Structuring social media assessments in employee selection

Christopher J. Hartwell¹ | Jake T. Harrison¹ | Rahul S. Chauhan² | Julia Levashina³ | Michael A. Campion⁴

¹Management Department, Utah State University, Logan, Utah, USA
²Department of Management and Information Systems, Kent State University, Kent, Ohio, USA
³Department of Marketing, Texas A&M University, College Station, Texas, USA
⁴Management Department, Purdue University, West Lafayette, Indiana, USA

Abstract
Social media assessments (SMAs) are a common, relatively new, practice in employee selection. However, SMAs are generally conducted in an informal way, leaving organizations with a practice low in reliability and validity, as well as opening up potential legal and ethical issues. We propose a framework of nine SMA structural components based on previously developed interview structural components. Effects of each component on reliability, validity, user reactions, and legality are discussed. We conduct two studies that measure the utilization of each structural component in research literature and in practice. Study 1 is a content analysis of prior SMA academic research. Results indicate the highest use for the procedural consistency and rating scales used and the lowest utilization of documentation. Study 2 surveyed hiring professionals to examine the structure level typically utilized for SMAs in practice. Results indicate low levels of structure for most components, which demonstrates a need for structural improvement when SMAs are used in hiring. We discuss the implications of our research and provide recommendations for structuring SMAs in both research and practice.

Keywords
employee selection, hiring, social media assessment, structure

Practitioner points
What is currently known about the topic:
- SMAs are a common practice during employee selection processes.
- SMAs are typically done covertly and inconsistently, not formally part of the process.
- Adding structure increase reliability and validity of selection processes (e.g., interviews).

What our paper adds to previous knowledge:
- Many structural components developed in an interviewing context can be applied to SMAs.
- SMA structural components are identified, with definitions of low, medium, and high levels.

Portions of this paper, including preliminary results of Study 1, were previously published in an undergraduate-focused research journal (Harrison & Hartwell, 2020).
• Usage of SMA structure in the research literature is examined using content analysis.
• Usage of SMA structure in practice is studied using a survey of hiring professionals.

Implications of our findings for practitioners:
• Implementing SMA structure could increase validity, reliability, user reactions, and legality.
• Low usage rates of SMA structure among practitioners demonstrate potential for improvement.
• The SMA structural framework provides a blueprint for how to add structure to SMAs.

1 | INTRODUCTION

Social media (SM) has been extensively used by organizations for employment decisions. Between 40% and 80% of staffing professionals use SM in hiring (e.g., CareerBuilder, 2018; Hartwell & Campion, 2020; SHRM, 2013), and screening candidates using SM has increased significantly in recent years (CareerBuilder, 2017). Despite this increased attention to SM in practice, there are many legal, ethical, and validity concerns about SM assessments (SMAs; Slovensky & Ross, 2012). One concern is that SMAs are typically informal and inconsistent, likely reducing their validity (Van Iddekinge et al., 2016). Many authors suggest that formally structuring SMAs is one way to maximize reliability and validity, while minimizing legal concerns (e.g., McFarland & Ployhart, 2015; Roth et al., 2016). Research is needed on how SMAs can be improved through structure and the extent such practices are currently used in academic literature and practice.

To answer this study call, this study focuses on SMA structure and provides three major contributions. First, based on the structured interviewing literature, we develop a framework for SMAs in employee selection that will be useful for designing SMAs in academic research and hiring practice. The framework includes empirically based structural components that are operationalized using high, medium, and low levels. The components’ potential impacts on reliability, validity, user reactions, and legality are also discussed. Second, we conduct a content analysis of prior SMA research in Study 1 to examine how these structural components have been utilized, providing insight into how SMAs are currently operationalized in academic research. Third, we survey hiring professionals in Study 2 to examine how the structural components are utilized in SMA practice, providing practical insight into how structured SMAs are currently used in employee selection.

1.1 | Development of the structural framework

Empirical evidence for the benefits of structured over unstructured interviews abounds, including improving validity, reliability, and legal acceptability (e.g., Conway et al., 1995; Cronshaw & Wiesner, 1989; Huffcutt & Arthur, 1994; McDaniel et al., 1994). Campion et al. (1997) proposed 15 structural components for a selection interview. Seven of these were theorized to influence the content of the interviews, or the nature of the information elicited: (1) conducting a job analysis, (2) asking the same questions, (3) limiting prompting, (4) asking valid question types, (5) conducting longer interviews, (6) controlling ancillary information, and (7) not allowing questions from applicants. The other eight components influence the evaluation of the interview or the processing and judgment of the information elicited: (8) rating answers, (9) using anchored rating scales, (10) taking detailed notes, (11) using multiple interviewers per applicant, (12) using the same interviewer(s) across candidates, (13) not comparing applicants between interviews, (14) providing interviewer training, and (15) using statistical prediction.

Although much research covers interview structure, there is a need for research on SMA structure. We developed our SMA structure framework by comparing interview and SMA contexts, and determining which structured interview components were relevant to SMAs. The goal of both methods is to acquire candidate information to predict performance, and validity, reliability, and legality are relevant in both contexts. One major difference is interviewers interact with applicants to elicit specific information (an active assessment), while such interaction is not present in SMAs. Rather, SMA raters make inferences from already available information (a passive assessment; Hartwell & Campion, 2020). Thus, some structural components that affect the interview content are not applicable in SMAs, including limiting prompting, conducting longer interviews, not allowing questions from applicants, and not comparing applicants between interviews. However, by altering other components to the SMA context (e.g., replacing interview questions with topics to evaluate) many interview components are applicable in SMAs, despite the lack of personal interaction with applicants.

