Journal of Management
Vol. 42 No. 7, November 2016 1811–1835
DOI: 10.1177/0149206313515524

Social Media for Selection? Validity and Adverse Impact Potential of a Facebook-Based Assessment

Chad H. Van Iddekinge
Florida State University
Stephen E. Lanivich
Old Dominion University
Philip L. Roth
Clemson University
Elliott Junco
Independent Consultant

Recent reports suggest that an increasing number of organizations are using information from social media platforms such as Facebook.com to screen job applicants. Unfortunately, empirical research concerning the potential implications of this practice is extremely limited. We address the use of social media for selection by examining how recruiter ratings of Facebook profiles fare with respect to two important criteria on which selection procedures are evaluated: criterion-related validity and subgroup differences (which can lead to adverse impact). We captured Facebook profiles of college students who were applying for full-time jobs, and recruiters from various organizations reviewed the profiles and provided evaluations. We then followed up with applicants in their new jobs. Recruiter ratings of applicants' Facebook information were unrelated to supervisor ratings of job performance (rs = −.13 to –.04), turnover intentions (rs = −.05 to .00), and actual turnover (rs = −.01 to .01). In addition, Facebook ratings did not contribute to the prediction of these criteria beyond more traditional predictors, including cognitive ability, self-efficacy, and personality. Furthermore, there was evidence of subgroup difference in Facebook ratings that tended to favor female and White applicants. The overall results suggest that organizations should be very cautious about using social media information such as Facebook to assess job applicants.

Acknowledgments: This article was accepted under the editorship of Deborah E. Rupp. An earlier version of this article was presented at the 73rd Annual Meeting of the Academy of Management, Orlando, Florida, August 2013. The authors thank Myrna Hoover and the Florida State University Career Center for helping us recruit participants and for providing a site for the data collection. We also thank In-Sue Oh for his helpful comments on an earlier version of the article.

Corresponding author: Chad H. Van Iddekinge, College of Business, Florida State University, 821 Academic Way, P.O. Box 3061110, Tallahassee, FL 32306-1110, USA.

E-mail: cvanidde@fsu.edu

Keywords: *social media; Facebook; staffing; personnel selection; criterion-related validity; adverse impact; subgroup differences*

Recent reports suggest that many organizations are using the Internet to search for information about job applicants (e.g., Deschenaux, 2010; Levinson, 2010; Preston, 2011). This includes reviewing applicants' personal websites and blogs and performing web searches (e.g., via Google) on an applicant's name. The type of Internet information that has received the most attention is social media (SM), including SM platforms such as Facebook and Twitter. Several recent surveys have inquired about organizations' use of SM information for recruitment and selection. For example, a survey of 825 staffing professionals revealed that 73% of those surveyed use SM platforms such as Facebook to recruit applicants (Levinson, 2010). Furthermore, about one third of respondents indicated that they always check such SM information when vetting applicants. In a CareerBuilder survey of 2,303 hiring managers and human resources professionals, 37% of respondents reported they research applicants' SM information, and another 11% plan to start doing so in the future (Grasz, 2012). In addition, 29% of managers indicated that they had not hired applicants due to what they learned from applicants' SM information.

Several possible benefits of using SM and other Internet information for staffing have been noted. For example, unlike selection procedures such as interviews and work sample tests that can be costly and time-consuming to develop and administer, SM information already exists (i.e., it does not have to be developed), is free to view, and does not require applicants to be present. Furthermore, interviews and surveys with hiring officials suggest that many believe SM information can be used to help predict applicants' future performance. For example, one employer noted that "if a person runs competitive races, that is valuable information about how ambitious and results-oriented they are" (Willis, 2006). Another employer suggested that the types of information applicants post on SM platforms "demonstrate the level of an applicant's judgment and give employers insight into their personality" (Palank, 2006). And another suggested that "there is nothing that screams more accurately who you are than a Facebook page" (Meinert, 2011). Other reports indicate that some hiring officials believe SM information can provide information concerning person–organization fit. For instance, in the CareerBuilder survey mentioned above, one of the main reasons why respondents use SM is "to see if the candidate is a good fit for the company culture" (Grasz, 2012). Similarly, a Society for Human Resources Management (SHRM, 2008) survey found that 26% of staffing professionals use SM information to assess whether applicants would be a good fit for the organization. The focus on fit may indicate that hiring officials are using SM to help make predictions about whether applicants would be likely to quit or to remain with the company for the longer-term.

Some scholars also appear to believe that SM may hold promise for use in the selection process (e.g., Elzweig & Peeples, 2009; Karl, Peluchette, & Schlaegel, 2010; Kluemper & Rosen, 2009; Wilson, Gosling, & Graham, 2012). For instance, it has been suggested that SM information could be used to assess personality (Back et al., 2010; Kluemper & Rosen, 2009; Vazire & Gosling, 2004). As an example, the number of "friends" in an applicant's social network, and the comments these friends post about the applicant, may indicate traits such as

agreeableness and extraversion. The creativity with which applicants' arrange their Facebook profile, and the types of activities and quotes they post, may reflect their openness to experience. Other researchers have suggested that applicants who post inappropriate information online may have related problems at work, such as low conscientiousness or lack of integrity (Karl et al., 2010). Finally, some authors have suggested that organizations that do not check publically available SM information may be negligent in terms of failing to fully vet applicants (e.g., Elzweig & Peeples, 2009).

Despite the possible benefits, it is unclear whether assessments of SM information would satisfy certain standards for valid and legally defensible selection procedures. For example, the *Uniform Guidelines* (Equal Employment Opportunity Commission [EEOC], 1978) discuss the importance of having evidence of criterion-related validity prior to using tests as a basis for personnel decisions. Similarly, the *Principles* (Society for Industrial and Organizational Psychology [SIOP], 2003) note the primary inference of concern in an employment context is that test scores predict subsequent work behavior. Unfortunately, there is very little evidence for the criterion-related validity of inferences based on ratings of SM information. Furthermore, SM provides information that equal employment law prohibits, or severely limits, organizations from using for decision making. For instance, employers may discover, and find difficult to ignore, information concerning applicants' demographic characteristics (e.g., ethnicity, age), physical disabilities, religious beliefs, marital status, or sexual orientation. Moreover, blogs, Twitter posts, and other written material may document mental disabilities, substance abuse, arrests, or other life challenges.

Present Study

Unfortunately, at present, human resources (HR) professionals and managers have limited guidance concerning whether or how SM information should be reviewed during the staffing process. Although numerous popular press articles have addressed this issue, very little systematic research has been conducted to inform decision makers about the implications of using SM in selection.

The present study was conducted to address this important and timely issue. We focused on assessments of information from Facebook.com. Facebook is the most visited website in the world (Alexa.com, 2013) and is the most widely used SM platform, with over 1 billion monthly users (Fowler, 2012). We captured Facebook profiles of graduating college students who were applying for jobs. Recruiters viewed applicants' profiles and made judgments concerning applicants' suitability and their knowledge, skills, abilities, and other characteristics (KSAOs). We then correlated recruiter ratings with applicants' subsequent job performance and turnover.

The results help to address two critical gaps that exist in the knowledge of the SM for staffing phenomenon. First, there is virtually no research concerning whether or how assessments of SM information relate to criteria organizations attempt to affect through their staffing procedures (Brown & Vaughn, 2011; Davison, Maraist, & Bing, 2011). The only study we could find asked a few research assistants to review Facebook profiles of undergraduate students and rate students' personality and hireability (Kluemper, Rosen, & Mossholder, 2012). These ratings were made for a hypothetical job and then were correlated with supervisor ratings of job performance for a subset of the students (n = 56). The results were somewhat mixed in that

Facebook-based ratings of emotional stability and agreeableness correlated significantly with performance, whereas ratings of the remaining three Big Five factors did not.

The present study expands on previous work by examining how actual recruiters evaluate the Facebook profiles of graduating students applying for full-time jobs. In addition, we investigate the criterion-related validity of Facebook ratings with respect to both job performance and turnover. Examination of these issues is important because whether recruiter evaluations of SM information relate to valued criteria could affect the quality and diversity of human capital organizations acquire, as well as the legal defensibility of staffing decisions based on that information. Furthermore, it has been suggested that SM information may provide incremental validity beyond more traditional predictors (e.g., Kluemper & Rosen, 2009). We address this possibility by seeing whether SM assessments increment the validity of established predictors such as cognitive ability, self-efficacy, and personality.

Second, to our knowledge, there are no published data on whether SM assessments are similar or different across subgroups of applicants. This is unfortunate because subgroup differences can lead to adverse impact (e.g., Sackett & Ellingson, 1997), which can negatively affect organizational diversity and jeopardize the legal defensibility of selection decisions (e.g., Ployhart & Holtz, 2008). We address this gap by examining whether Facebook-based ratings differ by gender and ethnicity.

