Old Dominion University
ODU Digital Commons

Psychology Theses & Dissertations | Psychology

Summer 2018

Detecting Deceptive Impression Management Behaviors in Interviews Using Natural Language Processing

Elena Margaret Lawrence Auer
Old Dominion University, eauer22@gmail.com

Follow this and additional works at: https://digitalcommons.odu.edu/psychology_etds

Part of the Industrial and Organizational Psychology Commons

Recommended Citation
Auer, Elena Margaret L.. "Detecting Deceptive Impression Management Behaviors in Interviews Using Natural Language Processing" (2018). Master of Science (MS), Thesis, Psychology, Old Dominion University, DOI: 10.25777/yx69-dy97
https://digitalcommons.odu.edu/psychology_etds/70

This Thesis is brought to you for free and open access by the Psychology at ODU Digital Commons. It has been accepted for inclusion in Psychology Theses & Dissertations by an authorized administrator of ODU Digital Commons. For more information, please contact digitalcommons@odu.edu.

DETECTING DECEPTIVE IMPRESSION MANAGEMENT BEHAVIORS IN INTERVIEWS USING NATURAL LANGUAGE PROCESSING

by

Elena Margaret Lawrence Auer
B.A. May 2016, George Washington University

A Thesis Submitted to the Faculty of
Old Dominion University in Partial Fulfilment of the
Requirements for the Degree of

MASTER OF SCIENCE

PSYCHOLOGY

OLD DOMINION UNIVERSITY
August 2018

Approved by:

Richard N. Landers (Director)

Xiaoxiao Hu (Member)

Alan Meca (Member)

ABSTRACT

DETECTING DECEPTIVE IMPRESSION MANAGEMENT
BEHAVIORS IN INTERVIEWS USING NATURAL
LANGUAGE PROCESSING

Elena Margaret Lawrence Auer
Old Dominion University, 2018
Director: Dr. Richard N. Landers

Deceptive impression management (IM) is often used by applicants in employment interviews to improve their chances of receiving a job offer. Self-report measures of deceptive IM are typically used to evaluate interview faking in a lab setting but are limited when used in practice due to social desirability concerns. Given this limitation, natural language processing (NLP) has potential as a tool to unobtrusively assess raw interview content and measure deceptive IM. This study examined the use of open and closed-vocabulary NLP approaches for the detection of deceptive IM in mock employment interviews. In general, neither of these approaches successfully predicted deceptive IM. Several possible conclusions based on these findings are discussed. However, given the lack of empirical support for this method, organizations should proceed with caution when deciding to use NLP techniques to predict deceptive IM in employment interviews.

ACKNOWLEDGMENTS

Thank you to my advisor, Richard Landers, for supporting me throughout this entire process with your vast knowledge, incredible insight, and encouragement. Your mentorship has been invaluable these past two years and I am incredibly grateful for the countless hours and effort you have spent on advising me.

Thank you to my committee members, Xiaoxiao Hu and Alan Meca for reviewing this work and providing helpful feedback. Your advisement was critical to the success of this project. Thank you to Peggy Kinard, Linda Forchas, and Mary Boswell for your help with the administrative steps of this project and for all your help throughout these past two years.

Thank you to Bo Armstrong, Andrew Collmus, and Sebastian Marin for being the best lab mates anyone could ask for. Your mentorship, friendship, and comic relief kept me going and I seriously could not have done this without you all.

Thank you to my roommate, Kim Perry, for being a wonderful friend. Your constant encouragement and support means the world to me. Your willingness to watch trash TV with me means even more.

Thank you to my dad, Ed Auer, for being my infallible support system. Even when things get tough, you never fail to make me laugh or help put things in perspective. Thank you to my mom, Sandra Lawrence-Clarke, for your endless encouragement and wisdom.

Thank you to my cats, Tito and Roux, for always being there to cuddle at the end of the day, even if it was just to get wet food.

Lastly, thank you to Matt Newman for being my rock. Your love, support, and friendship kept me going throughout this process and I could not ask for a better person to share these successes with.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Deceptive impression management (IM) in an employment interview setting has broad and significant implications for both individuals and organizations. Applicants invent and distort their interview answers to improve interview outcomes and create false but favorable impressions (Levashina & Campion, 2007). This deceptive IM can affect the outcomes of interviews by creating construct contamination (Donovan, Dwight, Schneider, 2014), ultimately attenuating the relationship between observed scores and future job performance (Levashina & Campion, 2006). Understanding if and to what extent applicants engage in deceptive IM is especially important because these behaviors affect personnel decisions, which ultimately play a role in broader organizational success (Huselid, 1995; Becker & Gerhart 1996; Bowen & Ostroff, 2004).

Deceptive IM is frequently used by interviewees to increase their chances of obtaining employment opportunities, which can harm the validity of an interview. Across multiple studies, over 90% of applicants reported deceptive IM behavior during employment interviews (Levashina & Campion, 2007). However, not all applicants fake to the same extent, meaning some applicants obtain higher deceptive IM scores than others based on their deceptive IM behaviors (Levashina & Campion, 2007). Deceptive IM behavior also affects interview ratings. For example, extensive image creation increases the probability of getting an interview or job offer even after interview experience and GPA are accounted for (Levashina & Campion, 2007). Further, deceptive IM often goes undetected by the interviewer, making it difficult to control for in interviewer ratings (Roulin, Bangerter & Levashina, 2014). Lastly, although structured interviews tend to reduce the occurrence of faking, deceptive IM occurs in both structured and unstructured interviews and is thus difficult to prevent by solely changing interview methods (Ellis, West, Ryan & DeShon, 2002; Tsai, Chen & Chiu, 2005; Levashina & Campion, 2006). In summary, deceptive IM is prevalent, effective, goes undetected, and occurs despite the structure of an interview, ultimately threatening the validity of an interview.

Self-report measures of deceptive IM are typically used to evaluate interview faking in the lab, both because individuals are aware of their own tactics (Bolino & Turnley, 1999) and because third-party raters cannot accurately detect deception (DePaulo, Stone & Lassiter, 1985, Bond & DePaulo, 2006). However, there are several drawbacks to using self-report measures in applied contexts. Interviewees may be reluctant to report deceptive tactics due to social desirability influences (Ellingson, Smith, & Sackett, 2001; Pauls & Crost, 2003) or fear of adverse effects on employment opportunities, and these field effects can influence the observed validity of self-reported deceptive IM. Thus, it would be beneficial to be able to measure deceptive IM unobtrusively. Some researchers attempted to do this using third-party ratings of impression management (Stevens & Kristof, 1995), but the raters found it difficult to identify deceptive IM because intent cannot be directly observed (Levashina & Campion, 2007).

Given the limitations of both traditional self-report deceptive IM measurement and third-party ratings, natural language processing (NLP), which broadly refers to the creation of datasets from unstructured text sources, has potential as a tool to unobtrusively assess raw interview content and measure deceptive IM without many of those limitations. The use of NLP in psychology generally refers to one or two major general approaches for analyzing and representing text: open-vocabulary (O'Connor, Bamman & Smith, 2011; Grimmer & Stewart, 2013; Park et al., 2014) and closed-vocabulary approaches (Golbeck, Robles, & Turner, 2011; Holtgraves, 2011). Open-vocabulary approaches involve data-driven models of language (Blei, Ng & Jordan, 2003). For example, in one type of open-vocabulary approach, linguistic features that are most predictive of the target outcome will emerge by predicting an outcome from raw word count data (e.g., Park et. al., 2015). In contrast, closed-vocabulary approaches are often based upon existing theory and empirical evidence on linguistic features. These approaches typically involve the use of word dictionaries, which relate words and word families with some target characteristic, such as anxiety or happiness. Using this approach in the deceptive IM context, existing theory on linguistic features associated with deception, impression management, and faking in interviews can be used to inform specific hypotheses about which linguistic features are most likely related to deceptive IM (e.g., Hauch, Gitlin, Masip & Sporer, 2015; Culbertson, Weyhrauch & Waples 2016). However, it is not clear which approach is best, in terms of both predictive power and psychometric soundness, for capturing psychological behavior through text. The purpose of the present study is therefore to fill this gap by examining these two NLP approaches to measuring deceptive IM in interviews.