In fact, nonempirical SMA articles have suggested similar structural components to those in interviewing, such as conducting a job analysis or identifying job-relevant criteria (Davison et al., 2012; Ployhart, 2012; Slovensky & Ross, 2012), establishing policies for
consistent procedures across candidates (Clark & Roberts, 2010; Davison et al., 2012; Elzweig & Peeples, 2009), developing standardized rating forms (Davison et al., 2012; Kluemper, 2013; Ployhart, 2012; Van Iddekinge et al., 2016), taking detailed notes (Byrnside, 2008; Ployhart, 2012; Slovensky & Ross, 2012), using multiple raters (Brown & Vaughn, 2011; Davison et al., 2012; Kluemper, 2013), training raters (Elzweig & Peeples, 2009; Kluemper, 2013; Ployhart, 2012), and combining ratings (Roth et al., 2016). They also recommend structural elements that affect the acceptability of SMAs (in a legal sense and by the applicant): informed consent and notification of results (Smith & Kidder, 2010). These components are specific to passive selection processes that do not require direct interaction with the applicant (e.g., SMAs, background screens, credit checks, reference checks).

As SMAs have grown in popularity, concerns regarding the reliability and validity of social media data have developed (Hartwell & Eggli, 2020). Evidence for SMA validity is mixed. SMAs using Facebook have shown some validity when measuring personality and predicting self-rated personality, supervisor-rater performance, and college GPA (Kluemper et al., 2012), and LinkedIn-focused SMAs show promise in validly measuring some general competencies (e.g., leadership, communication) and in predicting some career-related outcomes (promotions, degree-related jobs) (Roulin & Levashina, 2019). But other studies suggest SMAs may not have much criterion-related validity (Van Iddekinge et al., 2016; Zhang et al., 2020). Given this lack of proven reliability and criterion-related validity, SMA utility may be limited. We suggest that validity and reliability issues could be resolved, at least in part, by structuring the process.

Drawing from Campion et al.’s (1997) interview structure components and recommendations of SMA literature, we developed a framework of six structural components that directly relate to one or more structured interview components, and two additional components specific to the passive SMA procedure. Table 1 defines each component, highlights the Campion et al. (1997) correlating component (where applicable), and defines levels of structure (low/medium/high).

Campion et al.’s (1997) structured interview framework examined each structural component’s impact on reliability, validity, and user reactions. Legality is another measure of a selection method’s value (Roulin & Levashina, 2019). The passive nature of the SMAs, requiring no direct input from the applicant, may create legal and ethical implications. SMAs may constitute an invasion of privacy (Brandenburg, 2008; Clark & Roberts, 2010), defamation (Byrside, 2008; Davison et al., 2012), or a violation of the Fair Credit Reporting Act (FCRA; Byrside, 2008; Davison et al., 2012; Smith & Kidder, 2010). They also have legal exposure to other selection procedures, such as adverse impact (Byrside, 2008; Roth et al., 2016). Therefore, we also examine each SMA structural component’s potential influence on validity, reliability, user reactions, and legality.

It is worth noting differences in SM platforms. Some focus on text, while others focus on pictures or video. Some have unique features, such as Twitter’s character limit or Snapchat’s short availability of videos. One distinction relevant to SMAs is between personal and professional SM platforms. For personal SM platforms, user content is largely unconstrained, while professional SM platforms focus on career and job-related content (Hartwell & Campion, 2020). Because personal SM content is generally less structured than professional SM content, structuring the SMA will likely have a greater impact for personal SM compared to professional SM—though benefits are likely to occur for both. Where applicable in the components outlined below, we note likely differences between SMAs using personal versus professional SM platforms.

### 1.1.1 Job-relatedness

The first component of SMA structure focuses on the job-relatedness of the constructs being measured. In the lowest level of structure, the SMA focuses on overall applicant impressions, such as overall qualifications (Bohnert & Ross, 2010), hireability (Kluemper et al., 2012), or suitability/fit (Van Iddekinge et al., 2016). Medium structure uses general (non-job-specific) knowledge, skills, abilities, and other attributes (KSAOs), including cognitive ability (Kluemper & Rosen, 2009; Van Iddekinge et al., 2016) and personality attributes (Bohnert & Ross, 2010; Kluemper & Ross, 2009; Kluemper et al., 2012). High structure for this component includes measuring specific job-related KSAOs.

**Effects on validity, reliability, user reactions, and legality**

*Effects on validity, reliability, user reactions, and legality.* The higher the level of job-relatedness, the closer the measurements will relate to the specific job. Some research indicates that increased job-relatedness in SMA ratings creates higher reliability and lower adverse impact than overall, general LinkedIn ratings (Roulin & Levashina, 2019). However, Hartwell (2015) proposed that when there is inconsistency of available information across applicants (as is the case specifically for personal SM platforms), then broader constructs (e.g., person-job fit, hireability) may be a more valid and reliable approach because it can accommodate a wider range of information. Thus, professional SM may lend itself to the highest level of structure, while medium structure may better match personal SM.

User reactions (both applicant and SMA rater) are likely to positively relate due to the increased face validity of the process as job-relatedness of constructs increases. Finally, using higher levels more closely aligns the ratings to the specific job. Since demonstrating job-relatedness is a part of the legality of selection procedures (Uniform Guidelines on Employee Selection Procedures, 1978), higher levels of the job-relatedness component should increase legality.