Criterion-Related Validity of Assessments of SM Information

A key aspect in the evaluation of a selection procedure is whether scores on the procedure predict future performance on the job. Withdrawal behaviors, such as absenteeism and turnover, also may represent appropriate criteria to evaluate selection procedures (SIOP, 2003). A careful analysis suggests several factors could inhibit the ability of decision makers to draw valid inferences about applicants' KSAOs and future performance based on SM information. First, as discussed, SM platforms such as Facebook are designed to network with friends and family rather than to measure job-relevant attributes. Indeed, most SM information pertains to applicants' outside-of-work interests and activities, which may have little bearing on work behavior. This factor, in and of itself, may be enough to suggest that criterion-related validity for SM assessments may be low.

Second, the sheer volume of SM information also may inhibit decision makers from drawing valid inferences. For example, Facebook profiles include four different sections of information, and each section can comprise the equivalent of several pages of background information, writing, pictures, videos, and other types of information about an applicant and their social network. This large amount of information may put demands on decision makers' ability to process all the potential cues and to determine what information (if any) is relevant and what is not. This situation may cause decision makers to rely on biases and cognitive heuristics may reduce validity. For example, decision makers may focus on information that is particularly salient (e.g., attractiveness, social activities, political beliefs), but that may not be job-relevant (i.e., the availability heuristic; Tversky & Kahneman, 1974).

Finally, inaccurate information may undermine the criterion-related validity of SM assessments. For example, the desire to be perceived as socially desirable may lead applicants to embellish or fabricate information they post on SM, such as experience, qualifications, and achievements. Furthermore, because other people can post information about

applicants on SM platforms (e.g., Facebook), applicants do not have complete control of their information. As such, applicants may be unduly "penalized" for what others post. In fact, one study found that comments posted by others on one's Facebook profile had a greater effect on observers' impressions than did one's own comments (Walther, Van Der Heide, Kim, Westerman, & Tong, 2008). Last, the lack of concrete identifying information about an applicant, such as background information or recent pictures, may lead to issues of mistaken identity.

In sum, although SM platforms such as Facebook and Twitter may contain some potentially job-relevant information, identifying and accurately rating this information could be quite difficult. Indeed, such platforms are not designed to elicit job-relevant information and, in fact, contain various types of job-irrelevant information that could influence judgments about applicants. Considering all these factors, we predict the following:

Hypothesis 1: SM assessments will not predict on-the-job performance.
Hypothesis 2: SM assessments will not predict turnover intentions.
Hypothesis 3: SM assessments will not predict turnover.

Because these hypotheses predict null effects, we adopt an approach Cortina and Folger (1998) described for dealing with null-effect hypotheses. Using the *triangulation* approach, the researcher (a) includes multiple operationalizations of the focal predictor to show that the null effects are consistent across measures, (b) calculates confidence intervals around the effects for the focal predictor to show the null effects are not likely due to sampling error, and (c) measures other predictors to show that the dependent variable can indeed be predicted. Regarding Step a, we ask recruiters to make both suitability and KSAO ratings based on applicants' SM information. We also calculate and report confidence intervals for relations between the two sets of SM assessments and the criteria, per Step b.

With respect to Step c, we measure several variables previous research has shown to predict job performance and turnover. Cognitive ability is regarded as one of the best predictors of performance, particularly task performance (e.g., Schmidt & Hunter, 1998). General self-efficacy reflects an individual's perception of his or her ability to perform across a variety of different situations (Judge, Erez, & Bono, 1998) and has been shown to relate to job performance (e.g., Judge & Bono, 2001). Core self-evaluation (CSE) is a related construct that refers to appraisals people make concerning their self-worth, competence, and capabilities (Judge, Locke, & Durham, 1997). A meta-analysis (Chang, Ferris, Johnson, Rosen, & Tan, 2012) found that CSE is related to both task performance and organizational citizenship behaviors. We also examine grade point average (GPA), which is thought to reflect a combination of ability and motivation, and predicts job performance (Roth, BeVier, Switzer, & Schippmann, 1996).

Finally, the Big Five personality factors have been the most widely studied set of noncognitive predictors in the selection literature over the past two decades. Although validity evidence for the Big Five is not as strong as the evidence for some of the other constructs, certain traits can predict performance. For example, self-reports of conscientiousness are a consistent, albeit relatively modest, predictor of overall job performance (e.g., Barrick & Mount, 1991), and extraversion and agreeableness have been shown to predict extra-role performance (e.g., Hurtz & Donovan, 2000). Given the track record of these variables for predicting performance, we hypothesize the following:

Hypothesis 4a: Cognitive ability, self-efficacy, core self-evaluation, GPA, and personality (i.e., conscientiousness) will predict in-role performance.

Hypothesis 4b: Core self-evaluation and personality (i.e., agreeableness and extraversion) will predict extra-role performance.

Several of these constructs also have been shown to predict turnover intentions or actual turnover. Specifically, agreeableness, conscientiousness, and emotional stability tend to relate modestly to turnover intentions and actual turnover (Zimmerman, 2008), and CSE has been shown to relate negatively to turnover intentions (Chang et al., 2012). Evidence concerning relations between the other predictors and turnover is more limited or equivocal. For example, research suggests that relations between cognitive ability and turnover tend to be quite small and depend on the cognitive demands of the job (Maltarich, Nyberg, & Reilly, 2010). Moreover, we are not aware of any research linking self-efficacy or GPA to turnover intentions or actual turnover. Thus, we explore possible relations between these variables and turnover intentions and actual turnover.

Hypothesis 5a: Core self-evaluation and personality (i.e., agreeableness, conscientiousness, and emotional stability) will predict turnover intentions.

Hypothesis 5b: Core self-evaluation and personality (i.e., agreeableness, conscientiousness, and emotional stability) will predict actual turnover.

Subgroup Differences in Assessments of SM Information

Another concern is that there may be subgroup differences in assessments of SM information, which could lead to adverse impact. We first note reasons why SM assessments generally may produce differences among different subgroups of applicants. We then focus more specifically on possible subgroup differences with regard to gender and ethnicity.

First, SM platforms such as Facebook provide direct information about applicants' gender, ethnicity, age, and other personal characteristics (e.g., political orientation, religious affiliation), which equal employment law typically prohibits organizations from using for decision making. Thus, a variety of demographic information is available to decision makers who choose to review applicants' SM information. This is a concern because it may be difficult to refute the argument that once demographic information was seen, it was not considered in the decision making (Diamond, 2012; Karl et al., 2010).

Second, searching the Internet for information about applicants represents a fairly weak situation (Mischel, 1979). For one, SM platforms such as Facebook may not yield job-related information because they are designed for social interaction rather than for personnel selection. In addition, the process of searching for and evaluating applicants' SM information is likely to be rather unstructured. For example, searching applicants' SM information can be accomplished at one's desk and outside a formal assessment process in which an applicant is present. Also, unlike assessments such as structured interviews or assessment centers, recruiters may be likely to have a list of specific KSAOs they are trying to assess when reviewing SM information. In a weak situation such as this, job-irrelevant factors (e.g., demographic characteristics) may be more likely to influence recruiters' judgments than in more structured types of assessments (Brown & Vaughn, 2011).

Gender. In addition to the factors noted above, there may be other reasons to expect gender subgroup differences in SM assessments in particular. First, researchers have found that females are less likely than males to post problematic content, such as substance abuse and sexual exploits, on SM platforms (Karl et al., 2010; Peluchette & Karl, 2010). Second, there are gender differences on some of the constructs that could be assessed based on SM information. For example, females tend to have higher verbal ability and writing ability than males (Hough, Oswald, & Ployhart, 2001; Roth, Buster, & Barnes-Farrell, 2010). Such differences could be evident in audio, video, or written SM material.

In addition, some researchers have suggested that organizations could use SM information to assess personality-related constructs (e.g., Kluemper et al., 2012). If such constructs can be assessed from SM information, there may be subgroup differences in the ratings. As an example, Strano (2008) found that compared to males, female Facebook profile pictures were more likely to show females with friends and/or smiling and appearing happy. This is relevant because females tend to possess somewhat higher levels of agreeableness than males (e.g., Hough et al., 2001). Similarly, research has found that females score higher than males on situational judgment tests associated with agreeableness and conscientiousness (Whetzel, McDaniel, & Nguyen, 2008). These factors lead us to predict the following:

Hypothesis 6: SM assessments will be higher for female participants than for male participants.

Ethnicity. We also suspect that there may be ethnic subgroup differences in assessments of SM information. For instance, researchers have found that Black and Hispanic individuals are more likely to post quotes related to their ethnic heritage (Grasmuck, Martin, & Zhao, 2009). Compared to Whites, racioethnic minorities also are more likely to participate in social and political causes via SM (Lieu, 2012). The similarity-attraction paradigm (Byrne, 1971) would predict that the salience of information about minority applicants' ethnicity and political beliefs could generate perceptions of dissimilarity among recruiters from majority groups (i.e., Whites, who constituted the majority of recruiters in the present sample). In turn, recruiters may tend to give higher ratings to applicants they perceive to be more similar to them and lower ratings to participants who may appear less similar.