Deceptive Impression Management

Deceptive IM in the interview context is defined as the "conscious distortions of answers to the interview questions in order to obtain a better score on the interview and/or otherwise create favorable perceptions" (Levashina & Campbell, 2007, p. 1639). This definition integrates literature on impression management from both a personality perspective (Paulhus, 1984), which differentiates between intentional and unintentional tactics, and the social behavioral perspective (Schlenker, 1980; Gilmore, Stevens, Harrell-Cook, & Ferris, 1999), which differentiates between honest and dishonest IM. In contrast to honest IM, which involves highlighting one's attributes and credentials without distortion, deceptive IM is intentional misrepresentation. In contrast to unintentional IM tactics, such as in self-deception, deceptive IM is intentional. The taxonomy of interview faking by Levashina and Campion (2007) identifies four factors: slight image creation, extensive image creation, image protection, and ingratiation. Slight image creation includes embellishing and tailoring answers, or fit enhancing to make an image of a good job candidate. Extensive image creation includes constructing stories or experiences, inventing better answers, and borrowing experiences or accomplishments of others to invent an image of a good job candidate. Image protecting includes omitting information, masking details of the past, or distancing oneself from negative events to defend an image of a good job candidate. Finally, ingratiation includes conforming one's expressed opinions to those of the interviewer or organization, as well as praising or complementing the interviewer or organization to gain favor with the interviewer and improve the appearance of a job candidate.

Interviewees who are successful at deceptive IM are likely not representing their "true" self in an interview (Levashina & Campion, 2007), impacting how well an interviewer can assess interviewee competencies and future job performance and potentially leading to poor hiring decisions (Donovan, Dwight & Schneider, 2014). For example, IM tactics, which includes deceptive IM, are only modestly correlated with job performance but impact interviewer ratings more than validated predictors of performance, including cognitive ability and conscientiousness (Barrick, Shaffer & DeGrassi, 2009). Carlson, Carlson, and Ferguson (2011) found that employee use of deceptive IM in the workplace was associated with low ratings of promotability and low leader-member exchange. Further, job applicants who used deceptive IM during the selection process engaged in more counterproductive work behaviors once they were on the job (O'Neill, Lee, Radan, Law, Lewis & Carswel, 2013).

Developing an approach to text-based measurement of deceptive IM may help account for these concerns in both traditional and future selection practices by informing organizations to take a second, more critical review of applicants when high deceptive IM is detected. In the case of traditional interviews, where an interviewer is rating an applicant based on their performance, using NLP as an additional screening tool to alert an interviewer of deception could be used to mitigate some of the impact on validity. Additionally, using NLP to detect deception could be especially useful as organizations increasingly use completely algorithmic interview scoring (Chapman & Webster, 2003; Chen et al., 2016; Chen et al., 2017). For example, the company HireVue uses a combination of NLP, voice analysis, and facial expressions to assess future job performance of applicants in asynchronous video interviews (hirevue.com). If the applicant is faking, the prediction accuracy of future job performance using that data may be adversely impacted (Donovan, Dwight, Schneider, 2013).

Natural Language Processing

NLP refers to a range of techniques for analyzing and representing text to achieve human-like language processing by computers (Liddy, 2001). NLP approaches can be classified into one or more of three primary goals: syntactic (structural aspects of language), semantic (meaning of language), and pragmatic (context-dependent meaning) understanding of text (Cambria & White, 2014). For example, in the sentence "I'd like to meet for lunch today," the structural aspects refer to the parts of speech used (i.e., pronouns usage, punctuation, etc.). The semantic understanding refers to the meaning of the sentence, which in this case refers to meeting for lunch. Finally, the pragmatic interpretation of this sentence would be in the context of the rest of the conversation and, for example, convey that a supervisor wants to have lunch with an employee to discuss performance. Choosing an NLP approach depends on the goal of the research or project. In psychological research, primarily due to practical limitations, researchers have traditionally focused on syntactic and semantic understanding of text using theory-driven closed-vocabulary approaches (Pennebaker, Booth, Boyd & Francis, 2015). More recently, a few psychological studies have used open-vocabulary, data-driven approaches (Schwartz et al., 2013; Kern et al., 2014; Park et al., 2014; Campion, Campion, Campion & Reider, 2016; Kulkarni et al., 2017). Although some general comparisons have been made between open and closed-vocabulary approaches, differences in these approaches in the context of deceptive IM are not yet clear.

Closed-vocabulary approaches, also referred to as using user-defined dictionaries or lexica, is historically the most common NLP approach to text analysis in psychology. In these approaches, lists of words are grouped together by a theoretical psychological cause, and total word frequency within those lists yields a score for that cause. The Linguistic Inquiry and Word Count software (LIWC) is the most common example of this method (Pennebaker, Francis & Booth, 2001). From a psychological perspective, this approach allows researchers to take a theory-driven approach to measurement using word frequencies. For example, Pennebaker and Stone (2003) measured self-focus of aging participants using pronoun usage, because it is theorized that self-focus can be captured by pronoun use. However, there are drawbacks to closed-vocabulary approaches because they are limited to a priori assumptions about relevant features which, although derived from psychological theory, is uncommon in relation to how other disciplines use NLP. For example, in computer and data science, open-vocabulary approaches are dominant. Further, closed-vocabulary approaches are not sensitive to context, irony, sarcasm, or idiomatic expression.

In contrast, open-vocabulary approaches take a data-driven approach to text analysis using machine learning algorithms to comprehensively extract language features (Kosinski, Stillwell & Graepel, 2013; Park et. al., 2015). A few examples of open-vocabulary techniques include topic modeling (Blei, Ng & Jordan, 2003), using words and phrases as predictors (n-grams; e.g., Schwartz et al., 2013), and automatic summarization (Mani, 2001). Because open-vocabulary approaches are data-driven, they are not limited to a priori assumptions about which linguistic features are relevant to measuring the psychological construct, so they are more robust to unconventional language use and can provide insight about constructs lacking theoretical links to language use. These approaches also typically include more information (i.e., more words) in the predictive models, improving prediction capabilities (Iacobelli, Gill, Nowson & Oberlander, 2011). However, the development of statistical models using open-vocabulary approaches typically requires a dramatically greater sample size to avoid the number of variables (i.e., words and word groupings) dramatically exceeding the sample size. For example, in a simple bag of words approach, individual counts of all words found across a dataset might be used as predictors in a statistical model of an outcome of interest. However, depending on the amount of text analyzed, there might be tens of thousands of words, making sample size requirements extraordinarily high. Topic modeling, which is the open-vocabulary technique used in the current study, is a form of data reduction that can be used to avoid using single words as predictors and reduce the sample size requirement.

Open and closed-vocabulary approaches both offer benefits and drawbacks in the context of measuring psychological constructs. Open-vocabulary approaches tend to surpass theory-driven closed-vocab text analysis approaches in terms of predictive power (Iacobelli, Gill, Nowson, and Oberlander, 2011; Schwartz et al., 2013) because of the greater amount of information used for prediction. For example, Newman, Pennebaker, Berry, and Richards's (2003) closed-vocabulary approach predicted deception when talking about abortion-related views with 67% accuracy, whereas an open-vocabulary approach was found to predict deceptive product reviews with over 90% accuracy (Li, Cardie & Li, 2013). However, in contrast to open-vocabulary approaches, closed-vocabulary dictionaries provide a consistent form of construct measurement across studies. In open-vocabulary approaches, because variables are derived from the words used in a particular dataset, the resulting models can be much more difficult to explain and generalize to other contexts. For example, a recent meta-analysis examined computer-detected deception of studies primarily using LIWC dictionaries (Hauch, Gitlin, Masip & Sporer, 2015), which was only possible because these dictionaries were consistent across studies and utilized a closed-vocabulary approach. Additionally, previous studies have also found evidence of poor discriminative and convergent validity when using an open-vocabulary approach, which can be problematic when using this approach as a form of psychometric measurement (Park et al., 2014). Thus, considering the potential benefits and drawbacks of both open and closed-approaches when measuring psychological constructs, it is important to examine both approaches for both predictive power and psychometric soundness in the prediction of deceptive IM. Additionally, considering the lack of clear theory when it comes to detecting deceptive IM using language, both a theoretically-driven and data-driven approach are useful for better understanding how language is used differently by interviewees engaging in deceptive IM.