### 1.1.2 Procedural consistency

This component entails the uniformity of SMA procedures across applicants. The higher the structure is obtained the more applicants are treated equally. Low structure is signified by process
<table>
<thead>
<tr>
<th>Structural component</th>
<th>Definition</th>
<th>Corresponding interview components</th>
<th>Levels of structure</th>
</tr>
</thead>
</table>
| Job-relatedness      | Level of constructs used | - Job analysis  
- Valid question types | High: SMA measures job-specific KSAOs.  
Med: SMA measures general KSAOs.  
Low: SMA measures overall impressions. |
| Procedural consistency | Uniformity of procedures across applicants | - Same questions  
- Same interviewers | High: Exact same set of procedures are followed for all applicants.  
Med: General process is roughly similar for all applicants.  
Low: Same set of procedures is not followed for all applicants. |
| Rating scales used    | Level of measurement detail | - Rating each question  
- Anchored rating scales  
- Statistical prediction | High: Each trait is measured using multiple rating scale items.  
Med: Each trait is measured with a single rating scale item.  
Low: No rating scales are used. |
| Documentation         | Notes taken and records kept | - Taking detailed notes | High: Detailed records kept regarding information found during SMA.  
Med: General notes are made during the SMA.  
Low: No records kept regarding the SMA. |
| Assessor training     | Level of training provided to assessors | - Interviewer training | High: Those conducting SMAs are provided with comprehensive training on how to do so effectively.  
Med: Those conducting SMAs are provided with basic instructions.  
Low: Those conducting SMAs are not trained. |
| Use multiple raters   | Number of raters for each profile | - Multiple interviewers | High: Having three or more SMA raters.  
Med: Having two SMA raters.  
Low: Having only one SMA rater. |
| Separate rater(s) from decision maker(s) | Having rater(s) other than the decision maker(s) conduct the SMA | - Ancillary information | High: SMA rater(s) and decision maker(s) are completely separate.  
Med: SMA rater(s) are part of the decision-making group.  
Low: SMA rater and decision-maker are the same individual(s). |
| Informed consent      | Level of applicant consent gathered in relation to the SMA | N/A | High: Applicants give informed consent specifically for the SMA.  
Med: Applicants consent to general background screen, part of which is the SMA.  
Low: No informed consent is given by the applicants. |
| Notification of results | Information given to applicant after the SMA | N/A | High: Applicant notified when SMA influences hiring decision, with chance to appeal.  
Med: Applicant notified when SMA influences hiring decision, without chance to appeal.  
Low: Applicants are not notified regarding SMA results. |

Abbreviations: KSAO, knowledge, skills, attributes, and other individual characteristics; N/A, not applicable; SMA, social media assessment.

*From Campion et al. (1997).*
inconsistency, such as only some applicants being subjected to the SMA or checking different SM platforms for different candidates. Medium structure includes all applicants going through a similar process with some inconsistencies. For example, assessing Facebook (FB) profiles of all candidates, but searching further into the past with some applicants than others or spending the same amount of time reviewing each applicant's SM profiles, but not searching the same platform(s) for each applicant. High structure consists of using the exact same procedures, such as reviewing the past 12 months of FB information for all applicants.

**Effects on validity, reliability, user reactions, and legality**

Procedural consistency is likely to have a direct positive impact on reliability (Campion et al., 1997). This is particularly true for personal SM, where a viewer can review multiple years (or even decades) of past content. Higher reliability increases the upper bounds of validity (Ghiselli et al., 1981). Consistency of administration is also likely to increase positive user reaction, including increased fairness perceptions (Smither et al., 1993) and legal defensibility (increased standardization) (Latham & Finnegan, 1993).

1.1.3 | Rating scales used

This component examines the level of measurement detail present, with structure increasing as detail level increases. Low structure relies on overall impressions (e.g., “This person seems like they would be a good leader”) rather than rating scales. Medium structure uses a single rating scale for each construct. High structure includes using multiple items for each rating scale. In essence, this structural component focuses on the specific measurement(s) utilized to assess the constructs of interest in the job-relatedness component.

**Effects on validity, reliability, user reactions, and legality**

Utilizing multiple ratings is likely to have a positive impact on all four outcomes. Multiple ratings will increase reliability over no-ratings (“gut” impressions) or single ratings, and reliability is a necessary condition for validity (Campion et al., 1997; Conway et al., 1995). Validity may also increase because multiple ratings allow more job-related content to be evaluated. Thus, increasing validity leads to higher legal defensibility for the SMA. Finally, applicants may appreciate that multiple rating scales provide more opportunities to perform in any given area than scales with single or no ratings. But SMA raters may find the process more tedious with multiple multiple scales.

1.1.4 | Documentation

Documentation refers to the notes taken during the SMA; the more specific the documentation, the higher the structure. Low structure includes no SMA records. Medium structure includes general SMA notes (e.g., “removed from consideration based on lack of relevant job experience”). High structure includes detailed notes regarding information found during the SMA that influenced specific ratings and/or decision-making.

**Effects on validity, reliability, user reactions, and legality**

As with interviews, taking notes during the SMA should increase reliability and validity by acting as justification for the ratings and decisions (Middendorf & Macan, 2002). Notes assist raters in making more reliable and accurate ratings. Applicant reactions are unlikely to be affected by note-taking, as they would not be present during the process or privy to those notes (unless part of the notification of results component discussed later). SMA rater may react positively, being able to note the process behind the ratings, or negatively, being forced to keep notes. If note taking is objective and job-related, it will have a positive effect on SMA legality, as the notes provide justification for the ratings. However, if notes include non-job-related or legally protected information (e.g., race, sexual orientation, religion, etc.) or subjective evaluations (e.g., “immature” or “rock star!”) instead of direct behaviors/content (e.g., “made negative comment about current work,” or “previous experience matches open job”), there could be a negative effect on legality and validity (Burnett et al., 1998; Middendorf & Macan, 2002). This is a particular concern for personal SM platforms, where non-job-related information is more abundant.