There also may be existing ethnic subgroup differences in KSAOs recruiters may try to assess based on applicants' SM information. In particular, White individuals tend to be associated with higher levels of mental ability than Black and Hispanic individuals (e.g., Sackett, Schmitt, Ellingson, & Kabin, 2001). Such differences could be manifest in SM information, such as the level of verbal ability evidenced in applicants' Facebook and Twitter posts. Thus, to the extent that recruiters make inferences about intelligence on the basis of SM information, White applicants could receive higher ratings. These factors led us to predict the following:

Hypothesis 7: SM assessments will be higher for White participants than for Black and Hispanic participants.

Method

Participants and Procedures

Three sets of participants provided data for the study. First, we contacted undergraduate and graduate students at a large university in the southeastern United States. These students were near graduation and were looking for jobs, or were preparing to look for jobs. As an incentive, participants were entered into a drawing for various gift certificates. A total of 416 students agreed to participate. Participants were 63.2% female and an average age of 23.8 years (SD = 4.80). In terms of ethnicity, the sample was 78.1% White, 10.8% Hispanic, and 7.0% African American. The few remaining applicants were Asian, Native American, or multiracial.

Participants began by completing an initial online questionnaire that included items about their background, such as their year in school, major, GPA, and demographic characteristics. They also completed measures of the other predictors we examine (e.g., CSE, personality). Last, participants provided access to their Facebook profiles by "friending" a Facebook page set up for this study. Although organizations might not be able to access the Facebook profile of every applicant like we were able to do, many Facebook users do not restrict access to their profiles (Palis, 2012). Even when users do restrict access, viewers often can still see some basic information, including photos. Furthermore, many applicants "friend" organizations (or vice versa) at some point during the staffing process, and some organizations even require applicants to provide access to their SM information (Valdes & McFarland, 2012). Thus, our approach allowed us to obtain a standard set of information for each participant (as we describe below) and was realistic in terms of the Facebook information some organizations may obtain from applicants.

Once participants friended us, we saved scrollable screenshots of their Facebook profile. We saved the four main pages common to all profiles at the time of the study, including Wall, Info, Photos, and Boxes. The *Wall* page contains the user's name, profile picture, demographic information, status message (e.g., what the user currently is doing), and list of friends (i.e., other Facebook users who are connected to the user). This page also includes messages posted by the user and/or the user's friends, as well as information the user wants to share with others, including pictures, videos, and links to others sites. The *Info* page contains additional information about the user, such as his or her contact information, current employer, educational background, and interests groups to which the user belongs. The *Photos* page contains photographs the user or the user's friends have posted. Finally, the *Boxes* page provides the user a place to paste small pictures (e.g., of celebrities they like), cartoons, and other images that reflect the person's interests. The order and appearance of each page was the same as what was on the participant's Facebook profile at the time the information was captured.

Approximately one year later, we sent participants a follow-up questionnaire. A total of 292 of the original 416 participants (70.2%) responded. The questionnaire included various items about their job search activities, such as how many jobs they applied for, how many interviews they had, and whether they had obtained a job. For participants who were or had been employed, we asked them several questions about their job, including their intentions to remain in that job. For those who had left their first job, we asked them why they did so (almost all turnover was reported as voluntary). In addition, we asked participants to provide the name and contact information of their immediate supervisor.

Following the initial collection of participant Facebook information, we contacted recruiters from organizations who were scheduled to attend one of the university's career fairs and/or to conduct interviews on campus. A total of 86 recruiters agreed to participate in the study. Participants were HR staffing specialists (n = 40), hiring managers (n = 29), or employees in the job who were assigned to do recruiting (n = 17). These participants possessed an average of 6.40 years of experience recruiting and selecting employees (SD = 5.46). This set of participants was 55.8% female and 82.4% White (other participants were Black, Hispanic, or Asian), and their average age was 33.11 years (SD = 9.12).

We met with recruiters during the career fairs or at the university career center. We first provided recruiters with an overview of the study. Then, we asked them to evaluate potential applicants as if their organization was evaluating a process for assessing applicants' Facebook information. The Facebook profiles of the applicants were loaded on each of several laptop computers. There was a separate file for each applicant that contained all of the content recruiters would see if they accessed the Facebook profiles directly from the Internet. After recruiters reviewed each applicant's profile, we gave them a rating form and explained the ratings. Recruiters could refer back to the applicant's profile while making their ratings. We did not provide recruiters other information about participants (e.g., resumes), so that their evaluations would be based only on the Facebook profiles and not influenced by other information. Each recruiter evaluated an average of five applicants.

The last set of participants consisted of immediate supervisors of participants whose Facebook profiles we captured. We contacted supervisors by phone and asked them to evaluate participants' job performance. We were able to obtain performance information on 142 of the 292 participants who completed the follow-up questionnaire. Supervisors had worked with participants for an average of 13.7 months.

Measures

Unless noted, items within the measures described below were rated on a 5-point Likert-type scale with anchors that ranged from *strongly disagree* (1) to *strongly agree* (5).

Predictors: Facebook ratings. Relatively little is known about how organizations review and evaluate SM information during the staffing process. However, there is some indication that organizations review applicants' SM information early in the process and make fairly global assessments as to applicants' general suitability (Preston, 2011; SHRM, 2008). Therefore, we asked recruiters to rate the suitability of each participant using five items adapted from previous research (e.g., Adkins, Russell, & Werbel, 1994; Cable & Judge, 1997; Stevens & Kristof, 1995). Example items include "I can see how this person would be an attractive applicant to an organization," "I would further consider this person for employment if they had the skills to fill an open position," and "I would be hesitant to pursue this person as an applicant after viewing their Facebook profile" (reverse scored). We averaged the five ratings to form an overall composite of participant suitability (α = .93).

In addition, we identified 10 KSAOs that are relevant to performance in many types of jobs and that have been mentioned as attributes organizations do or could evaluate based on SM information (e.g., Brown & Vaughn, 2011; Davison, Maraist, Hamilton, & Bing, 2012; Grasz, 2012). These included adaptability, creativity, dependability, integrity, intelligence,

interpersonal skills, leadership, maturity and professionalism, work ethic, and written communication skills. Recruiters rated the level of each attribute on a scale with anchors that ranged from *very low* (1) to *very high* (5).

Based on previous analyses of ratings data (e.g., Pulakos & Schmitt, 1995; Viswesvaran, Schmidt, & Ones, 2005), we expected that recruiter ratings of the 10 KSAOs might reflect a more general impression of each participant. However, a principal axis factor analysis (with oblique rotation of factors) suggested that their ratings were somewhat differentiated and loaded on three main factors. The first factor comprised four items that appeared to reflect conscientiousness-related attributes (work ethic, dependability, integrity, and maturity). The second factor comprised two cognitive related items, namely intelligence and written communication skills. The third factor included three items that appeared more relevant to contextual aspects of work (interpersonal skills, adaptability, and creativity). Leadership ratings did not load on any of the factors. Thus, in addition to creating a variable that reflected the mean ratings across all 10 KSAOs, we created separate variables for each of the three factors that emerged. The internal consistency reliability (alpha) for the overall variable was .90, and the reliabilities of the conscientiousness, cognitive, and contextual variables were .86, .79, and .78, respectively.

Other predictors. We measured cognitive ability based on participants' self-reported ACT scores. Previous research has found that self-reported test scores correlate highly with actual scores. For example, a meta-analysis by Kuncel, Crede, and Thomas (2005) found a mean observed correlation of .82 ($k = 6$) between self-reported and actual SAT total scores. Reliability estimates for the ACT tend to be in the .90 range (ACT, 2007). CSE was measured using the scale developed by Judge, Erez, Bono, and Thoresen (2003). This scale includes 10 items, such as "When I try, I succeed," "I am capable of coping with most of my problems," and "Sometimes, I do not feel in control of my work" (reverse scored). The coefficient alpha for a composite of the 10 items was .86.

We used the International Personality Item Pool (IPIP; Goldberg et al., 2006) to measure the five-factor model dimensions. Each dimension was measured using 10 items. Coefficients alpha ranged from .69 for openness to .87 for emotional stability (see Table 1). We also measured general self-efficacy using an 8-item scale from the IPIP. This scale includes items such as "I can handle complex problems," "I am quick to understand things," and "I let myself be directed by others" (reverse scored). The coefficient alpha for a composite of the 8 items was .73. Finally, participants reported their overall GPA. Kuncel et al. (2005) found a mean observed correlation of .90 ($k = 12$) between reported GPA and actual GPA, but also large mean differences between reported and actual GPA (mean $d = 1.38$). These findings suggest a consistent tendency for research participants to inflate their grades.