Closed-vocabulary Approach: Linguistic Inquiry and Word Count (LIWC)

When using a closed-vocabulary approach to text analysis, existing theory and empirical evidence serve as the basis for a priori selection of linguistic features. In the context of deceptive IM, choosing linguistic features of interest can be informed by a blend of impression management and deception theory (e.g., Hauch, Gitlin, Masip & Sporer, 2015; Culbertson, Weyhrauch & Waples, 2016). Impression management theory posits that one way people alter a target’s opinion is though verbal behaviors and tactics such as making statements about one’s accomplishments or apologizing to the target (Tedeschi, 1981). For example, a person engaging in impression management in an interview might discuss their accomplishments in detail rather than discuss their weaknesses. An interviewee engaging in impression management also might apologize to the interviewer for not understanding the interview question. Thus, measuring altered verbal behavior is fundamental to measuring deceptive IM behaviors in interviews. Because measurement of verbal behaviors is typically operationalized by self-report measures of third-party raters, there is scant literature on specific linguistic features associated with either honest or deceptive impression management. However, there is extensive theoretical and empirical evidence for the use of linguistic features to identify deception, which can inform hypotheses about linguistic correlates of deceptive IM. Previous research on deception has generally identified four categories of linguistic features associated with deception: total number of words, pronoun use, emotion, and markers of cognitive complexity (Newman, Pennebaker, Berry & Richards, 2003; Hauch, Gitlin, Masip & Sporer, 2015). Existing research on linguistic cues related to IM or deception in interviews has focused on the first three (total number of words, pronoun use, and emotional valence of words) and have additionally explored which topics people discuss when engaging in impression management. Thus, it is likely that a dictionary-based measure of deceptive IM (total number of words, emotional valence of words, pronoun use, and topics) will predict self-reported scores of deceptive IM.

Total Number of Words. Fabricators tend to use fewer words than those telling the truth because deception is more cognitively demanding than truth-telling (Porter & Yuille, 1996; Vrij, 2000; Burgoon, Blair, Qin, & Nunamaker, 2003; DePaulo et al., 2003; Hauch, Gitlin, Masip & Sporer, 2015). Deception can be more demanding because it involves suppressing thoughts about the truth (Gombos, 2006), monitoring self-behaviors and observer reactions (Buller & Burgoon, 1996), and relying on semantic memory rather than episodic memory to construct a lie (Schank & Abelson, 1997). Because of these cognitive demands, theory and evidence posit that liars use fewer words due to the inherent cognitive difficulty of producing a lie (Hauch, Gitlin, Masip & Sporer, 2015). This finding is indirectly replicated in the literature on deception in interviews, such that applicants using deception (operationalized by being asked to lie in response to interview questions) in employment interviews were coded as using fewer details and descriptions than applicants not using deception (Culbertson, Weyhrauch & Wables, 2016). Interestingly, Schneider, Powell, and Roulin (2015) found that deceptive IM in interviews, particularly slight image creation and deceptive ingratiation, were associated with fewer pauses in speech, showing a lack of restraint of verbal behavior. The authors suggested that the added cognitive load in interviews may reduce a person's ability to regulate their output, making them more prone to pausing. While an interview context may limit the extent to which deception leads to reduced word use, overall word count will likely predict deceptive IM.

Hypothesis 1a. Total number of words used by the interviewee will predict self-reported scores of deceptive IM such that fewer words will be positively related to deceptive IM.

Word Emotional Valence. Fabricators tend to use more words associated with negative emotion (Hauch, Gitlin, Masip & Sporer, 2015). This is likely due to feelings of guilt, fear of getting caught, or feelings of being uncomfortable (Ekman, 2001; DePaulo et al, 2003; Vrij, 2008). Honesty has also been associated with positive word-use when examining text from Facebook posts (Hall & Pennington, 2013). Although negative word use is more prevalent in general deception, the use of negative words in a selection context may be more nuanced. Applicants may be purposely avoiding negative topics or trying to create a positive image of themselves to improve their chances of obtaining a position. However, there seems to be an overarching trend of more negative word use when engaging in deceptive IM. For example, when examining deceptive IM in interviews, Culbertson, Weyhrauch, and Waples (2016) found that applicants overall tend to use more negative statements and complaints when deceiving than applicants who were not engaging in deceptive impression management. Thus, most of the previous theoretical and empirical evidence suggests that, overall, negative word use will likely positively predict deceptive IM and positive word use will likely negatively predict deceptive IM.

Hypothesis 1b. Negative word-use will predict self-reported scores of deceptive IM such that increased negative word-use will be positively related to self-reported deceptive IM.

Hypothesis 1c. Positive word-use will predict self-reported scores of deceptive IM such that increased positive word-use will be negatively related to self-reported deceptive IM.

Pronoun usage. In attempting to distance themselves from events, those engaging in deception use fewer self-references (i.e., first-person pronouns) and more other-references (i.e., second and third person pronouns; DePaulo et al., 2003; Newman, Pennebaker, Berry & Richards, 2003). Using other-references is a way of avoiding ownership and responsibility for the described action or event. In other words, non-immediacy can indicate deception, because someone conveying a truthful message would have a strong and immediate connection to the event they were describing (e.g., Zhou & Zhang., 2004). In employment interviews, Culbertson, Wehrauch, and Waples (2016) found that verbal and vocal distancing occurred more frequently when applicants were engaging in deceptive IM. When applicants were engaging in deceptive IM, they were indirect, evasive, impersonal, and unclear. Another study examined the use of pronoun usage to detect IM in resumes (Weaver, 2017). Grounded in findings that self-monitoring is associated with second and third-person pronouns while Machiavellianism is associated with first-person pronouns (Ickes, Reidhead & Patterson, 1986), Weaver (2017) correlated pronoun usage with impression management in resumes. In contrast to Barnes and Ickes' results, Weaver found that self-oriented impression management, linked to self-monitoring, was negatively related with first-person pronouns and positively correlated with impersonal pronouns. Other-focused impression management was positively correlated with first-person pronouns, which again was contrary to Barnes and Ickes' findings. Weaver suggested, however, that deception may be an explanation for these unexpected findings, and that people who are deceiving may distance themselves from a lie using second and third-person pronouns. Thus, it is likely that pronoun usage will predict deceptive IM. However, given the lack of consistency of theoretical and empirical evidence regarding the directionality of each of these relationships, directionality was not explicitly hypothesized.

Hypothesis 1d. A dictionary-based measure of pronoun usage will predict self-reported scores of deceptive IM such that other-references pronouns (second, third-person, and impersonal (it, those, etc.) pronouns) will collectively be related to deceptive IM.

Hypothesis 1e. A dictionary-based measure of pronoun usage will predict self-reported scores of deceptive IM such that self-reference (first-person singular and plural) pronouns will collectively be related to deceptive IM.