1.1.5 | Assessor training

This refers to how well those conducting SMAs are trained on doing so in a reliable and valid manner. Structure increases as the breadth and depth of SMA training increase. The low structure includes no training for SMA assessors. With medium structure, the assessor is given basic instructions (e.g., “look for red flags” or “see if the applicant would be a good fit with our company”). High structure includes comprehensive training on how to effectively conduct SMAs. This could include frame-of-reference training, how to avoid common rating errors (e.g., leniency, contrast, halo), legal considerations (e.g., avoiding discrimination), how to interpret common SM information, or conducting practice ratings.

**Effects on validity, reliability, user reactions, and legality**

The realistic accuracy model (RAM; Funder, 1995, 2001) proposes that accurate assessment of individuals depends on (1) available information that is (2) relevant and that the information is properly (3) detected and (4) interpreted by the assessor. One major issue with SMAs is understanding what available information is relevant. Providing assessor training can help assessors recognize relevant SM content and what conclusions should be drawn (Elzweig & Peeples, 2009). Assessor training is designed to minimize judgment errors and increase rating accuracy—having a positive effect on reliability and validity (Pulakos & Schmitt, 1995). More in-depth training would likely be useful for assessing personal SM platforms compared to professional SM platforms. SMA assessor reactions are
likely to improve, as the SMA process and focus are more explicit (Campion et al., 1997). Applicant reactions are also likely to increase when they understand that assessors have been trained effectively (Latham & Finnegan, 1993). Finally, the legality of the process will increase, as assessor training can be used to demonstrate a standardized process focused on obtaining job-related ratings based on relevant information (Williamson et al., 1997).

### 1.1.6 | Use multiple raters

As has been demonstrated with interview and performance ratings, having multiple raters for SMAs can reduce idiosyncratic biases and tendencies, ensuring more reliable and valid ratings (Campion et al., 1997; Viswesvaran et al., 1996). Low structure is only having one rater for the SMA, medium structure is having two raters, and high structure is having three or more raters. Having three raters is typical of panel interviews (Campion et al., 1997), and having three unstructured serial interviews has similar validity to one structured interview (Schmidt & Zimmerman, 2004). Thus, we propose that having three SMA raters is sufficient to be considered high structure.

**Effects on validity, reliability, user reactions, and legality**

**Effects on validity, reliability, user reactions, and legality.** Having multiple raters will likely lead to higher reliability and validity of SMA ratings. The effect of user reactions is uncertain, particularly when using a personal SM. If applicants know multiple people are rating their SM, they could have more negative reactions than if it is a single rater. But they could also view the process more favorably if they understand that multiple raters will reduce bias. Negative applicant reactions are less likely (and positive reactions more likely) when using multiple raters for professional SM, as applicants tend to encourage organizations to view this content. SMA raters could react positively or negatively to having multiple raters. Legality is increased, as the effects of idiosyncratic biases and tendencies are reduced, and multiple people provide checks on one another’s biases (Williamson et al., 1997).

### 1.1.7 | Separate rater(s) from decision maker(s)

Those rating SM profiles are likely to encounter non-job-related and/or protected class information (e.g., age, race, religion, sexual identity, and political affiliation). Having a separate person or people making the hiring decision means that the decision maker(s) will not be exposed to non-job-related content that could consciously or unconsciously bias their decision (Fisher, 2011; Sprague, 2007). This relates to Campion et al.’s (1997) structured interview component that encourages interviewers not to have access to ancillary information (e.g., resumes) because such information is likely to bias interview ratings. It is also akin to maintaining applicant demographic information (e.g., gender, race, and military status) separately during the selection process for reporting purposes to not bias decisions.

Some employers might hire third parties to conduct SMAs to provide a barrier between the employer and potential protected class information available on SM, guarding against lawsuits (e.g., Levashina et al., 2017). The lowest level of structure is that the SMA rater and the decision maker are the same person (one person is involved). For medium structure, SMA raters and decision makers are the same person, but multiple people are involved. The highest level of structure includes separating the SMA rater(s) and the decision maker(s).

**Effects on validity, reliability, user reactions, and legality**

Separating rater(s) from decision maker(s) is likely to positively impact the reliability and validity of the overall hiring decisions by reducing contamination and bias. SM (especially personal SM) often includes salient non-job-related content, including information legally protected in many countries (e.g., race, religion, and sexual orientation). Separating SMA rater(s) and decision maker(s) means the hiring decision is less likely to be biased because the decision maker(s) was/were never exposed to SM content. Informing the applicant that SM raters are separate from decision makers could increase the perceived face validity of the process, making reactions more positive (Holden & Jackson, 1979). Finally, this structural element will increase the legality of the hiring process, as it reduces the chance for discriminatory or biased decisions, as described previously.

### 1.1.8 | Informed consent

This refers to the applicant being notified of and agreeing to the SMA. Low structure entails no informed consent. Medium structure includes the applicant’s consent to conduct a general background screen, which includes the SMA. High structure requires applicants’ consent to specifically conduct the SMA. Because professional SM profiles are typically created to promote one’s career, this component may not be as relevant. Some applicants may even list their professional SM site on their application. But informed consent is particularly an issue for personal SM platforms, where the organization is likely not the intended audience.

**Effects on validity, reliability, user reactions, and legality**

Getting the applicants’ informed consent before conducting SMAs may actually reduce validity. RAM (Funder, 1995, 2001) explains that accurate judgments about individuals depend on having information available. When consenting to the SMA, the applicant is notified that their SM will be examined. They may change privacy settings to restrict access or delete problematic content. Thus, judgment is potentially less accurate because less relevant information is available, reducing validity. However, the limited available information may increase reliability, as raters will be more likely to utilize the same information to base judgments. But these effects are speculative, and the true impact on validity and reliability is uncertain.