Criteria

Job performance. Supervisors rated participants' performance using seven items adapted from Williams and Anderson (1991). Three items measured in-role performance (e.g., "The employee performs tasks they are asked to complete"), three items measured extra-role performance (e.g., "The employee goes out of their way to help other employees"), and one item measured overall performance (i.e., "Overall, I am happy with this employee's performance"). A confirmatory factor analysis revealed a good fit for a two-factor model that

Table 1
Descriptive Statistics, Correlations, and Reliability Estimates for the Facebook Ratings, Other Predictors, and Criteria

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
Facebook ratings																					
1. Suitability	3.56	0.93	(.93)														−.04	−.10	−.08	.00	−.01
2. KSAOs	3.14	0.66	.55**	(.90)													−.09	−.15	−.14	−.05	.01
3. Conscientiousness	3.12	0.79	.59**	.92**	(.86)												−.15	−.22	−.21	−.03	.00
4. Cognitive	3.27	0.75	.51**	.83**	.74**	(.79)											.03	.04	−.00	−.10	−.18
5. Contextual	3.13	0.76	.26**	.80**	.57**	.51**	(.78)										−.03	−.08	−.07	−.03	.12
Other predictors																					
6. Cognitive ability	26.07	3.37	.02	.10	.05	.23**	.05	—									.30	.09	.23	−.08	−.16
7. Agreeableness	4.06	0.51	.10*	.05	.05	.00	.04	.03	(.85)								.17	.02	.12	.03	−.19
8. Conscientiousness	3.90	0.58	.14**	.06	.08	.00	.05	−.22**	.27**	(.83)							.03	−.07	−.02	−.10	−.22
9. Emotional stability	3.39	0.69	−.06	−.14**	−.09*	−.12**	−.15**	−.09	.05	.19**	(.87)						.02	.06	.05	−.12	−.15
10. Extraversion	3.48	0.67	−.04	−.02	−.05	−.06	.07	−.07	.36**	.19**	.14**	(.86)					.16	.28	.25	−.06	.06
11. Openness	3.91	0.53	.04	−.01	−.02	.00	.00	.18**	.20**	.03	.21**	.22**	(.69)				.02	−.02	.00	.04	−.10
12. CSE	3.85	0.53	.01	−.04	−.04	−.03	−.05	−.05	.16**	.38**	.61**	.33**	.21**	(.86)			.15	.21	.21	−.16	−.09
13. Self-efficacy	4.03	0.44	.03	.04	.02	.05	.05	.08	.20**	.30**	.32**	.30**	.51**	.50**	(.73)		.34	.18	.31	−.03	−.07
14. GPA	3.39	0.43	.25**	.21**	.24**	.26**	.05	.26**	.08	.10**	−.10*	−.12**	.01	.04	.11**	—	.08	.00	.05	−.13	−.20
Criteria																					
15. In-role performance	6.65	0.48	−.04	−.08	−.13	.03	−.03	.26**	.15*	.02	.02	.14	.02	.13	.30**	.07	(.74)				
16. Extra-role performance	6.26	0.88	−.08	−.13	−.18**	−.03	−.07	.07	.02	−.06	.05	.24**	−.01	.18**	.15*	.00	.45**	(.70)			
17. Overall performance	6.52	0.52	−.07	−.12	−.18**	−.00	−.06	.20*	.10	−.02	.04	.22*	.00	.18**	.27**	.04	.86**	.84**	(.75)		
18. Turnover intentions	2.68	1.18	.00	−.05	−.03	−.09	−.03	−.08	.03	−.10	−.12*	−.06	.04	−.15**	−.02	−.13**	.07	−.04	.02	(.89)	
19. Turnover	0.04	0.20	−.01	.01	.00	−.11*	.07	−.10	−.12*	−.14**	−.09	.04	−.06	−.06	−.04	−.13**	−.10	−.05	−.09	−.03	—

Note: $N = 416$ for the Facebook ratings and the other predictors, 142 for job performance ratings, 292 for turnover intentions, and 291 for turnover. Reliability estimates appear in parentheses on the diagonal. Observed correlations appear below the diagonal, and predictor–criterion correlations corrected for unreliability in the criterion (for job performance and turnover intentions) or to a .50 base rate (for turnover) appear above the diagonal. CSE = core self-evaluation; GPA = grade point average; KSAOs = knowledge, skills, abilities, and other characteristics.

*$p < .10$, two-tailed.

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

included an in-role factor and an extra-role factor (e.g., comparative fit index [CFI] = .99, root mean square error of approximation [RMSEA] = .04). This model also fit the data much better than a single-factor model (CFI = .91, RMSEA = .12). Thus, we created separate composites for in-role and extra-role performance. The factor loading for one of the extra-role items (i.e., "This employee makes suggestions to improve the organization") was weak, and including this item detracted from internal consistency reliability. Therefore, we excluded this item from the final composite. We also created an overall job performance composite based on the mean of the remaining six items. The coefficients alpha for the three composites were .74, .70, and .75, respectively. Because most participants had only one supervisor, we were unable to estimate the interrater reliability of the performance ratings.

Turnover intentions. Participants indicated their intentions to leave their present job using a five-item scale from Bozeman and Perrewé (2001). Example items included "I probably will look for a new job in the near future" and "I do not intend to quit my job" (reverse scored). The coefficient alpha for a composite of the five items was .89.

Turnover. We also asked participants whether they were still in, or had left, their first job after graduation. We focused on voluntary turnover, as this type of turnover tends to be more under employees' control (e.g., we excluded a few instances of involuntary turnover for reasons such as downsizing). However, the turnover base rate in the sample was low (4.1%), which likely reflects the modest time lag and the economic conditions at the time of the data collection. Because low base rate criteria can be difficult to predict, we corrected all the correlations involving turnover to reflect a 50-50 split between leavers and stayers (Kemery, Dunlap, & Griffeth, 1988; Zimmerman, 2008).

Results

Criterion-Related Validity of Facebook Ratings and Other Predictors

Table 1 presents descriptive statistics, reliability estimates, and intercorrelations for recruiter Facebook ratings, the other predictors, and the criteria. We also present correlations corrected for measurement error in the criteria (but not in the predictors) to estimate operational validity. However, to be consistent with prior research (e.g., Griffeth, Hom, & Gaertner, 2000; Harrison, Newman, & Roth, 2006; Zimmerman, 2008), we did not correct for measurement error in turnover, which is assumed to be measured with relatively little error. For all the results presented below, we note the statistical significance of the observed correlations and regression coefficients. However, in discussing the results, we focus on the corrected values to draw inferences regarding operational validity.

Bivariate validity estimates for Facebook ratings. Hypothesis 1 predicted that assessments of SM information would not be related to job performance. As Table 1 shows, neither suitability ratings nor overall KSAO ratings related significantly to any of the performance criteria. Operational validity estimates and 95% confidence intervals (CIs) for the suitability ratings were –.04 for in-role performance (CI = −.21 to .13), –.10 for extra-role performance (CI = −.27 to .07), and –.08 for overall performance (CI = −.25 to .09). The corresponding

validity estimates for overall KSAO ratings were –.09 (CI = −.26 to .08), –.15 (CI = −.32 to .01), and –.14 (CI = −.31 to .03), respectively. Thus, all of the CIs for the corrected correlations include zero. Similar results were found for the more specific KSAO factors. The one exception was that conscientiousness scores were significant and negatively related to extra-role performance ($r_{corrected}$ = −.22, CI = −.38 to –.06) and to overall performance ($r_{corrected}$ = −.21, CI = −.37 to –.05). Overall, these results provide support for Hypothesis 1.

Hypotheses 2 and 3 predicted that SM assessments would not predict turnover intentions and turnover, respectively. As with job performance, both sets of Facebook ratings were not significant predictors of turnover intentions. Operational validity estimates were .00 for suitability ratings (CI = −.17 to .17) and –.05 for KSAO ratings (CI = −.22 to .12). Furthermore, none of the KSAO factors were related to turnover intentions. Similarly, neither set of Facebook ratings related significantly to turnover. Operational validity estimates were –.01 for suitability ratings (CI = −.18 to .16) and .01 for KSAO ratings (CI = −.16 to .18). The one exception was that higher scores on the cognitive factor were associated with lower turnover ($r_{corrected}$ = −.18, CI = −.34 to −.02). The overall results provide support for Hypotheses 2 and 3.