Topics of Conversation. Words related to topics such as "fit," "achievement," "values," "family," and "leisure" have all be empirically linked to impression management (Ellis, West, Ryan & DeShon, 2002; Holoien & Fiske, 2013; Hall & Pennington, 2014; He, Glas, Kosinski & Veldkam, 2014; Waung, McAuslan, DiMambro, Miegoc, 2017). Of these, achievement, family, and leisure are LIWC dictionaries. Topics of achievement have been related to self-promotion impression management tactics (Holoien & Fiske, 2013). Achievement, in the context of an interview, is a relevant topic to discuss and thus more likely to occur despite the honest or deceptive intentions of the applicant. However, considering many of the verbal deceptive IM behaviors include enhancing, exaggerating inventing, and borrowing accomplishments, skills, and credentials, it is likely that applicants using deceptive IM will more often discuss the topic of achievement than those not using deceptive IM.

Additionally, the topics of “family” and “leisure” have been indirectly linked to deceptive impression management by way of association with both honesty and self-monitoring. Honesty is positively related to using words related to family (Hall and Pennington, 2013), and self-monitoring is negatively related to using words related to family (He, Glas, Kosinski & Veldkam, 2014). Given these relationships, it is likely that applicants using deceptive IM will not use words related to family (i.e., in the LIWC family dictionary). Similarly, talking about leisure has been positively related to self-monitoring (He, Glas, Kosinski, Stillwell & Veldkam, 2014) and positively related to deception (Hauch, Gitlin, Masip & Sporer, 2015). In contrast to discussing something personal, like family, leisure may be a topic more suited for deception. In an interview setting, this may be an ingratiation tactic to find shared interests with the interviewer even if the applicant is bending the truth about their actual interests.

Lastly, LIWC’s “authentic” dictionary will likely relate to deceptive IM. Authenticity, added to the 2015 set of LIWC dictionaries, is a summary variable using findings from Newman, Pennebaker, Berry, and Richards (2003)’s study that examined the use of linguistic features in detecting dishonesty. Because this summary variable is one of four proprietary variables in the LIWC program, it is unclear what the actual words on this list are. However, given the similarity of the findings of the Newman, Pennebaker, Berry, and Richards (2003) study to many of the hypothesized relationships, it is likely that authenticity will predict deceptive IM. Although some empirical evidence exists for the directionality of the relationships of these topics with deceptive IM, there is not strong enough theoretical and empirical evidence to explicitly hypothesize about the directionality of each relationship. Thus, it is hypothesized that as a group, these LIWC topics will predict deceptive IM.

Hypothesis 1f: A dictionary-based measure of topics will predict self-reported scores of deceptive IM as a group.

Open Vocabulary Approach: Latent Dirichlet Allocation (LDA) to Extract Topics

One prominent open-vocabulary approach is latent Dirichlet allocation (LDA) in which topics are extracted from unstructured text (Blei, Ng & Jordan, 2003). LDA, also referred to as topic modeling, is conceptually similar to factor analysis, such that words are grouped based on similarity. In the context of word usage, this similarity is determined by word co-occurrence. Thus, LDA in a measurement context can be used such that words (measures) are modeled as observed indicators of latent topics (constructs). This makes it a potentially useful way of identifying information from text that will detect deceptive IM usage. LDA operates under the assumption that document word counts (i.e., interviewee's response) are caused by a mixture of latent topics (Blei, Ng & Jordan, 2003). Using a Bayesian optimization technique, topics, which are clusters of related words, are extracted from text. For example, in a set of documents discussing pets, the topic "dog" will likely emerge and include related words such as "puppy", "fetch" and "leash." After topic extraction, participants are then represented as their probability of using the discovered topics. In other words, the probability of their word use is a function of the individuals' word use and the probability of the topic given that word. For example, a participant using the words "puppy" and "fetch" yields a higher probability of using the "dog" topic than a participant using the words “cat” and “mouse.” Importantly, any given text will produce a probability for each topic extracted, so topics are not mutually exclusive.

Because latent deceptive IM is theorized to cause both verbal behavior and self-reported faking, it is expected that the measures should converge in the interview context. Deceptive IM is the conscious altering of interview answers to obtain a better score on an interview (Levashina and Campion, 2007). Thus, when an interviewee is partaking in deceptive IM, the topics they choose to discuss are likely caused by latent deceptive IM. For example, the topic of achievement may emerge from the interviews and include words such as “performance”, “success”, and “effort.” A participant embellishing their achievements may yield a higher probability of using the “achievement” topic than a participant not embellishing. Open-vocabulary approaches have already been used for detecting deception. For example, topic modeling has been used to detect deceptive spam in product reviews (Li, Cardie & Li, 2013). Using LDA, the topic-based model differentiated between deceptive and truthful reviews with 95% accuracy. Similarly, topic usage as determined by LDA will likely predict deceptive IM.

Hypothesis 2. Interviewee LDA topic scores will predict self-reported deceptive IM.

Combining and comparing approaches

Because LDA- and LIWC-based approaches are different in the information they extract from text, in combination they will likely predict more variance in self-reported deceptive IM than either approach alone. Specifically, LDA- and LIWC-based approaches likely capture distinct and overlapping variance in word usage related to deceptive IM. A LIWC-based approach is limited to a priori theory about the relationship between linguistic features and deceptive IM, whereas an LDA-based approach can reveal linguistic features relevant to deceptive IM that have not yet been theorized, suggesting incremental variance in IM can be explained by LDA beyond LIWC. In contrast, LIWC-based prediction is tailored to the specific situational context in which its topics are to be used, which may enable the identification of more subtle constructs than LDA would identify, which is designed to maximize the amount of variance that can be explained in word usage using as few categories as possible. Thus, it is likely that LDA- and LIWC-based approaches independently contribute to the text-based prediction of deceptive IM.

Hypothesis 3a. In combination, an LDA- and LIWC-based approach will predict more variance in self-reported deceptive IM than LDA topics alone.

Hypothesis 3b. In combination, an LDA- and LIWC-based approach will predict more variance in self-reported deceptive IM than LIWC-based linguistic features alone.

CHAPTER 2

METHOD

Participants

An a priori power analysis was conducted to identify a target sample size. To detect small effect sizes for H1a-f, as estimated by studies examining deception and IM using LIWC dictionaries (Hauch, Gitlin, Masip & Sporer, 2015; He, Glas, Kosinski & Veldkam, 2014; Holoien & Fiske, 2013), a power analysis suggested a total sample size of 395 participants. To test H2, a total sample size of 167 participants was needed to achieve sufficient power. To test H3a-b, a total sample size of 215 participants was needed to achieve sufficient power. Although H1a-f required a much larger sample size to detect small effects, the statistical significance of each coefficient associated the LIWC categories was not practically relevant to the overall goal of detecting deceptive IM in interviews. Rather, the test of interest was their effect in combination. Thus, the focus was on obtaining enough participants to test H3a-b.

Due to external constraints, 165 participants were ultimately recruited in this study from a large public mid-Atlantic university in the United States. Most participants were recruited from a psychology student participant pool, in which participants received extra credit or fulfilled course requirements for their participation in the study (N = 133). Participants were also recruited in a general on-campus recruiting effort using email announcements, social media, and flyers and were offered a $10 Amazon gift card for their participation (N = 32). To ensure realistic participant effort in interviews, participant screening procedures from a similar study using a student sample was followed (Schneider, Powell & Roulin, 2015). Participants were screened using the post-interview filter question, “I took the mock interview as seriously as I would normally take a real interview.” Any participant indicating ‘strongly disagree’ or ‘disagree’ was removed from the sample, which resulted in the removal of 18 participants. Three attention check questions were also employed to ensure data quality. Six additional participants that did not pass two of the three attention checks were removed from the dataset, resulting in a final sample size of 141 participants. Sample characteristics are listed in Table 1.