While initial indications suggest that applicants may react negatively to SMAs (Drouin et al., 2015; Jeske & Shultz, 2019; Roth...
et al., 2016), reactions are likely to depend on a variety of things. First, previous research has demonstrated that applicants are likely to have negative reactions when potential employers review personal SM (e.g., Facebook, Twitter, and Instagram), but positive reactions when professional SM (e.g., LinkedIn) is reviewed (Nikolaou, 2014). Second, information provided to applicants is likely to influence their reactions. Prior literature has demonstrated that applicants have more positive reactions when they understand the job-relatedness of the selection procedure (Ryan & Ployhart, 2000). Providing the applicant with information regarding the purpose of the SMA could therefore reduce negative or increase positive applicant reactions. Effects on SMA rater reactions are unknown. They may appreciate the defensibility that informed consent provides or be frustrated by having to get informed content before conducting SMAs.

Finally, informed consent will increase the legality of the SMA. U.S. laws like the Fair Credit Reporting Act (FCRA) and Stored Communications Act (SCA), and other state-specific and international laws, could be applicable to SMAs (Byrnside, 2008; Kennedy & Macko, 2007; Smith & Kidder, 2010). FCRA requires that the applicant is informed of and consent to the information that will be screened. Similarly, the General Data Protection Regulation (GDPR) used in the European Union and United Kingdom (and a template for similar laws worldwide) requires consent for gathering personal data (Prictor et al., 2019). Thus, this structural component directly increases SMA legality in many countries.

### 1.1.9 Notification of results

While informed consent notifies the candidate before the SMA, notification of results is concerned with information provided to the applicant after the SMA, particularly when SMA results influence the hiring decision. Low structure entails not notifying applicants regarding SMA results. Medium structure includes notifying applicants when the SMA influences the hiring decision, but not allowing the applicant to appeal. High structure requires notifying the applicant and allowing for an appeal.

*Effects on validity, reliability, user reactions, and legality*

There is unlikely to be a direct result on validity or reliability from notification of results, other than potential correction of erroneous findings. For example, situations of mistaken identity are still possible. By notifying the applicant of the information found, a hiring manager may find that the SM profile was a different person with the same name.

Similar to the informed consent component, the effects on user reactions are unknown. Applicants who were unaware of the SMA may feel a violation of privacy if told the results (Barry & Fulmer, 2004; Potosky, 2008). However, negative feedback that can help applicants clean up their SM presence—or positive feedback regarding their SM—could be viewed favorably by applicants. SMA rater reactions to this component could similarly be positive or negative. This component will have a positive effect on legality by complying with legal standards in the FCRA, SCA, GDPR, and other laws (Byrnside, 2008; Smith & Kidder, 2010).

Utilizing these structural components, our two studies evaluate current SMAs practices in two settings. Study 1 analyzes prior academic literature on SMAs, coding the level of each structural component in each study. Study 2 uses a survey to understand current practice in how hiring managers utilize these structural components in SMAs. These two studies examine the frequency of use of our developed structural components in both research (Study 1) and practice (Study 2).

### 2 STUDY 1

Previous research has not explicitly addressed how to best structure SMAs, though some structural components are found in existing studies. We perform a content analysis that measures the extent to which SMA structural components are used and the level of structure in published academic studies. This provides useful information for researchers and a baseline for comparing the structural levels used by practitioners in Study 2.

#### Research Question 1: How structured are SMAs as operationalized in academic research?

### 2.1 Research method

We utilized PRISMA guidelines (Page et al., 2021) to conduct a literature review and report the results of a content analysis of SMA structure used in these studies. Studies were considered relevant if participants reviewed one or more SM profiles in an employment selection context. Our focus was on measuring the use of structure in SMA studies, using our developed framework.

We completed a database keyword search (Harari et al., 2020) of Business Source Premier and PsycINFO using the Boolean phrase "(Social Media OR ‘Social Networking Website’ OR ‘LinkedIn’ OR ‘LinkedIn Profile’) AND (Recruitment OR Hiring OR Selection OR Screening OR Cybervetting)" in May 2020 and November 2021. After reviewing abstracts, excluding articles focused on alternate topics such as computer programming or marketing and identifying articles that specifically included an SMA in an employment selection context, 15 articles met our criteria, from which 20 content analyses were conducted (some articles had more than one study). A backward search of references within identified articles did not uncover any additional studies.

Two authors separately rated each study on a scale of High (3), High/Medium (2.5), Medium (2), Medium/Low (1.5), or Low (1) structure for most structural components (see Table 1). Scores that end in ".5" indicate that the level of that component's structure falls between the measures for Low and Medium or Medium and High. For example, Kluemper and Rosen’s (2009) study indicates that some measures used single-item rating scales while others used multiple item rating scales, thus falling in between the Low and Medium levels of structure for 'Rating scales used' in our framework. Three
structural components (separate rater[s] from decision-maker[s], informed consent, and notification of results) required the SMA to be part of an actual hiring process, which was not the case for any of the studies evaluated. Initial interrater agreement ranged from 75% (assessor training) to 95% (documentation and multiple raters). Discrepancies were resolved through consensus after a close inspection of the source. Overall means and standard deviation (SD) were calculated for each component across studies.

2.2 | Results

Results of this content analysis can be seen in Table 2. Procedural consistency had the highest level of structure (M = 2.95, SD = 0.15). Multiple raters and rating scales used were moderately high in structure (Mscales = 2.48, SDscales = 0.58; Mraters = 2.43, SDraters = 0.73), while assessor training (M = 1.73; STradtraining = 0.84) and job-relatedness (M = 1.60, SD = 0.44) were both moderately low in structure. Lastly, documentation had the lowest level of structure (M = 1.13, SD of 0.44). Overall, our analysis shows that while some of the SMA structural components are generally high in structure (procedural consistency and rating scales used), other structural components are lower (assessor training, job-relatedness, and documentation). SDs for most components also demonstrate a lack of consistency in the use of SMA structure across studies.

