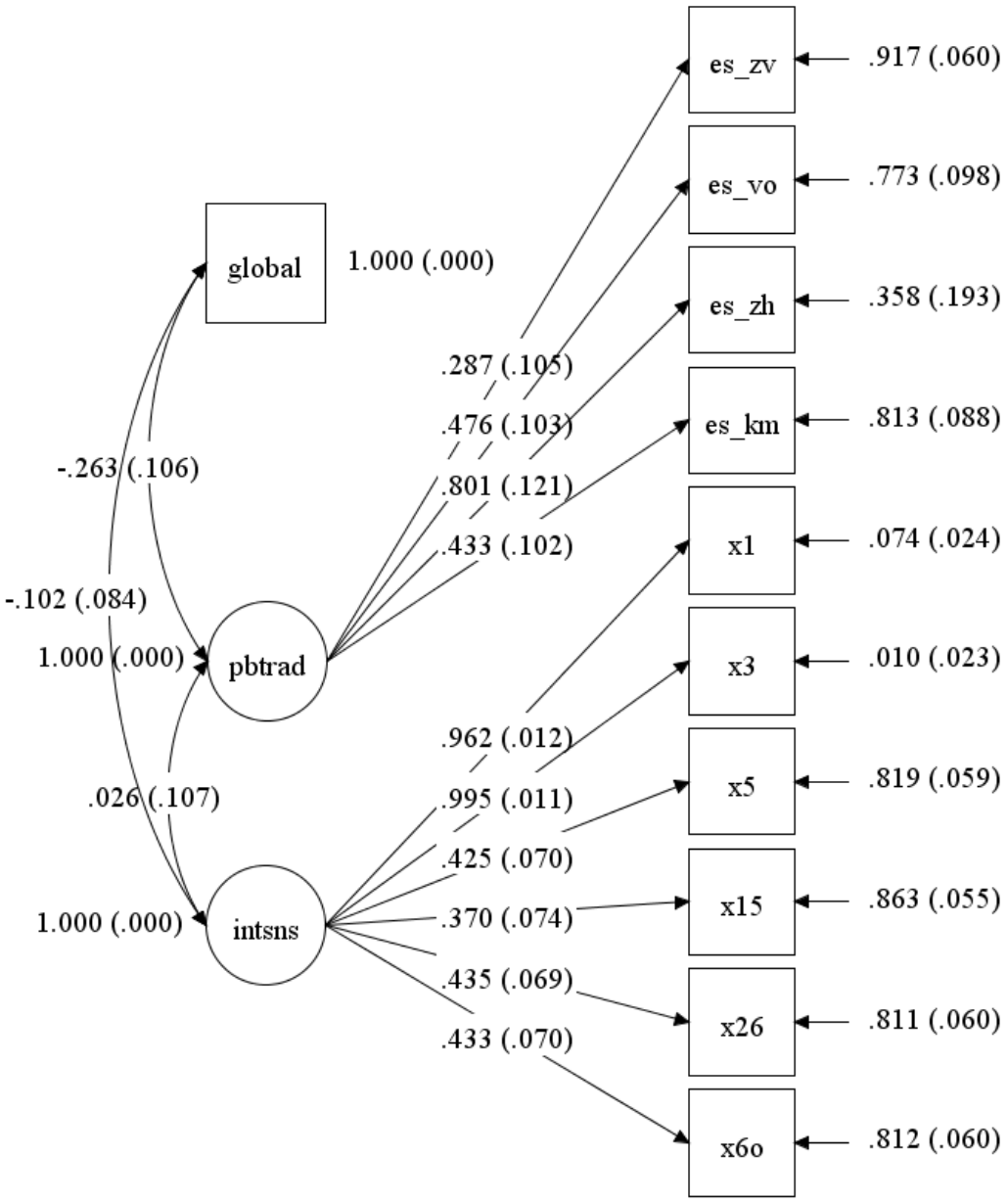
Confirmatory factor analysis for Impression Management. imtrad = traditional scale of Impression Management; imsns = Specific SNS Indicator scale of Impression Management; global = Global SNS Rating; imp1, imp2, imp3, and imp4 refer to the four Impression Management parcels. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



Confirmatory factor analysis for Verbal Ability. verbtrad = traditional scale of Verbal Ability; vasns = Specific SNS Indicator scale of Verbal Ability; global = Global SNS Rating. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



Confirmatory factor analysis for Overt Integrity. Ovtrad = traditional scale of Overt Integrity; intsns = Specific SNS Indicator scale of Integrity; global = Global SNS Rating; es_vt = trust; es_gv = perceived norms; es_nr = (non)rationalization; es_va = intentions/fantasies. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.