Table 1

Sample Characteristics (N=141)

	N	%
Race		
European American or White	47	33.33%
African American or Black	58	41.13%
Asian American	3	2.13%
Native American or Native Alaskan	1	0.71%
Pacific Islander or Native Hawaiian	2	1.42%
two or more races	18	12.77%
“other’ American	8	5.67%
“other” not American	4	2.84%
Gender		
Male	24	17.02%
Female	117	82.98%
Age		
<24 years old	118	83.69%
> 24 years old	23	16.31%
Employment Status		
Full-Time	24	17.02%
Part-Time	73	51.77%
Unemployed	44	31.21%
GPA		
< 2.00	57	40.43%
>2.00	84	59.57%
Year in School		
Freshman	48	34.04%
Sophomore	28	19.86%
Junior	28	19.86%
Senior	31	21.99%
Graduate	2	1.42%

Non-traditional	1	0.71%
Non-degree seeking	3	2.13%
Has attended an Interview Workshop		
No	118	83.69%
Yes	23	16.31%
# of Employment Interviews (Past Year)		
0	41	29.08%
1	42	29.79%
2	29	20.57%
3	19	13.48%
4+	10	7.09%
# of Employment Interviews (Lifetime)		
0	6	4.26%
1	11	7.80%
2	23	16.31%
3	19	13.48%
4	26	18.44%
5+	56	39.72%
Looking for Job in the Next Year		
Yes	97	68.79%
No	44	31.20%

Materials and Measures

Interview. Participants were asked to complete an asynchronous interview using the online interview platform HireVue. Asynchronous interviews, also known as one-way interviews, require applicants to record a video of themselves answering a predetermined set of interview questions. This method is becoming increasingly popular in organizations and interviews are starting to be automatically scored using NLP techniques (Chen et al 2016; Feloni, 2017). Given this shift towards using NLP to score one-way interviews, examining deceptive IM using NLP is especially relevant and practical in this context. Although differences in deceptive IM in asynchronous versus synchronous interviews have yet to be explored, deceptive IM has been empirically demonstrated to occur in asynchronous mock interviews (Silva, 2016) like this one.

Demographics. Participants were asked to report their age, gender, race, and employment status. Participants were also asked to report their year in school and GPA.

Self-Report Deceptive Impression Management. Self-reported deceptive IM was measured using Levashina and Campion's Interview Faking Behavior Scale (Appendix B). This scale consists of four factors: slight image creation, extensive image creation, image protection, and ingratiation. All subscales were administered using a 5- point Likert scale (1 = to a little extent, 5 = to a great extent) and scores of each scale were averaged to yield an overall deceptive IM composite score (Cronbach's $\alpha = .94$). This scale was validated in a six-part study that examined content-related, convergent, discriminant, and criterion-related validity evidence in addition to evidence of test-retest reliability (Levashina & Campion, 2007). Content validity was established using an expert panel of judges, and convergent and discriminant validity were established by examining the correlation between the Interview Faking Behavior scale and related (social desirability, attitudes towards dishonesty, self-monitoring, and Machiavellianism) and unrelated measures (gender and GPA). Test-retest reliability was established by administering the scale two times using one-month intervals and criterion-related validity was established by examining the effect of faking on interview outcomes.

Slight image creation scale. Slight image creation (Cronbach's $\alpha = .86$) is the attempt of an interviewee to create an image as a good job candidate. There are three subscales of slight image behavior: embellishing, tailoring, and fit enhancing. Embellishment, which consists of four items, is embellishing answers beyond the truth and includes items like "I exaggerated my future goals." Tailoring, which consists of six items, is modifying answers to fit the job that the applicant is applying for and includes items like “I distorted my work experience to fit the interviewer’s view of the position.” The third subscale, fit-enhancing, is not relevant in a mock-interview context and therefore were not be measured.

Extensive image creation scale. Extensive image creation (Cronbach’s α = .89) is the attempt of an interviewee to invent an image as a good job candidate. There are three subscales of extensive image behavior: Constructing, inventing, and borrowing. Constructing, which consists of seven items, is building stories using previous work experience and includes items like “I told fictional stories prepared in advance of the interview to best present my credentials.” Inventing, which consists of seven items, is creating better answers and includes items like “I stretched the truth to five a good answer.” Borrowing, which consists of three times, is answering the question using other’s experiences or accomplishments and includes items like “I described team accomplishments as primarily my own.”

Image protection scale. Image protection (Cronbach’s α = .89) is the attempt of an interviewee to defend an image as a good job candidate. There are three subscales of image protection: Distancing, masking, and omitting. Distancing, which consists of three items, is improving answers by separating them from negative experiences and includes items like “I tried to suppress my connection to negative events in work history.” Masking, which consists of four items, is creating better answers by concealing aspects of background and includes items like “I covered up some “‘skeletons in my closet’.” Omitting, which consists of four times, is not mentioning things to improve answers and includes items like “I tried to avoid discussing my lack of skills or experience.”

Ingratiation scale. Ingratiation (Cronbach’s α = .90) is the attempt of an interviewee to gain favor with the interviewer. There are two subscales of image protection: Opinion conforming and interviewer or organization enhancing. Opinion conforming, which consists of eight items, is expressing beliefs, values or attitudes similar to that of the interviewer and includes items like “I did not express my opinions when they contradicted the interviewer’s opinions.” Interviewer or organization enhancing, which consists of four items, is creating better answers by concealing aspects of background and includes items like “I covered up some “‘skeletons in my closet’.” Interviewer or organization enhancing, the second subscale, is not relevant in a mock-interview context and therefore were not be measured.

Filter Questions. Three attention check questions were employed to ensure data quality (Appendix C). Participants were asked to indicate the desired answer for two of the questions (i.e., please select “To no extent.”) and were also asked to respond to the question (“I have lived in Antarctica”). Participants were also asked to indicate on a five-point Likert-type question the degree to which they agree or disagree to the statement “I took the mock interview as seriously as I would normally take a real interview” (Schneider, Powell & Roulin, 2015).

Interview Experience. Participants were asked to indicate their interview experience using a 4-item interview experience scale (Silva, 2016; Appendix D). Participants reported the number of employment interviews they have completed, both in the past year and in their lifetime. The participants were also asked to indicate “yes” or “no” to the questions: “Will you be looking for a job in the next year?” and “Have you ever attended an interview workshop.”

Topic Choice. Topics were extracted from the interview transcripts using LDA. LDA extracts topics from a corpus of text and calculates individual’s topic usage probability, providing one probability per topic for each text processed (Blei, Ng & Jordan, 2003). Deriving topics from unstructured text using LDA is accomplished with four distinct steps: data wrangling, pre-processing, dataset generation, and topic extraction (Landers, 2017). First, data wrangling involves the cleaning of text data and the generating of a corpus in which to employ NLP. To do this, text from the transcript of each interview question was added to a dataset and keyed with a participant identifier. Second, pre-processing alters the text in the corpus so that it contains only linguistically and analytically meaningful units. To do this, all words were changed to lower case, words were stemmed, all numbers and punctuation were removed, and all functional words were removed. Third, dataset generation converts the cleaned text into a dataset that can be analyzed; in other words, individual words or phrases are assigned numerical identifiers, essentially coding a word or phrase so that it can be used in LDA. Thus, a document-term matrix (dtm) was created, which became the input for the LDA model. Fourth, using the LDA function in the R *topicmodels* package, an LDA model was developed using the dtm.