2.3 | Discussion

As evidenced by our content analysis, there is room for improvement in SMA structure in academic research studies. The high structure found in the procedural consistency suggests that SMA research studies generally used the same process across participants. However, the lower structure found in assessor training, job-relatedness, and documentation indicate that SMAs could be better structured in academic research. There may be various reasons that academic studies do not use more structured SMAs. They may be trying to match the structure of SMAs used in practice or may be purposefully using only the elements of structure they are testing in their research design. If the goal of a research study is to determine whether SMAs can validly predict job-related criteria (KSAOs, job performance, etc.), we suggest that using highly structured SMAs is most appropriate. However, specific research examining the effects of each structural component is needed to understand their direct (an incremental) influence on SMA ratings and validity.

It is likely that the SMA structural components are observed even less in practice than in research. This supposition is supported by reports that there is little to no structure in SMAs as they are currently practiced in organizations (e.g., Roth et al., 2016). Our second study moves from research to practice by assessing the frequency that structural components are used in SMAs in actual hiring decisions.

3 | STUDY 2

This study moves from the research context of Study 1 to examine the use of SMA structure in practice. Organizations that conduct ill-structured SMAs are at risk of making ineffective and potentially unethical or illegal employment decisions. The structural components highlighted previously are recommended ways to reduce bias and make more valid and effective hiring decisions based on SMAs. It is often assumed that SMAs are not consistently used and there is little structure in the process (Ross & Slovensky, 2012; Van Iddekinge et al., 2016), yet there has been no scientific inquiry to verify this assumption. There is some initial evidence that just over half of organizations (57%) have formal or informal policies regarding SMAs (SHRM, 2013). However, given that only 20% of respondents report using SMAs in the selection, it is quite possible those respondents who have an organizational SMA policy actually use SMAs.

Because SMAs can be quickly, covertly, and easily accomplished without the knowledge of the applicant or even others in the organization, and because of the typical informal nature SMAs, it is hypothesized that current SMAs are low on all structure components.

Hypothesis 1: SMAs, as currently practiced, are significantly below Level 2 (medium structure) on each of the SMA structural components.

3.1 | Research method

3.1.1 | Participants

The sample consisted of HR professionals with current recruiting or hiring responsibilities. Surveys were emailed to members of HR professional associations, recruiters of college graduates, and alumni of HR Master’s programs at two universities. The survey link was also posted online for members of SHRM Connect (an SM platform for HR professionals; http://community.shrm.org/home) and six LinkedIn groups for HR professionals (HR Professionals Association, Human Resource Professionals, Linked: HR, PHR Linked, SHRM Official, and SHRM Networking). These various recruitment methods allowed for a broad sample representing a variety of industries, locations, and organizations.

Data for Study 2 came out of a larger study examining HR manager perceptions and the use of SMAs (Hartwell & Campion, 2020). The current study focuses on a subset of HR managers that indicated they utilize SMAs (82% of the overall sample). This resulted in 167 respondents from 32 U.S. states and 21 additional countries (78% of respondents were in the United States). The top industries represented by respondents were business services (29%), manufacturing (20%), health care (7%), education (6%), government (5%), and transportation (5%). Average experience with recruiting and/or hiring was over 12 years (M = 12.16, SD = 6.96), and recruitment and/or hiring was an average of 50% of their job duties (M = 50.12%, SD = 32.19%). On a scale from 1 (not at all familiar) and 4 (extremely familiar), the average participant indicated strong familiarity with SM, both Facebook (M = 3.60; SD = 0.67) and LinkedIn (M = 3.71;
<table>
<thead>
<tr>
<th>Study</th>
<th>Job-relatedness</th>
<th>Procedural consistency</th>
<th>Rating scales used</th>
<th>Documentation</th>
<th>Assessor training</th>
<th>Multiple raters</th>
<th>Social media platform used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kluemper and Rosen (2009)</td>
<td>2</td>
<td>3</td>
<td>2.5</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>Facebook</td>
</tr>
<tr>
<td>Bohnert and Ross (2010)</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook*</td>
</tr>
<tr>
<td>Kluemper et al. (2012)-S1</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>Facebook</td>
</tr>
<tr>
<td>Kluemper et al. (2012)-S2</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>2.5</td>
<td>Facebook</td>
</tr>
<tr>
<td>Roulin and Bangerter (2013)</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook</td>
</tr>
<tr>
<td>Van Iddekinge et al. (2016)</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1.5</td>
<td>1</td>
<td>Facebook</td>
</tr>
<tr>
<td>van de Ven et al. (2017)-S1</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>van de Ven et al. (2017)-S2</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>Pike et al. (2018)</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook* and LinkedIn*</td>
</tr>
<tr>
<td>Rosen et al. (2018)</td>
<td>1.5</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>Twitter</td>
</tr>
<tr>
<td>Roulin and Levashina (2019)-S1</td>
<td>2</td>
<td>2.5</td>
<td>2</td>
<td>1</td>
<td>2.5</td>
<td>2</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>Roulin and Levashina (2019)-S2</td>
<td>1.5</td>
<td>2.5</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>Becton et al. (2019)</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1.5</td>
<td>3</td>
<td>Facebook*</td>
</tr>
<tr>
<td>Roth et al. (2020)-S1</td>
<td>1.5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook*</td>
</tr>
<tr>
<td>Roth et al. (2020)-S2</td>
<td>1.5</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook*</td>
</tr>
<tr>
<td>Zhang et al. (2020)-S3 (Unst)</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook*</td>
</tr>
<tr>
<td>Zhang et al. (2020)-S3 (St)</td>
<td>2</td>
<td>3</td>
<td>2</td>
<td>1.5</td>
<td>3</td>
<td>2</td>
<td>Facebook</td>
</tr>
<tr>
<td>Wade et al. (2020)</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>Facebook* and LinkedIn*</td>
</tr>
<tr>
<td>Cubrich et al. (2021)</td>
<td>1</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>Fernandez et al. (2021)</td>
<td>2</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>LinkedIn</td>
</tr>
<tr>
<td>Overall mean (SD)</td>
<td>1.60 (0.44)</td>
<td>2.95 (0.15)</td>
<td>2.43 (0.58)</td>
<td>1.13 (0.44)</td>
<td>1.73 (0.84)</td>
<td>2.48 (0.73)</td>
<td></td>
</tr>
</tbody>
</table>