Confirmatory factor analysis for Personality-Based Integrity. Pbtrad = traditional scale of Personality-Based Integrity; intsns = Specific SNS Indicator scale of Integrity; global = Global SNS Rating; es_zv = reliability; es_vo = (low)stimulus seeking; es_zh = (low)manipulativeness; es_km = trouble avoidance. The standardized estimates are presented with the standard errors in parentheses. Values of 1.000 represent that the variances were fixed at 1 to give a standardized solution.

Appendix B: Ethics Approval Forms

Department of Psychology The University of Western Ontario
Room 7418 Social Sciences Centre,
London, ON, Canada N6A 5C1
Telephone: (519) 661-2067Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	**12 12 08**	**Approval Date**	**12 12 14**
Principal Investigator	**Rick Goffin/Travis Schneider**	**End Date**	**13 12 10**
Protocol Title	**Expert ratings of social networking site indicators**		
Sponsor	**n/a**		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

[redacted]

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PREB are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson Albert Katz (Department Chair), Steve Lupker, and TBA (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

Department of Psychology The University of Western Ontario
Room 7418 Social Sciences Centre,
London, ON, Canada N6A 5C1
Telephone: (519) 661-2067Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	**13 02 10**	**Approval Date**	**13 02 13**
Principal Investigator	**Rick Goffin/Travis Schneider**	**End Date**	**14 02 12**
Protocol Title	**Social networking and job selection**		
Sponsor	**n/a**		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

[redacted]

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PREB are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson Albert Katz (Department Chair), Steve Lupker, and TBA (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

Department of Psychology The University of Western Ontario
Room 7418 Social Sciences Centre,
London, ON, Canada N6A 5C1
Telephone: (519) 661-2067Fax: (519) 661-3961

Use of Human Subjects - Ethics Approval Notice

Review Number	12 09 05	**Approval Date**	12 09 13
Principal Investigator	Travis Schneider/MASS TESTING (dept)	**End Date**	13 04 30
Protocol Title	Mass testing		
Sponsor	n/a		

This is to notify you that The University of Western Ontario Department of Psychology Research Ethics Board (PREB) has granted expedited ethics approval to the above named research study on the date noted above.

The PREB is a sub-REB of The University of Western Ontario's Research Ethics Board for Non-Medical Research Involving Human Subjects (NMREB) which is organized and operates according to the Tri-Council Policy Statement and the applicable laws and regulations of Ontario. (See Office of Research Ethics web site: http://www.uwo.ca/research/ethics/)

This approval shall remain valid until end date noted above assuming timely and acceptable responses to the University's periodic requests for surveillance and monitoring information.

During the course of the research, no deviations from, or changes to, the protocol or consent form may be initiated without prior written approval from the PREB except when necessary to eliminate immediate hazards to the subject or when the change(s) involve only logistical or administrative aspects of the study (e.g. change of research assistant, telephone number etc). Subjects must receive a copy of the information/consent documentation.

Investigators must promptly also report to the PREB:
a) changes increasing the risk to the participant(s) and/or affecting significantly the conduct of the study;
b) all adverse and unexpected experiences or events that are both serious and unexpected;
c) new information that may adversely affect the safety of the subjects or the conduct of the study.

If these changes/adverse events require a change to the information/consent documentation, and/or recruitment advertisement, the newly revised information/consent documentation, and/or advertisement, must be submitted to the PREB for approval.

Members of the PREB who are named as investigators in research studies, or declare a conflict of interest, do not participate in discussion related to, nor vote on, such studies when they are presented to the PREB.

[redacted]

Clive Seligman Ph.D.

Chair, Psychology Expedited Research Ethics Board (PREB)

The other members of the 2012-2013 PREB are: Mike Atkinson (Introductory Psychology Coordinator), Rick Goffin, Riley Hinson Albert Katz (Department Chair), Steve Lupker, and TBA (Graduate Student Representative)

CC: UWO Office of Research Ethics

This is an official document. Please retain the original in your files

Curriculum Vitae

Travis Schneider

Degrees and Certifications

Associate Certified Coach, International Coaching Federation (September, 2015 - present)
Training Through College of Executive Coaching

Ph.D., Industrial/Organizational Psychology (September, 2010 - present)
University of Western Ontario
Dissertation Title: Social Networking Sites and Personnel Selection: An Initial Validity Assessment

M.Sc., Industrial/Organizational Psychology (September, 2008 - August, 2010)
University of Western Ontario
Thesis Title: Perceived Ability to Deceive as a Moderator of the Criterion-Related Validity of Pre-Employment Personality Testing

B.A., **Psychology (September, 2003 - April, 2008)**
University of British Columbia

Work Experience

Consultant (September, 2014 - present)
Employer: Leadership Success Group
Description of role: Through leadership coaching, worked one-on-one with clients to help them achieve their full leadership potential.