To use the LDA function, the number of topics (K) had to first be specified, which is a main challenge of using LDA. To find the number of topics, the package *ldatuning* (Nikita, 2016) was used, which estimates the best fitting number of topics to the dataset by visualizing four different approaches to selecting the number of topics (Griffiths & Steyvers, 2004; Cao, Xia, Li, Zhang & Tang, 2009; Arun, Suresh, Madhavan & Murthy, 2010; Deveaund, SanJuan & Bellot, 2014). This process is conceptually similar to the use of a scree plot when determining the appropriate number of factors in an exploratory factor analysis in that it requires some degree of interpretation. Where scree plot interpretation generally requires the visual identification of an “elbow,” these techniques require visual identification of the minimum or maximum of a parabola. Broadly, these tuning approaches split documents in half to calculate the number of topics and then compare those results with calculated topics in the second half of the document. More specifically, the metrics developed by Arun et al. (2010) identify the point at which the difference between the two document samples is smallest in terms of Kullback-Liebler divergence scores, which represent the degree of similarity between word probabilities for each topic and the distribution of topics in the documents. The Cao et al. (2009) metric identifies the point at which the difference between the two document samples is smallest in terms of the average cosine similarity between each of the topic distributions. The Griffiths et al. (2004) metric maximizes the mean of the log-likelihoods of the data given the number of topics and thus maximizes the distances between all topic pairs. Deveaud et al.'s (2014) metric identifies the maximum mean distance, calculated using the Jensen-Shannon distance, between the topic distribution pairs. Thus, when interpreting the metrics, the goal is to identify the number of topics at the point of convergence between minimum values of metrics by Arun et al. (2010) and Cao et al. (2009) and maximum values of metrics by Griffiths et al. (2004) and Deveaud et al. (2014).

Initially, a range of 2 through 100 topics was examined using *ldatuning*. The optimal number of topics was decided on by the majority agreement of the metrics, an approach that has been used in other studies that have used the ldatuning package (e, g., El Mezouar, Zhang, and Zou, 2017; Zahedi, Babar & Treude, 2018). The Deveaud et al. (2014) metric yielded inconsistent results from the majority of the metrics and was therefore excluded from the interpretation. The Arun et al. (2010), Cao et al. (2009), and Griffiths et al. (2004) metrics appeared to converge between 2 and 32 topics. This range was then reduced to examine the ideal number of topics between 2 and 32 at a finer level of specificity. Figure 1 shows this analysis from which it was concluded that about 22 topics would be appropriate.

Next, LDA was used to model each document (i.e., interview transcript) as a mixture of 22 topics. Table 2 shows the top words associated with each topic. Using the LDA model, a probability matrix was generated that estimates the proportion of words in the document that are associated with each topic (Blei, Ng & Jordan, 2003), yielding a probability for each participant on each topic. These topic probabilities were used as predictors in the tests of H2 and H3.

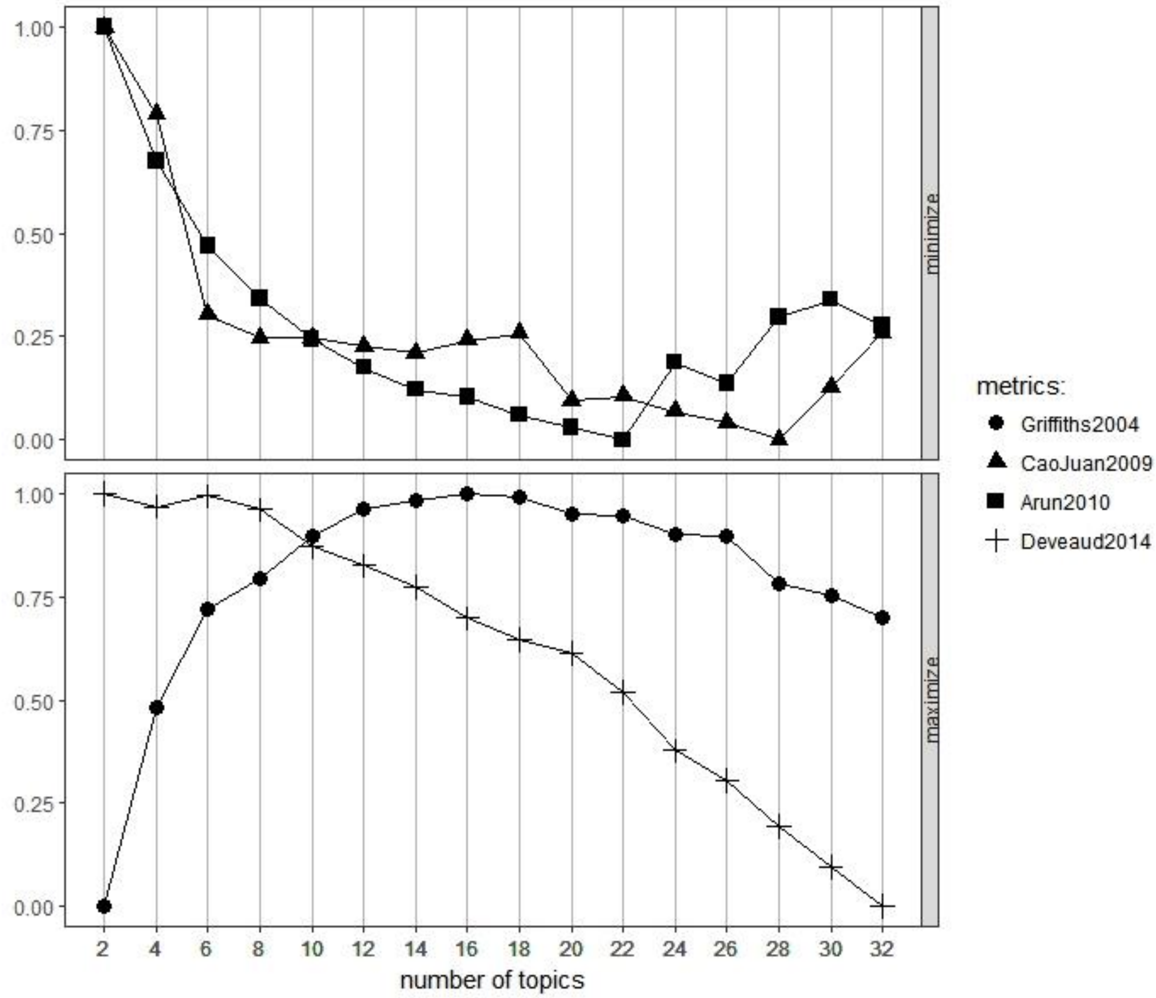

Figure 1.

Table 2

Top 10 Words Associated with each LDA Topic

Topic 1	know	peopl	can	get	make	one	see	thing	well	think
Topic 2	team	see	idea	work	get	new	thing	can	learn	communic
Topic 3	tri	new	one	way	get	learn	chang	make	peopl	thing
Topic 4	can	peopl	think	new	idea	know	way	tri	get	kind
Topic 5	know	work	someth	think	idea	tri	way	realli	team	learn
Topic 6	make	idea	get	new	sure	tri	know	team	alway	learn
Topic 7	make	peopl	know	thing	work	see	sure	get	good	can
Topic 8	idea	tri	team	work	everyon	make	way	new	peopl	sure
Topic 9	know	tri	kind	definit	new	thing	get	task	differ	idea
Topic 10	make	work	sure	compani	idea	may	abl	everybodi	might	new
Topic 11	know	learn	differ	can	will	also	come	member	someth	team
Topic 12	team	thing	peopl	get	work	new	way	learn	tri	idea
Topic 13	know	way	peopl	can	get	task	work	idea	everyon	alway
Topic 14	idea	want	know	see	work	can	alway	way	get	peopl
Topic 15	know	way	thing	can	idea	tri	everybodi	work	good	think
Topic 16	think	know	import	peopl	alway	idea	make	can	kind	someth
Topic 17	tri	can	know	way	new	person	one	team	learn	order
Topic 18	know	peopl	idea	work	thing	team	see	use	learn	everybodi
Topic 19	know	get	everybodi	peopl	idea	use	actual	time	tri	better
Topic 20	will	tri	idea	know	get	can	make	peopl	one	see
Topic 21	know	think	get	basic	peopl	can	come	job	place	communic
Topic 22	can	work	peopl	make	one	get	want	new	feel	come