Note: Social media platforms marked with an asterisk (*) were artificially created to mimic a real social media platform; 1 = low structure, 2 = medium structure, 3 = low structure, S1 = Study 1, S2 = Study 2, S3 = Study 3, Unst = unstructured condition, St = structured condition. Three structural components (Separate rater[s] from decision maker[s], Informed Consent, and Notification of Results) were not applicable, as none of the studies reported SMA use in actual hiring decisions.
SD = 0.56); 99% of participants had a LinkedIn account, while 90% had a Facebook profile. Forty-four percent of participants (n = 74) were recruited by email and 56% (n = 93) via SM. To accurately represent current practices in selection, we limited the sample to those presently involved in recruiting and/or hiring. Participants were offered a report of the survey results for their participation.

3.1.2 Measures of SMA structure

Respondents were asked eight items about their typical SMA process, each corresponding to one structural component. Respondents chose one of three options, one at each level of structure (see full items in an online supplement). A sample item was labeled “Process Used” (which corresponds with the procedural consistency component) and included the following three options: The exact same procedures are used for all applicants (rated “3”), The general process is roughly similar for all applicants (rated “2”), and The same procedures are not used for all applicants (rated “1”). Because data collection occurred during a previous iteration of our model, the separate rater(s) from decision maker(s) component is slightly different than the final iteration (see specific options in the online supplement) and the multiple raters component was not included.

3.2 Results

There were few significant differences in SM structure use when comparing respondent attributes. Participants in the United States (coded 0) were less likely to use documentation than respondents from other countries (coded 1) (r = −.22, p = .001). Those with a higher percentage of their job focused on hiring were more likely to use higher levels of training in SMAs (r = .20, p = .028). Finally, more experience was correlated with higher levels of informed consent (r = .23, p = .011). No other relationships were significant.

### Table 3 Utilization of SMA structural components in practice

<table>
<thead>
<tr>
<th>Structural component</th>
<th>% of respondents selecting each level</th>
<th>Mean (SD)</th>
<th>t</th>
<th>Study 1 means (SDs) for comparison</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level 1</td>
<td>Level 2</td>
<td>Level 3</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Job-relatedness</td>
<td>51</td>
<td>17</td>
<td>32</td>
<td>1.81 (0.90)</td>
</tr>
<tr>
<td>Procedural consistency</td>
<td>14</td>
<td>58</td>
<td>28</td>
<td>2.14 (0.63)</td>
</tr>
<tr>
<td>Rating scales used</td>
<td>86</td>
<td>6</td>
<td>8</td>
<td>1.22 (0.58)</td>
</tr>
<tr>
<td>Documentation</td>
<td>61</td>
<td>31</td>
<td>8</td>
<td>1.47 (0.64)</td>
</tr>
<tr>
<td>Assessor training</td>
<td>57</td>
<td>25</td>
<td>18</td>
<td>1.62 (0.78)</td>
</tr>
<tr>
<td>Separate rater(s) from decision maker(s)</td>
<td>16</td>
<td>43</td>
<td>41</td>
<td>2.26 (0.71)</td>
</tr>
<tr>
<td>Informed consent</td>
<td>75</td>
<td>18</td>
<td>7</td>
<td>1.31 (0.59)</td>
</tr>
<tr>
<td>Notification of results</td>
<td>86</td>
<td>4</td>
<td>11</td>
<td>1.25 (0.64)</td>
</tr>
</tbody>
</table>

Note: Ns between 156 and 165; t tests compare the mean value against the scale midpoint of “2”; percentages may not add up to 100% due to rounding; Level definitions can be found in the text and in Table 1 (People Involved levels are only found in the text). Abbreviation: N/A, not applicable.

**p < .01; ***p < .001.

3.2.1 Structure of SMAs

- Hypothesis 1 was based on the previously non-tested assumption in the literature that SMAs are informal, inconsistent, and unstructured. Hypothesis 1 predicted that, on average, survey respondents would utilize each of the structural components at a level below the scale midpoint. Table 3 indicates the percentage of respondents at each level of each structural component, along with means, SDs, and t-test results testing whether the mean level of each component is different from average level (Level 2). All eight components’ means were significantly different from “2,” with six of the eight supporting hypotheses by being significantly lower: job-relatedness, rating scales used, documentation, assessor training, informed consent, and notification of results. Two components were higher than the midpoint—procedural consistency and separate rater(s) from decision maker(s).

3.3 Discussion

SMAs have made noteworthy strides recently within academic literature (Davison et al., 2012, 2016; Kluemper, 2013; Ployhart, 2012; Van Iddekinge et al., 2016). However, SMAs in applied practice are seemingly lacking in both validity and ethicality (Jeske & Shultz, 2016; Roth et al., 2016; Schmidt & O’Connor, 2016). In the academic setting, one bright outlook corresponds to the procedural consistency and measurement elements of SMAs, as demonstrated in Study 1. This makes intuitive sense, given that academic empirical studies are often tightly structured to avoid confounding errors. However, academic research is still deficient in structure of assessor training, job-relatedness, and documentation.