Bivariate validity estimates for the other predictors. Recall that we included measures of several other predictors to show that the dependent variables could indeed be predicted (per Cortina & Folger, 1998). In partial support of Hypothesis 4a, cognitive ability and self-efficacy were statistically significant predictors of in-role performance ($r_{corrected}$ = .30 and $r_{corrected}$ = .34, respectively; see Table 1). Agreeableness also predicted in-role performance ($r_{corrected}$ = .17). In contrast, CSE, GPA, and conscientiousness were not significant predictors of in-role performance. In partial support of Hypothesis 4b, CSE ($r_{corrected}$ = .21) and extraversion ($r_{corrected}$ = .28) predicted extra-role performance, as did self-efficacy ($r_{corrected}$ = .18). However, agreeableness did not predict extra-role performance. Finally, cognitive ability ($r_{corrected}$ = .23), extraversion ($r_{corrected}$ = .25), CSE ($r_{corrected}$ = .21), and self-efficacy ($r_{corrected}$ = .31) were significant predictors of overall performance.

Hypothesis 5a predicted that personality (i.e., agreeableness, conscientiousness, and emotional stability) and CSE would predict turnover intentions. In partial support of this hypothesis, emotional stability ($r_{corrected}$ = −.12) and CSE ($r_{corrected}$ = −.16) were significant predictors of turnover intentions, as was GPA ($r_{corrected}$ = −.13). However, agreeableness and conscientiousness were not significant predictors. Hypothesis 5b predicted that these same variables would predict turnover. Agreeableness and conscientiousness were significant predictors of turnover, as was GPA. The base-rate-adjusted validity estimates for these variables were –.19, –.22, and –.20, respectively. In contrast, emotional stability and CSE were not significantly related to turnover. These results provide partial support for the hypothesis.

In sum, several of the other variables predicted job performance, turnover intentions, and/or actual turnover. This suggests that the small (and largely nonsignificant) relations between Facebook ratings and performance and turnover are unlikely due to problems with the criterion measures.

Incremental validity of Facebook ratings. We also regressed each performance criterion on the predictors to examine whether the Facebook ratings provided any incremental validity beyond the other predictors. The results of these analyses are shown in Table 2. As with the bivariate relationships, significance tests were conducted on the observed regression coefficients, but we focus on the corrected coefficients given our interest in operational validity.

Table 2
Incremental Validity of Facebook Ratings Over the Other Predictors

Criterion/Model/Predictor	β	R	R^2	ΔR^2
Criterion: In-role performance				
Model 1: Other predictors		.40** (.46)	.16 (.21)	
Model 2: Other predictors + Facebook ratings		.42** (.48)	.18 (.23)	.02 (.02)
Cognitive ability	.24** (.28)			
Agreeableness	.08 (.09)			
Conscientiousness	–.03 (–.03)			
Emotional stability	–.09 (–.11)			
Extraversion	.05 (.06)			
CSE	.04 (.05)			
Self-efficacy	.28** (.31)			
GPA	–.01 (–.01)			
Facebook suitability ratings	.01 (.02)			
Facebook KSAO ratings	–.13 (–.16)			
Criterion: Extra-role performance				
Model 1: Other predictors		.33* (.38)	.11 (.15)	
Model 2: Other predictors + Facebook ratings		.35* (.41)	.12 (.17)	.02 (.02)
Cognitive ability	.06 (.07)			
Agreeableness	–.06 (–.07)			
Conscientiousness	–.15 (–.17)			
Emotional stability	–.10 (–.12)			
Extraversion	.22** (.26)			
CSE	.19 (.22)			
Self-efficacy	.07 (.09)			
GPA	.03 (.04)			
Facebook suitability ratings	.02 (.01)			
Facebook KSAO ratings	–.14 (–.16)			
Criterion: Overall performance				
Model 1: Other predictors		.38** (.44)	.15 (.19)	
Model 2: Other predictors + Facebook ratings		.41** (.47)	.17 (.22)	.02 (.03)
Cognitive ability	.18** (.21)			
Agreeableness	.01 (.02)			
Conscientiousness	−.10 (–.11)			
Emotional stability	−.12 (–.13)			
Extraversion	.15 (.17)			
CSE	.13 (.15)			
Self-efficacy	.21** (.24)			
GPA	.01 (.02)			
Facebook suitability ratings	.02 (.02)			
Facebook KSAO ratings	−.16 (–.19)			
Criterion: Turnover intentions				
Model 1: Other predictors		.26** (.26)	.07 (.07)	
Model 2: Other predictors + Facebook ratings		.27** (.27)	.07 (.07)	.00 (.00)
Cognitive ability	−.09 (–.09)			
Agreeableness	.10 (.10)			
Conscientiousness	−.10 (–.10)			

(continued)

Table 2 (continued)

Criterion/Model/Predictor	β	R	R^2	ΔR^2
Emotional stability	−.10 (–.09)			
Extraversion	−.09 (–.08)			
CSE	−.10 (–.12)			
Self-efficacy	.12* (.11)			
GPA	−.13** (–.13)			
Facebook suitability ratings	.07 (.07)			
Facebook KSAO ratings	−.07 (–.07)			
Criterion: Turnover				
Model 1: Other predictors		.26** (.41)	.07 (.17)	
Model 2: Other predictors + Facebook ratings		.26** (.41)	.07 (.17)	.00 (.00)
Cognitive ability	−.11* (–.18)			
Agreeableness	−.11* (–.17)			
Conscientiousness	−.15** (–.24)			
Emotional stability	−.12 (–.21)			
Extraversion	.08 (.13)			
CSE	.04 (.09)			
Self-efficacy	.04 (.05)			
GPA	−.10 (–.15)			
Facebook suitability ratings	.03 (.06)			
Facebook KSAO ratings	.03 (.03)			

Note: $N = 416$ for the Facebook ratings and the other predictors, 142 for job performance ratings, 292 for turnover intentions, and 291 for turnover. Values in parentheses reflect regression results based on predictor-criterion correlations corrected for unreliability in the criterion (for job performance and turnover intentions) or to a .50 base rate (for turnover). β = standardized regression coefficients from the other predictors + Facebook ratings model; CSE = core self-evaluation; GPA = grade point average; KSAOs = knowledge, skills, abilities, and other characteristics.
*$p < .10$.
**$p < .05$.

The other predictors, as a set, were significantly related to in-role performance ($R_{corrected} = .46$, $R^2_{corrected} = .21$).[1] Including Facebook ratings did not increment the validity of the other predictors ($\Delta R^2_{corrected} = .02$, $p = .34$).[2] The other predictors also related significantly to extra-role performance ($R_{corrected} = .38$, $R^2_{corrected} = .15$), and adding Facebook ratings did not significantly increase prediction ($\Delta R^2_{corrected} = .02$, $p = .31$). Finally, the other predictors related significantly to overall performance ($R_{corrected} = .44$, $R^2_{corrected} = .19$). Once again, including Facebook ratings did not demonstrate significant increment validity ($\Delta R^2_{corrected} = .03$, $p = .21$).

The other predictors also related significantly to turnover intentions ($R_{corrected} = .26$, $R^2_{corrected} = .07$). As with job performance, adding the Facebook ratings did not increment the validity of the other predictors ($\Delta R^2_{corrected} = .00$, $p = .53$). Finally, the other predictors related significantly to turnover ($R = .41$, $R^2 = .17$), and adding the Facebook ratings did not provide any incremental prediction ($\Delta R^2 = .01$, $p = .74$). Thus, in no instance did the Facebook ratings provide significant incremental validity beyond the other predictors.

Subgroup Differences in Facebook Ratings

Table 3 displays descriptive statistics and standardized mean differences (*d*) for Facebook ratings by gender and ethnicity. Hypothesis 6 predicted that SM assessments would be higher for females than males. In support of this hypothesis, scores for female participants were significantly higher than male participants for both suitability ($d = -0.35$, $p < .01$) and overall KSAO ratings ($d = -0.20$, $p = .03$). Differences on the KSAO ratings appeared to be due primarily to the fact that females received higher ratings on the conscientiousness factor ($d = -0.25$, $p < .01$), particularly on items related to maturity and integrity.

Hypothesis 7 predicted that SM assessments would be higher for White participants than for Black and Hispanic participants. In partial support of this hypothesis, Whites received significantly higher suitability ratings than both Blacks ($d = 0.35$, $p = .03$) and Hispanics ($d = 0.34$, $p = .01$), but not on overall KSAO ratings. In addition, White participants tended to receive higher scores than Black participants on the cognitive factor ($d = 0.32$, $p = .06$) and lower scores than Hispanic participants on the contextual factor ($d = -0.34$, $p = .03$).

One consideration in evaluating subgroup differences in SM assessments is how they correspond to more traditional measures of similar constructs.[3] To explore this issue, we calculated *d* statistics for ACT scores and IPIP conscientiousness scores. We then compared these statistics to the *d*s for recruiters' ratings of cognitive- and conscientiousness-related KSAOs. Male–female *d*s for recruiter-rated conscientiousness and IPIP conscientiousness scores were −0.25 and −0.28, respectively. White–Black *d*s for recruiter-rated cognitive ability and ACT scores were 0.32 and 0.40 (respectively), and the *d*s for recruiter-rated and IPIP conscientiousness scores were 0.09 and 0.04. In comparing White and Hispanic participants, *d*s for recruiter-rated cognitive ability and ACT scores were −0.10 and −0.05 (respectively), and the *d*s for recruiter-rated and IPIP conscientiousness scores were 0.08 and 0.01. Thus, the direction and magnitude of subgroup differences generally were consistent between recruiter ratings and self-report scores. The one exception was the male–female *d*s for cognitive ability, which were −0.17 based on recruiter ratings and 0.08 based on ACT scores.