Deceptive Impression Management-related LIWC Dictionaries. The Linguistic Inquiry and Word Count (LIWC) program, which is a set of text analysis algorithms that count words using pre-defined categories of language, was used to calculate word usage for the closed-vocabulary approach. The program was originally developed to measure health through participant's writing after difficulty obtaining rater agreement on writing samples (Pennebaker & Beall, 1986). LIWC consists of two features: processing and dictionaries (Pennebaker, Booth, Boyd & Francis, 2015). Processing is the identification of each word and comparison of each word to each of the dictionaries. Dictionaries are the actual collection of words that define the 90 output variables (i.e., first-person pronouns, authenticity, health; Pennebaker, Booth, Boyd & Francis, 2015). The LIWC2015 dictionaries are composed of about 6,400 words (Pennebaker, Booth, Boyd & Francis, 2015). The dictionaries were developed using a 7-step process, where developers collected relevant words, had judges assess fit of each word in the categories, analyzed base rates, and evaluated the psychometric properties of each category. Corrected internal consistency for each category to be used ranged from .55 to .88 (Pennebaker, Booth, Boyd & Francis, 2015): achievement dictionary was .81, leisure was .86, family was .88, pronoun usage ranges ranged from .61 to .84, positive emotion was .64, and negative emotion was .55. Reliability for word count and the authentic dictionary were not reported. LIWC2015, on average, captures about 86% of word use based on a base rate analysis of text from blogs, expressive writing by participants, novels, natural speech, the New York Times, and Twitter. In the present study, raw text from the interviews was analyzed using the LIWC program, which automatically processes the text and produces output. The output is expressed as a percentage of total words, with exception to word count and the summary variables (Analytic, Clout, Authentic, and Tone; (Pennebaker, Booth, Boyd & Francis, 2015). The summary variables are standardized composites that have been converted to percentiles.

Procedure

Participants were asked to first record themselves in an asynchronous mock-interview and take a post-interview survey. All interviews were recorded on HireVue. In a survey on Qualtrics, the participants were asked to sign a digital consent form. Similar to procedures used by previous deceptive IM studies, participants were instructed that the purpose of the study was to examine different behaviors that are used to impress interviewers and increase interview scores and that this was an opportunity for them to prepare themselves for a real future employment opportunity (cf. Levashina & Campion, 2007). The participants were also told that a researcher would be scoring their performance at the end of the interview. The participants were redirected to the asynchronous interview software and were asked a standard series of five interview questions (Appendix A) that were not job or organization specific. Interview questions were chosen from interview questions used in similar studies (Levashina & Campion, 2007) that were believed likely to elicit behaviors consistent with deceptive IM. The candidates were given 30 seconds to prepare their answer and three minutes to respond (Chen, Zhao, Leong, Lehman, Feng & Hoque, 2017). Each participant then completed an online survey. Interview videos were automatically transcribed into text using the machine transcription services used by HireVue.

CHAPTER 3

RESULTS

Data Cleaning and Assumption Checks

Prior to conducting analyses, the dataset was cleaned and variables were checked for missing data, outliers, and normal distribution of the dependent variable. There was no missing data. To check for influential outliers, a Cook's distance plot was created for each regression model indicating any observations in which Cook's distance was greater than one (Cook, 2000). These plots did not indicate any influential outliers. Overall deceptive IM was slightly skewed and platykurtic (skewness = 0.8657, kurtosis = 2.782). However, deceptive IM did not exceed the generally recommended thresholds of non-normality (e.g., Bulmer, 1979).

The data were checked for four assumptions of linear regression, including 1) linearity of residuals, 2) normal distribution of residuals, 3) equal variance of residuals, 4) and lack of multicollinearity among predictors (Cohen, Cohen, West, and Aiken, 2003). First, linearity between predictor and outcome variables was examined plotting the residuals against deceptive IM. The relationships between all independent variables and dependent variables appeared linear. Second, the residual errors were checked for normality using a Q-Q plot and were found to be normally distributed. Third, all models were additionally checked for heteroscedasticity using the Breusch-Pagan test (Breusch & Pagan, 1979). Heteroscedasticity was present in the model used to test H2 and H3, where topic scores were regressed onto deceptive IM scores. To account for this heteroscedasticity, a Box-Cox transformation was used to transform deceptive IM scores in those models (Box & Cox, 1964). Fourth, multicollinearity in the multivariate models was checked by calculating variance inflation factors for each independent variable and checking for values above 10 (O'Brien, 2007). No models demonstrated problematically high multicollinearity. Descriptive statistics and a correlation matrix of all variables are presented in Table 3.