In applied practice, our second study demonstrates that SMAs are even further lacking in structure. With a sample of HR practitioners that perform recruitment and selection duties, most of the structured assessment elements (job-relatedness, rating scales
used, documentation, assessor training, informed consent, and notification of results) fall significantly below the mid-range level.

While these findings primarily support our hypotheses, two interesting elements that were in opposition to our prediction were procedural consistency and separate rater(s) from decision maker(s). Procedural consistency is seemingly practiced in both the academic and practitioner realms (Clark & Roberts, 2010; Davison et al., 2012; Elzewig & Peeples, 2009). This finding has important implications, as a lack of procedural consistency will introduce error, subjecting an organization to undesirable consequences such as discriminatory practices and corresponding lawsuits. SMAs that include procedural consistency are likely to reduce discrimination across candidates. With regard to separating rater(s) from decision maker(s), this result is not quite as surprising as it may initially seem. Management that will supervise the applicants is generally the primary decision makers on selection, whereas HR is more often associated with screening protocols (Heneman et al., 2019). We may assume that HR practitioners in this sample performed SMAs, and forwarded their results to direct management for final hiring decisions. This may not be an indication of purposeful or strategic thinking, as much as an inherent part of the process.

4 | RECOMMENDATIONS

Introducing higher structure has demonstrated higher validity and reliability in a number of selection practices, such as rating resumes (Buster et al., 2005), letters of recommendation (Aamodt et al., 1993), and especially interviews (Levashina et al., 2017). We theorize that adding structure to the SMA has the potential to increase SMA validity and reliability as well. For both academia and industry, we strongly suggest that a core component of all SMAs include procedural consistency. This element was salient in both academic and applied practice and provides a degree of baseline reliability for SMAs. This foundational element should aid in reducing discriminatory practices across applicants (Davison et al., 2016).

For academic research on SM and HR, there are a few considerations worth noting before implementing structure into SMAs. In the field studies, many of the structural elements apply and will largely strengthen a study’s quality. In a study where stimulus content is synthesized and designed to mimic real SM information, some structural elements such as notification will not necessarily apply. However, incorporating applicable structured elements should only act to strengthen the study’s quality. We emphasize that a study’s SMA structure can impact the results (e.g., underestimating validity when the structure is low). We encourage highly structured SMAs in academic research. However, the relevance of SMA structure in the academic setting will largely depend on the study objectives and research questions and should be implemented accordingly.

We also recommend that practitioners improve SMAs to at minimum a medium level of structure for each component. While organizations should strive for a high degree of structure for each element, we encourage organizations to avoid a low level and develop at least a medium level of structure, particularly if the resources or skills required to reach a high level of structure are unfeasible. However, we encourage SMAs to use the highest structure feasible.

5 | CONTRIBUTIONS

This study is the first to strategically integrate validated structured interview principles into the growing practice of SMAs. Integrating structural components into SMAs is likely to have a positive impact on validity, reliability, user reactions, and legality (Brandenburg, 2008). We demonstrate through two studies that SMA structural components are utilized in SMA research and in practice, though usage rates tend to be lower during actual hiring situations than in academic research. Results also show that the low utilization of SMA structure, particularly in organizational hiring contexts, leaves substantial room for improved standardization.

We hope that this study will act as a foundation for applied SMAs, as well as provide avenues for subsequent SMA research. We have provided a pathway toward developing a comprehensive model of SMAs, and it is our hope that our contributions to the literature will provide a framework that remains relevant even as SM platforms continuously evolve.

6 | LIMITATIONS AND FUTURE RESEARCH

This article conducted two independent studies to inform our results, but certain limitations are worth noting. First, as noted previously, the informed consent and notification of results components in our model are new additions from the Campion et al. (1997) interview structure framework that apply specifically in an SMA setting. These components warrant dedicated research that specifically examines their dynamics across a variety of jobs and industries to determine
whether standardization can be achieved and the effect of these components on validity and reliability.

Second, subsequent research should measure the impact of structural components on SMA outcomes (validity, reliability, user reactions, etc.). Given the mixed results regarding the predictive validity of SMAs, it is imperative to understand whether adding structure increases the predictiveness of job-related outcomes (e.g., job performance, job tenure, and organizational commitment). While this paper proposes generally positive effects of SMA structure on reliability, validity, user reactions, and legality, these propositions require empirical tests to examine whether they hold true in practice.

Third, the data gathered from HR professionals in Study 2 is self-reported. While attempts were made to reduce social desirability responding—such as ensuring that data would remain confidential—participants may still tend to respond in a socially desirable manner. There is also the possibility that their perceptions of SMA processes may not be accurate, even if they are attempting to respond accurately. Future research could measure the structure of actual SMAs directly. Another potential measurement limitation is that HR professionals responded to only one item per structural component to assess their level of SMA structure. This was done to limit fatigue, as this was part of a much larger survey, and because we felt the items were clearly measured in a single item. However, more items per competency would likely give more nuanced and accurate ratings.

7 | CONCLUSION

This study takes a well-validated, commonly utilized structural framework for traditional personnel selection and applies it to the growing practice of SMAs. Our findings provide the foundation for a viable approach toward SMAs, using levels of structure that can be implemented in research and practice. While certain elements should be investigated further, such as the ethical considerations associated with the acceptability of SMAs, other components such as procedural consistency can be promptly implemented. We are hopeful that this study will act as a foundation for subsequent academic research, and simultaneously provide strong recommendations for SMAs conducted by applied practitioners.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Christopher J. Hartwell http://orcid.org/0000-0002-3405-3229
Rahul S. Chauhan http://orcid.org/0000-0003-2389-2219

REFERENCES


CareerBuilder. (2018, August 9). More than half of employers have found content on social media that caused them NOT to hire a candidate, according to recent CareerBuilder survey. http://press.careerbuilder.com


HARTWELL ET AL.