Discussion

The Internet has had a profound effect on the way organizations recruit and select employees. One way organizations are using the Internet is to search for SM information about job applicants (Grasz, 2012; Levinson, 2010; Preston, 2011). Unfortunately, these organizations appear to be assessing SM information in the absence of data concerning the validity of inferences made on the basis of such information. As such, organizations that assess SM information probably do not know whether use of such assessments leads to better decisions in terms of selecting employees who demonstrate effective job performance or longer tenure. In addition, we are not aware of any previous published data on whether SM assessments are similar for applicants of different subgroups. This could leave organizations in a precarious situation if use of such information yields adverse impact against protected groups of applicants (e.g., racial minorities, older applicants) and there is no validity evidence to defend the use of the procedure.

Table 3

Descriptive Statistics and Standardized Mean Differences for Facebook Ratings by Gender and Ethnicity

					KSAO Ratings										
		Suitability Ratings			Overall			Conscientiousness			Cognitive			Contextual		
Subgroup	*N*	*M*	*SD*	*d*	*M*	*SD*	*d*	*M*	*SD*	*d*	*M*	*SD*	*d*	*M*	*SD*	*d*
Gender																
Males	153	3.37	0.98	−0.35**	3.05	0.67	−0.20**	2.99	0.82	−0.25**	3.18	0.77	−0.17**	3.07	0.77	−0.07
Females	263	3.70	0.90		3.18	0.65		3.19	0.77		3.32	0.74		3.13	0.80	
Ethnicity																
Whites	324	3.64	0.92		3.13	0.67		3.13	0.79		3.28	0.75		3.07	0.81	
Blacks	29	3.30	1.00	0.35**	3.09	0.62	0.06	3.07	0.62	0.09	3.05	0.69	0.32*	3.13	0.76	−0.07
Hispanics	45	3.31	1.00	0.34**	3.19	0.61	−0.10	3.07	0.76	0.08	3.36	0.79	−0.10	3.31	0.68	−0.32**

Note: d = standardized mean difference, which was calculated as $M_{\text{majority group}} - M_{\text{minority group}}/SD_{\text{pooled}}$. KSAOs = knowledge, skills, abilities, and other characteristics.
*$p < .10$, one-tailed.
**$p < .05$, one-tailed.

Main Findings and Implications

The present study was conducted to further examine the potential implications of using SM information for selection. We asked recruiters to review and evaluate Facebook profiles of graduating college students who were looking for jobs. Results suggest that recruiter ratings generally are unrelated to graduates' subsequent job performance, turnover intentions, and turnover. This trend holds across overall suitability ratings, aggregate KSAO ratings, and the three KSAO factors that emerged (i.e., conscientiousness, cognitive, and contextual). In addition, Facebook ratings do not provide incremental prediction of these criteria beyond more established predictors, including cognitive ability, personality, self-efficacy, CSE, and GPA.

We also discovered that Facebook ratings tend to be higher for females than for males, and, in several instances, higher for White individuals than for Black and Hispanic individuals. Gender and ethnic differences do vary somewhat by the KSAOs recruiters rated. For example, females tend to score higher on conscientiousness-related KSAOs, whereas Whites tend to score higher on cognitively related KSAOs (e.g., writing ability and intelligence). The overall pattern of subgroup differences for the Facebook ratings generally is consistent with the direction and magnitude of subgroup differences typically found with other measures of constructs related to those the recruiters evaluated. Lower SM assessments for racioethnic minorities could lead to a serious problem because once an adverse impact is established, organizations often must provide evidence of criterion-related validity or business necessity (Arvey & Faley, 1988). If there is no evidence of validity (as was the case in the current study), organizations may be less able to defend use of the procedure against claims of adverse impact.

Recommendations

Overall, the present findings cast serious doubts concerning the appropriateness of considering applicants' SM information (Facebook ratings in the current study) during the selection process. To date, there is little evidence to support the criterion-related validity of inferences based on SM assessments. Our results also suggest there may be subgroup differences in such assessments that could lead to ethnicity-based adverse impact.

Furthermore, when it is not feasible to establish the criterion-related validity of a selection procedure (e.g., for small-*N* jobs or when appropriate criterion measures are not available; Sackett & Arvey, 1993), organizations often must rely on evidence of content-related validity. Inferences concerning content validity are enhanced when the process of developing a selection procedure is based on a systematic analysis of the tasks and KSAOs a particular job requires (Arvey & Faley, 1988; Goldstein, Zedeck, & Schneider, 1993). Content validity also is enhanced to the extent the selection procedure requires applicants to respond to questions or perform tasks that are similar to what employees actually do on the job (EEOC, 1978). Given the social purpose of SM platforms such as Facebook, and the fact that applicants do not use such platforms to apply for a particular job, it might be quite difficult to demonstrate the validity of content-related inferences based on Facebook and other forms of SM.

Finally, there are potential issues related to the availability of job-irrelevant information on SM and the lack of consistent information across applicants. Indeed, although most of

the recruiters in our sample indicated that they use SM information, they do not appear to seek such information in every situation or for every applicant. Such issues could make it challenging to use SM and other Internet information in a standardized way. Furthermore, because SM platforms such as Facebook expose decision makers to factors that equal employment law restricts, the burden of proof may be on the organization to demonstrate that group membership did not factor into the hiring decision (and this may be difficult to accomplish).

On the basis of these factors, we strongly encourage organizations to refrain from using SM (e.g., Facebook) and other Internet information (e.g., Google searches) until methods for collecting and evaluating such information are shown to be reliable and valid. We also recommend that organizations develop clear policies concerning the use of SM and other Internet information. Indeed, the results of an SHRM (2011) survey suggested that most organizations do not have a policy regarding the use of Internet information. Some of the background data we collected for the present study suggest that recruiting managers and employees are more likely to use SM information during the staffing process than are HR specialists. Therefore, it may be important for HR staff to educate managers and employees about the potential problems with using this information.

Strengths and Limitations

The present study possessed several strengths that enabled us to examine the potential use of SM assessments for staffing. For example, the results are based on evaluations of (a) actual Facebook profiles of active applicants made by (b) practicing recruiters from a range of organizations. In addition, we captured participants' SM information when they were applying for jobs and then followed up with them several months later to measure their on-the-job performance and turnover. This predictive design may parallel when and how organizations would assess SM information in an operational selection context. This design also allowed us to sample a large number of applicants early in the job search process before other predictors (e.g., interviews) winnowed the field down to a more restricted group. Finally, we collected data on a diverse set of measures from three different sources (i.e., applicants, recruiters, and supervisors), which should further minimize the potential influence of method variance.

Despite these strengths, several factors may limit the inferences and generalizations that can be drawn from some of the results. First, although graduating college students would seem like an appropriate population from which to sample, some of the results may not apply to evaluation of more experienced applicants. For instance, more experienced applicants may tend to exhibit greater maturity and more vigilance regarding what they post online. If so, there could be less variance in SM assessments of such applicants, which actually could further attenuate the criterion-related validity of such assessments.

Second, participants represented a range of jobs and organizations. Although this design may enhance generalizability, it also fails to control factors that could influence some of the variables and relationships, such as predictors that vary in relevance across jobs (e.g., some of the five-factor model dimensions). Related to this, we asked recruiters from one set of organizations to evaluate participants' overall suitability as a job applicant, as well as their standing on KSAOs relevant to performance in many types of jobs. We then asked supervisors from a different set of organizations to evaluate participants' performance in a particular

job. Thus, it is possible that relations between the SM assessments and the criteria could have been somewhat larger had we been able to collect ratings from recruiters and supervisors from the same organizations.

Third, given the dearth of research on the use of SM information in staffing, we wanted to focus on the relationship between recruiter evaluations of Facebook and the outcomes. For example, we chose not to provide recruiters additional information (e.g., resumes, selection test scores) that could have made it difficult to isolate the effects of the Facebook information. As such, our results probably are most relevant to organizations that review SM information early in the selection process, before other predictor information is reviewed. Our findings may be less relevant to organizations that consider such information later in the selection process in conjunction with other predictors or after applicants have been screened in on other predictors.

Finally, the turnover rate in the present sample was low. Although we corrected correlations involving turnover to reflect a 50-50 split between leavers and stayers, the validity results for turnover may be somewhat unstable due to the small number of participants who turned over during the study period.