Consultant (September, 2014 - present)
Employer: Calgary Career Counselling
Description of role: Led a team of consultants and served as a career counsellor to help clients better understand their career path and fulfill their career goals.

Consultant (August, 2014 - present)
Employer: Work EvOHlution
Description of role: Engaged in business development and offered technical expertise in creating, implementing, analyzing, and validating organizational development and selection assessments within organizations.

Consultant (January, 2013 - August, 2014)
Employer: Research Psychologists Press
Description of role: Updated personality and vocational interests assessments, the validation of these assessments, and documenting any findings (i.e., for publication).

Work Experience (continued)

Consulting Internship (January, 2013 - April, 2013)
As part of an awarded MITACS scholarship, partnered with the university and Research Psychologists Press to update an existing personality assessment using polytomous IRT and classical test theory.

Consultant (May, 2012 - present)
Employer: Prestige Hotels and Resorts
Description of role: Coordinated, implemented, and analyzed/described the results of an ongoing work engagement study to help the organization retain quality employees.

Mass Testing Research Coordinator (September, 2010 - August, 2014)
Employer: University of Western Ontario
Involves: Coordinated with students and faculty, and created and maintained the Mass Testing survey which is used by the psychology department for pre-selection and for many other uses.

Teaching Assistant (September, 2008 - present)
Employer: University of Western Ontario
Courses: The Human Mind (Psych 2010), Introduction to I/O Psychology (Psych 2660B), Teams and Work Groups in Organizations (Psych 3694), Understanding and Managing Behaviour in Organizations (Psych 3690G), Human Sexuality (Psych 2075).

Research Assistant (August, 2008 - present)
Employer: Dr. Richard Goffin
Area: Industrial/Organizational Psychology
Key Contributions: Assembled videos of trainers and set up a study whereby participants would rate the job performance of the trainers. In another project, made use of structural equation modelling to analyze data from an assessment centre.

Technical Consultant (August, 2007 - April, 2008)
Employer: Dr. Todd Handy and Dr. Jonathan Schooler
Area: Social and Cognitive Psychology
Key Contributions: Set up the technological and scientific equipment for the lab and programmed the ePrime software to execute the research designs.

Teaching Experience

Lab Instructor (September, 2009 - April, 2010; September, 2013 - April, 2014)
Employer: University of Western Ontario
Lecturer: Riley Hinson
Lab Course: Research Methods and Statistical Analysis in Psychology (Psych 2820E)

Publications

Schneider, T. J., Goffin, R. D., & Daljeet, K. N. (2015). "Give us your social networking passwords": Implications for personnel selection and personality. *Personality and Individual Differences, 73*, 78-83. doi:10.1016/j.paid.2014.09.026

- *Found that 57.87% of job applicants would remove themselves from the selection process when asked for their social networking site password by a hypothetical employer*
- *Discovered that the request also caused adverse impact for several minority groups*

McLarnon, M. J. W., Carswell, J. J., & Schneider, T. J. (2014). A case of mistaken identity? Latent profiles in vocational interests. *Journal of Career Assessment, 23*, 166-185. doi: 10.1177/1069072714523251

- *Investigated the RIASEC model of career interests using a person-centered approach instead of the traditional variable-centered approaches*
- *Found eight qualitatively and quantitatively distinct types of individuals who complete career interest surveys*

O'Neill, T. A., McLarnon, M. J. W., Schneider, T. J., & Gardner, R.C. (2013). Current misuses of multiple regression for investigating bivariate hypotheses: An example from the organizational domain. *Behavior Research Methods, 46*, 798-807. doi: 10.3758/s13428-013-0407-1

- *Examined the potential misuse of multiple regression in measuring bivariate relations*
- *In 44% of the articles reviewed, the results would have been different if correlations were used rather than multiple regression*

Schneider, T. J., & Goffin. (2012). Perceived ability to deceive and incremental prediction in pre-employment personality testing. *Personality and Individual Differences, 52*, 806-811. doi: 10.1016/j.paid.2012.01.015

- *Dealt with the perennial problem of faking (deliberate attempts to respond to personality items in a way that advances one's position) when used for hiring purposes*
- *Developed and examined a new measure that may be related to faking*