Future Research

Clearly, there is a tremendous need and opportunity for additional research in this area. For one, future studies could replicate and extend our work on criterion-related validity. For example, we examined recruiter ratings of Facebook, which is the most popular SM platform in the world and one that many recruiters appear to consult. Future studies could explore other Internet resources, such as LinkedIn and Twitter.

Future studies also might explore different ways to assess SM information. We asked recruiters to evaluate participants on overall suitability as well as on several KSAOs. This approach may be more structured than the ways organizations may evaluate applicants' Internet information, such as using a single overall rating of hireability, or perhaps no quantitative ratings at all. Future studies might explore other ways to structure SM assessments to see if validity can be improved. For example, it might be possible to adapt strategies researchers have identified for structuring selection interviews (e.g., Campion, Palmer, & Campion, 1997, 1998). Perhaps SM information could be reviewed to identify behavioral examples that reflect job-related KSAOs. These examples then could be used to create rating scales decision makers could use to evaluate applicants' SM information (Brown & Vaughn, 2011; Slovensky & Ross, 2012). However, even if researchers could identify a small set of well-defined KSAOs to measure, and develop a structured way to measure the KSAOs, there still would be concerns about the availability of job-irrelevant information, lack of consistent information across applicants, and so forth.

Furthermore, future research could investigate other ways organizations might use SM and other Internet information. For instance, some organizations might use SM assessments to screen out applicants whose websites contain problematic content, such as substance abuse or illegal acts. As an example, we reviewed the Facebook profiles of 36 participants whose mean suitability rating was 2.00 or lower to see if we could identify why recruiters may have provided low ratings for these individuals.[4] Although we did not observe any one theme relevant to most or all of these participants' profiles, we did notice three factors that

were relevant to some of the profiles. Specifically, 19 profiles included profane language, 17 profiles included photos that showed the participant at parties or drinking, and 13 profiles were of participants who had traditionally non-White names and/or who were clearly non-White. Other less prevalent themes included strange profile pictures, religious quotes, and sexual references. Future studies could explore in more detail the types of SM information that leads recruiters to screen out certain applicants and whether applicants whose websites contain such information, in fact, have lower performance, higher turnover, and so forth.

If future research discovers that SM assessments can predict criteria such as job performance, a next step would be to examine how such assessments might influence, or work in conjunction with, other predictors. For example, does access to SM information influence the types of questions interviewers ask? How do decision makers weigh SM information relative to other predictor information?

Additional construct validity research within this domain also may be useful. Several studies have examined the convergence between self-ratings of personality and other ratings of personality based on SM information. For example, Vazire and Gosling (2004) reported a mean self-observer correlation of .31 across the Big Five factors based on the average ratings of 11 undergraduate raters. Similarly, Marcus, Machilek, and Schütz (2006) reported a mean self-observer correlation of .20 across the Big Five based on the average ratings from five undergraduate raters (also see Back et al., 2010).

These findings suggest that correlations between self-ratings of personality and observer ratings of personality (based on SM information) are small to moderate. However, these values are based on ratings from 5 to 10 observers, which likely is more than the number of recruiters or managers who may evaluate the SM information of any given applicant. Furthermore, the previous studies collected self and observer ratings from the same measure (i.e., self and observer forms of a Big Five inventory). In the present study, we found lower levels of convergence (e.g., $r = .08$ for conscientiousness; see Table 1) when observer ratings were based on a single recruiter and different measurement methods (although the correlation between recruiter ratings of cognitive-related KSAOs and self-reported ACT scores [.23] was somewhat larger). Overall, we encourage more research that examines the construct-related validity of recruiter or manager ratings of job-relevant constructs based on SM information.

The reliability of SM assessments also deserves attention. For example, we are not aware of any studies that have estimated the interrater reliability of recruiter or manager ratings of SM information. In an effort to provide some initial data, we collected SM ratings from a second recruiter for 90 of the participants in our study.[5] We then calculated intraclass correlation coefficients (ICC,2) between the mean ratings of the two recruiters. The interrater reliability estimate for the suitability ratings was .23. However, the interrater correlation for one of the five items from the suitability scale ($r = -.02$) was notably lower than the interrater correlations for the other four items. Excluding this item from the scale resulted in an interrater estimate of .31. The interrater reliability of the mean of the KSAO ratings was .14. Here again, the interrater correlation for one of the 10 KSAOs ($r = -.20$, leadership) was notably lower than the mean of the other correlations. Excluding this item resulted in an interrater estimate of .29. These results suggest that different recruiters' ratings of the same applicants may be only modestly consistent.

The present study appears to be one of the first to examine subgroup differences in SM assessments. Future studies might examine different subgroups (e.g., Asians, younger versus older applicants), as well as continue to build a base of data for other groups such as Blacks and Hispanics. In addition, SM platforms such as Facebook contain a range of other personal information that could affect recruiter evaluations. Future studies could code SM for information on physical attractiveness, sexual orientation, or disability status to see if such factors influence recruiter evaluations.

Future research also should investigate how applicants react to organizations that review their Internet information. Applicant reactions in this domain are somewhat different in that applicants do not take an assessment per se. Indeed, reviews of SM and other Internet information likely are conducted without applicants' knowledge. What is germane is how applicants react to the fact that organizations are reviewing SM information. We suspect that use of SM assessments could engender negative reactions from some applicants. For example, such assessments might score low on perceptions of job relatedness, opportunity to perform, two-way communication, and transparency (Bauer et al., 2001; Gilliland, 1993; Hausknecht, Day, & Thomas, 2004). Such concerns could be particularly prevalent when SM assessments include requests for passwords or to "friend" a representative of the organization.

Finally, there may be opportunities for qualitative research in this area. For example, to better understand how decision makers use Internet information, researchers could ask hiring officials to "talk out" what they are looking for in applicants' SM information and how they use that information to make decisions. Also, what inferences do recruiters make when applicants do not participate in SM platforms such as Facebook, or when they do, but privacy settings prevent recruiters from reviewing some or all of the information?

Concluding Thoughts

Although many organizations are using SM and other Internet information to help make decisions about job applicants, there is extremely limited empirical evidence to support this practice. The results of the present study revealed that recruiter ratings of Facebook profiles correlate essentially zero with job performance, turnover intentions, and turnover. There also was evidence of subgroup differences in recruiters' ratings that tended to favor female and White applicants. Overall, we are concerned that the ways organizations are using Facebook and related SM platforms to facilitate staffing decisions may not yield valid inferences about future performance and may yield subgroup differences, and that recruiters may not use SM assessments in standardized ways. We look forward to future research that attempts to shed light on the many questions and concerns about this phenomenon.

Notes

1. Openness to experience—which correlated moderately with some of the other predictors (e.g., core self-evaluation), but was unrelated to all the criteria—appeared to function as a suppressor variable in several of the regression models. Thus, we excluded openness from these analyses.

2. These and subsequent incremental validity results were highly similar regardless of whether we included overall knowledge, skills, abilities, and other characteristics (KSAO) scores or scores for each of the three KSAO factors. For ease of interpretation, we report results using the overall scores only.

3. We thank an anonymous reviewer for raising this issue.
4. We thank Action Editor Fred Oswald for suggesting this analysis.
5. We thank an anonymous reviewer for suggesting this.

References

ACT. 2007. *The ACT technical manual*. Iowa City, IA: Author.

Adkins, C. L., Russell, C. J., & Werbel, J. D. 1994. Judgments of fit in the selection process: The role of work value congruence. *Personnel Psychology*, 47: 605-623.

Alexa.com. 2013. *How popular is Facebook.com?* Retrieved from http://www.alexa.com

Arvey, R. D., & Faley, R. H. 1988. *Fairness in selecting employees* (2nd ed.). Reading, MA: Addison-Wesley.

Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. 2010. Facebook profiles reflect actual personality, not self-idealization. *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.

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. 2001. Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, 54: 388-420.

Bozeman, D. P., & Perrewé, P. L. 2001. The effect of item content overlap on Organizational Commitment Questionnaire-turnover cognitions relationships. *Journal of Applied Psychology*, 86: 161-173.

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: 219-255.

Byrne, D. E. 1971. *The attraction paradigm*. New York: Academic Press.

Cable, D. M., & Judge, T. A. 1997. Interviewer perceptions of person–organization fit and organizational selection decisions. *Journal of Applied Psychology*, 82: 546-561.

Campion, M. A., Palmer, D. K., & Campion, J. E. 1997. A review of structure in the selection interview. *Personnel Psychology*, 50: 655-702.

Campion, M. A., Palmer, D. K., & Campion, J. E. 1998. Structuring employment interviews to improve reliability, validity, and users' reactions. *Current Directions in Psychological Science*, 7: 77-82.

Chang, C.-H., Ferris, D. L., Johnson, R. E., Rosen, C. C., & Tan, J. A. 2012. Core self-evaluations: A review and evaluation of the literature. *Journal of Management*, 38: 81-128.

Cortina, J. M., & Folger, R. G. 1998. When is it acceptable to accept a null hypothesis: No way, Jose? *Organizational Research Methods*, 1: 334-350.

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

Davison, H. K., Maraist, C., Hamilton, R. H., & Bing, M. 2012. To screen or not to screen? Using the Internet for selection decisions. *Employee Responsibilities and Rights Journal*, 24: 1-21.

Deschenaux, J. 2010. Beware using social networking sites for hiring. *Staffing Management*, 6: 40.

Diamond, T. E. 2012. Social media in the workplace: A blessing and curse. *EEO Insight*, 4: 25-30.

Elzweig, B., & Peeples, D. 2009. Using social networking web sites in hiring and retention decisions. *S.A.M. Advanced Management Journal*, 74: 27-35.

Equal Employment Opportunity Commission, C. S. C., Department of Labor, and Department of Justice. 1978. Uniform guidelines on employee selection procedures. *Federal Register*, 43(166): 38290-38315.

Fowler, G. A. 2012. Facebook: One billion and counting. *Wall Street Journal*, October 4. Retrieved from http://search.proquest.com

Gilliland, S. W. 1993. The perceived fairness of selection systems: An organizational justice perspective. *Academy of Management Review*, 18: 694-734.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. 2006. The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality*, 40: 84-96.

Goldstein, I. L., Zedeck, S., & Schneider, B. 1993. An exploration of the job analysis-content validity process. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*: 3-34. San Francisco: Jossey-Bass.

Grasmuck, S., Martin, J., & Zhao, S. 2009. Ethno-racial identity displays on Facebook. *Journal of Computer-Mediated Communication*, 15: 158-188.

Grasz, J. 2012. Thirty-seven percent of companies use social networks to research potential job candidates, according to new CareerBuilder survey. *CareerBuilder*, April 18. Retrieved from http://www.careerbuilder.com

Griffeth, R. W., Hom, P. W., & Gaertner, S. 2000. A meta-analysis of antecedents and correlates of employee turnover: Update, moderator tests, and research implications for the next millennium. *Journal of Management*, 26: 463-488.

Harrison, D. A., Newman, D. A., & Roth, P. L. 2006. How important are job attitudes? Meta-analytic comparisons of integrative behavioral outcomes and time sequences. *Academy of Management Journal*, 49: 305-325.

Hausknecht, J. P., Day, D. V., & Thomas, S. C. 2004. Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57: 639-683.

Hough, L. M., Oswald, F. L., & Ployhart, R. E. 2001. Determinants, detection, and amelioration of adverse impact in personnel selection procedures: Issues, evidence, and lessons learned. *International Journal of Selection and Assessment*, 9: 152-194.

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

Judge, T. A., & Bono, J. E. 2001. Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86: 80-92.

Judge, T. A., Erez, A., & Bono, J. E. 1998. The power of being positive: The relation between positive self-concept and job performance. *Human Performance*, 11: 167-187.

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. 2003. The Core Self-Evaluations Scale: Development of a measure. *Personnel Psychology*, 56: 303-331.

Judge, T. A., Locke, E. A., & Durham, C. C. 1997. The dispositional causes of job satisfaction: A core evaluations approach. *Research in Organizational Behavior*, 19: 151-188.

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: 174-186.

Kemery, E., Dunlap, W., & Griffeth, R. W. 1988. Correction for range restriction in point-biserial correlations. *Journal of Applied Psychology*, 73: 688-691.

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

Kluemper, D. H., Rosen, P. A., & Mossholder, K. 2012. Social networking websites, personality ratings, and the organizational context: More than meets the eye? *Journal of Applied Social Psychology*, 42: 1143-1172.

Kuncel, N. R., Crede, M., & Thomas, L. L. 2005. The validity of self-reported grade point averages, class ranks, and test scores: A meta-analysis and review of the literature. *Review of Education Research*, 75: 63-82.

Levinson, M. 2010. Social networking ever more critical to job search success. *CIO.com*, June. Retrieved from http://www.cio.com

Lieu, D. 2012. Hispanics and Blacks more likely than Whites to support causes on line. *Chronicle of Philanthropy*, June. Retrieved from http://philanthropy.com

Maltarich, M. A., Nyberg, A. J., & Reilly, G. 2010. A conceptual and empirical analysis of the cognitive ability-voluntary turnover relationship. *Journal of Applied Psychology*, 95: 1058-1070.

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: 1014-1031.

Meinert, D. 2011. Seeing behind the mask. *HR Magazine*, 56: 31-32, 34, 36-37.

Mischel, W. 1979. On the interface of cognition and personality: Beyond the person-situation debate. *American Psychologist*, 34: 740-754.

Palank, J. 2006. Face it: "Book" no secret to employers; social sites used as background check. *Washington Times*, July 17: A1. Retrieved from http://www.lexisnexis.com

Palis, C. 2012. Facebook privacy options ignored by millions of users: Consumer Reports. *Huffington Post*, May 3. Retrieved from http://www.huffingtonpost.com

Peluchette, J., & Karl, K. 2010. Examining students' intended image on Facebook: "What were they thinking?!" *Journal of Education for Business*, 85: 30-37.

Ployhart, R. E., & Holtz, B. C. 2008. The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology*, 61: 153-172.

Preston, J. 2011. Social media history becomes a new job hurdle. *StarNews*, July. Retrieved from http://iw.newsbank.com

Pulakos, E. D., & Schmitt, N. 1995. Experience-based and situational interview questions: Studies of validity. *Personnel Psychology*, 48: 289-308.

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

Roth, P. L., Buster, M. A., & Barnes-Farrell, J. 2010. Work sample exams and gender adverse impact potential: The influence of self-concept, social skills, and written skills. *International Journal of Selection and Assessment*, 18: 117-130.

Sackett, P. R., & Arvey, R. D. 1993. Selection in small *N* settings. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations*: 418-447. San Francisco: Jossey-Bass.

Sackett, P. R., & Ellingson, J. E. 1997. The effects of forming multi-predictor composites on group differences and adverse impact. *Personnel Psychology*, 50: 707-721.

Sackett, P. R., Schmitt, N., Ellingson, J. E., & Kabin, M. B. 2001. High-stakes testing in employment, credentialing, and higher education: Prospects in a post-affirmative action world. *American Psychologist*, 56: 302-318.

Schmidt, F. L., & Hunter, J. E. 1998. The validity and utility of selection methods in personnel selection: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, 124: 262-274.

Slovensky, R., & Ross, W. H. 2012. Should human resource managers use social media information to screen job applicants? *Managerial and Legal Issues in the USA*, 14: 55-59.

Society for Human Resources Management. 2008. Online technologies and their impact on recruitment strategies. *SHRM Staffing Research*, July-September.

Society for Human Resources Management. 2011. The use of social networking websites and online search engines in screening job candidates. *SHRM Staffing Research*, August.

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

Stevens, C. K., & Kristof, A. L. 1995. Making the right impression: A field study of applicant impression management during job interviews. *Journal of Applied Psychology*, 80: 587-606.

Strano, M. M. 2008. User descriptions and interpretations of self-presentation through Facebook profile images. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace*, 2: 1-11.

Tversky, A., & Kahneman, D. 1974. Judgment under uncertainty: Heuristics and biases. *Science*, 185: 1124-1131.

Valdes, M., & McFarland, S. 2012. Job seekers getting asked for Facebook passwords. *Associated Press*, March 20. Retrieved from http://finance.yahoo.com

Vazire, S., & Gosling, S. D. 2004. e-Perceptions: Personality impressions based on personal websites. *Journal of Personality and Social Psychology*, 87: 123-132.

Viswesvaran, C., Schmidt, F. L., & Ones, D. S. 2005. Is there a general factor in ratings of job performance? A meta-analysis framework for disentangling substantive and error influences. *Journal of Applied Psychology*, 90: 108-131.

Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. 2008. The role of friends' appearance and behavior on evaluations of individuals on Facebook: Are we known by the company we keep? *Human Communication Research*, 34: 28-49.

Whetzel, D. L., McDaniel, M. A., & Nguyen, N. 2008. Subgroup differences in situational judgment test performance: A meta-analysis. *Human Performance*, 21: 291-308.

Williams, L. J., & Anderson, S. E. 1991. Job satisfaction and organizational commitment as predictors of organizational citizenship and in-role behaviors. *Journal of Management*, 17: 601-617.

Willis, D. P. 2006. On the lookout. *Asbury Park Press*, July 31. Retrieved from http://proquest.umi.com

Wilson, R. E., Gosling, S. D., & Graham, L. T. 2012. A review of Facebook research in the social sciences. *Perspectives on Psychological Science*, 7: 203-220.

Zimmerman, R. D. 2008. Understanding the impact of personality traits on individuals' turnover decisions: A meta-analytic path model. *Personnel Psychology*, 61: 309-348.