Table 3

Descriptive Statistics and Pearson Correlations Between All Study Variables

	M	SD	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1 Deceptive IM	1.51	.48																																			
2 Word Count	637.37	370.33	.07																																		
3 1st Person Singular	5.92	2.18	.00	-.26 *																																	
4 1st Person Plural	.82	.68	.19 *	-.03	.01																																
5 2rd Person	2.22	1.85	-.03	.17	-.47 *	.00																															
6 3rd Person Singular	.30	.32	.00	.06	-.14	.03	.21 *																														
7 3rd Person Plural	1.49	.78	-.09	.08	.01	-.14	-.14	-.06																													
8 Impersonal Pronouns	7.99	1.78	-.14	.23 *	-.21 *	-.03	.05	-.13	-.02																												
9 Positive Emotion	2.94	1.05	-.07	.03	-.08	-.02	.01	-.03	-.10	.08																											
10 Negative Emotion	.72	.47	.01	-.04	-.02	.00	-.10	.05	.14	-.16	.01																										
11 Family	.08	.14	.07	.04	-.04	.15	.08	-.07	-.02	-.03	-.09	.05																									
12 Achievement	3.36	1.26	-.05	-.05	.00	-.12	-.25 *	-.09	.00	-.10	.20 *	-.06	-.17 *																								
13 Leisure	.76	.48	.00	-.25 *	.07	.00	.01	-.03	.07	-.18 *	-.06	-.10	-.01	.34 *																							
14 Authentic	53.19	16.54	-.03	.15	.29 *	-.09	.01	-.04	-.10	.09	-.07	-.19	.04	-.11	-.18 *																						
15 LDA Topic 1	.01	.12	-.01	.14	-.03	.24 *	.00	-.06	-.15	-.02	-.02	.07	.22 *	-.11	-.07	.00																					
16 LDA Topic 2	.06	.21	.00	-.04	-.04	.03	.03	-.09	.06	.00	.08	-.01	-.05	-.01	.13	-.10	-.03																				
17 LDA Topic 3	.04	.19	.03	-.06	.04	.00	-.06	.00	-.07	.05	-.12	.01	-.09	-.12	-.10	.05	-.03	-.06																			
18 LDA Topic 4	.06	.23	.17 *	.06	-.12	-.06	.14	.10	.00	-.05	-.05	.01	-.08	-.18 *	-.10	.05	-.03	-.07	-.06																		
19 LDA Topic 5	.03	.17	-.08	.19 *	-.07	-.08	.08	-.04	-.01	.16	.11	.08	.10	.06	-.05	.03	-.02	-.03	-.04	-.05																	
20 LDA Topic 6	.06	.23	.13	-.06	.01	.04	-.12	.00	-.04	-.24 *	.01	.04	-.01	.08	.12	-.11	-.03	-.05	-.06	-.07	-.05																
21 LDA Topic 7	.04	.19	.01	.10	.07	.02	-.08	.03	-.01	.01	.16	-.02	-.05	-.03	-.05	-.05	-.03	-.02	-.05	-.05	-.04	-.06															
22 LDA Topic 8	.06	.22	.17 *	-.07	.07	-.12	-.15	-.20 *	.08	-.01	.07	.05	-.06	.23 *	.08	-.08	-.03	-.07	-.06	-.07	-.05	-.07	-.06														
23 LDA Topic 9	.05	.21	.02	-.12	.04	-.05	.06	.26	-.03	-.23 *	.09	-.02	-.03	-.06	-.11	-.01	-.03	-.06	-.05	-.06	-.05	-.07	-.05	-.06													
24 LDA Topic 10	.02	.12	-.02	.33 *	-.10	.01	.00	.09	.02	.15	.05	-.07	-.06	.10	-.12	-.13	-.02	-.04	-.03	-.03	-.03	-.04	-.03	-.04	-.03												
25 LDA Topic 11	.03	.15	-.10	.03	-.06	-.03	-.03	.11	-.01	.04	.03	-.08	-.10	-.01	.01	-.01	-.02	-.04	-.04	-.04	-.03	-.05	-.04	-.05	-.04	-.02											
26 LDA Topic 12	.10	.28	-.19 *	-.19 *	.08	.06	-.08	-.01	-.06	.01	-.18 *	-.15	-.05	.17 *	.22 *	.00	-.04	-.07	-.06	-.09	-.06	-.10	-.05	-.10	-.09	-.05	-.06										
27 LDA Topic 13	.04	.19	-.10	.06	-.13	-.15	.11	.00	-.04	.06	.10	-.01	.02	.00	-.04	-.07	-.03	-.05	-.05	-.06	-.04	-.04	-.05	-.05	-.05	-.03	-.04	-.08									
28 LDA Topic 14	.01	.12	.13	.21 *	.08	.01	.03	.01	-.11	.05	.01	-.04	.13	-.07	-.05	.12	-.01	-.03	-.03	-.03	-.02	-.03	-.03	-.03	-.03	-.02	-.02	-.04	-.03								
29 LDA Topic 15	.05	.20	-.09	.01	-.08	.06	.03	-.06	-.04	.19 *	.04	.01	-.01	-.07	-.02	.01	-.03	-.07	-.03	-.06	-.02	-.06	-.05	-.01	-.03	-.03	-.04	-.07	-.06	-.03							
30 LDA Topic 16	.04	.17	-.10	-.01	-.05	-.01	.16	.04	-.12	.07	.24 *	-.08	.02	-.03	-.04	.08	-.03	-.07	-.04	-.03	-.05	-.05	-.05	-.04	.01	-.04	.02	-.05	-.06	-.03	-.06						
31 LDA Topic 17	.04	.20	-.01	.03	-.03	-.13	-.07	.02	.05	.00	-.03	.00	.01	.17 *	-.04	.08	-.03	-.06	-.05	-.05	-.04	-.06	-.04	-.06	-.05	-.03	-.04	-.07	-.05	-.03	-.05	-.05					
32 LDA Topic 18	.09	.27	-.03	-.28 *	.19 *	-.03	-.14	-.09	.23 *	-.03	-.21 *	.15	.19 *	-.06	.01	-.01	-.04	-.09	-.07	-.06	-.06	-.07	-.07	-.09	-.08	-.04	-.05	-.12	-.05	-.04	-.08	-.04	-.07				
33 LDA Topic 19	.04	.19	.02	.05	.00	.02	-.03	.02	.06	.00	-.09	.01	-.03	-.05	-.11	.20 *	-.03	-.02	-.05	-.06	-.03	-.06	-.05	-.04	-.06	.03	-.04	-.06	-.05	-.03	-.06	-.06	-.02	-.08			
34 LDA Topic 20	.05	.21	.06	-.02	-.02	.08	-.01	.00	.00	.02	-.01	-.07	.01	.05	.04	.03	-.03	-.06	-.04	-.06	-.05	-.04	-.05	-.06	-.06	-.03	-.04	-.07	-.05	-.03	-.06	-.06	-.05	-.06	-.06		
35 LDA Topic 21	.03	.14	.20 *	.04	.21 *	.11	.30 *	-.06	-.12	-.08	-.10	-.10	.04	-.16	.16	.13	-.02	.00	-.04	-.05	-.04	-.04	-.04	.03	-.05	-.03	-.03	-.03	-.04	-.02	-.05	-.05	-.04	-.06	-.05	-.03	
36 LDA Topic 22	.04	.19	-.14	.14	-.14	.08	.06	-.01	.10	.03	.02	.17	-.01	-.06	-.07	-.11	-.03	-.06	-.05	-.05	-.03	-.06	-.04	-.06	-.05	-.03	-.04	-.07	-.05	-.03	-.05	-.02	-.04	-.07	-.05	-.05	-.04

Note. *Correlation is significant at the 0.05 level (2-tailed).

The final dataset included participant reported composite deceptive IM scores, the percentage of use for each LIWC dictionary category (closed-approach), and the probability of topic use for each LDA topic. The linguistic features were then used as predictors in the regression models. Ordinary least squares regression was appropriate for addressing the proposed hypotheses because it assesses the extent to which a linear relationship between each predictor (i.e., linguistic features) and the outcome (i.e., deceptive IM) exists (Aiken, West & Reno, 1991).

Hypothesis Tests

To test H1 a-f, simple and multiple ordinary-least-squares regression were used to determine if each predictor or set of predictors' relationships with deceptive IM were statistically significant. Hypotheses 1 a-f were not supported, as shown in table 4.

Table 4

Regression results for Hypothesis 1 predicting Deceptive IM with Closed-Vocabulary Linguistic Features

Hypothesis	Predictor	*t*	*p*	*β*	F	*df*	*p*	R^2	adj R^2
1a	WC	0.89	.377	0.08				0.00	
1b	Negative Emotion Words	0.11	.911	0.01				0.01	
1c	Positive Emotion Words	-0.87	.387	-0.07				0.01	
1d	Overall				1.05	4,136	.383	0.03	0.00
	2rd Person Pronouns	-0.40	0.69	-0.04					
	3rd Person Singular Pronouns	-0.18	0.85	-0.02					
	3rd Person Plural Pronouns	-1.16	0.25	-0.10					
	Impersonal Pronouns	-1.67	0.10	-0.14					
1e	Overall				2.64	2,138	.075	0.04	0.02
	1st Person Singular Pronouns	-0.04	0.97	0.00					
	1st Person Plural Pronouns	2.30	0.02	0.19					
1f	Overall				0.25	4,136	.911	0.01	-0.02
	Authentic	-0.37	.715	-.032					
	Family	0.69	.494	.059					
	Leisure	0.08	.937	.007					
	Achievement	-0.52	.606	-.048					

To test H2, self-reported deceptive IM was regressed onto all topic probabilities determined by the LDA, but only the overall model R^2 was interpreted. Because topic probabilities sum to one within each case (i.e., there is a 100% chance that each case is assigned to one of the 22 topics extracted), the final topic (i.e., Topic 22) was chosen as a reference group. Thus, when all predictors equaled zero, the predicted value was the mean of Deceptive IM for cases assigned to Topic 22. Hypothesis 2 predicted that this system of interviewee LDA topic scores would predict deceptive IM. The data did not support this hypothesis $F(21,119) = 1.18$, $p = .287$, $R^2 = .173$ (Table 5).

Table 5
Regression results for Hypothesis 2 predicting Deceptive IM with LDA Topic Scores

	t	*p*	*β*	F	*df*	*p*	R^2	*adj* R^2
Overall Model				1.18	21, 119	.278	0.173	0.027
Topic 1	0.71	0.476	0.07					
Topic 2	1.08	0.283	0.13					
Topic 3	1.35	0.181	0.15					
Topic 4	2.17	0.032	0.27					
Topic 5	0.32	0.750	0.03					
Topic 6	1.90	0.060	0.24					
Topic 7	1.11	0.270	0.13					
Topic 8	2.07	0.041	0.25					
Topic 9	1.31	0.193	0.16					
Topic 10	0.42	0.676	0.04					
Topic 11	-0.03	0.976	0.00					
Topic 12	0.08	0.939	0.01					
Topic 13	0.28	0.778	0.03					
Topic 14	2.13	0.035	0.21					
Topic 15	0.41	0.685	0.05					
Topic 16	0.39	0.697	0.04					
Topic 17	1.33	0.185	0.15					
Topic 18	1.14	0.259	0.16					
Topic 19	1.28	0.202	0.15					
Topic 20	1.24	0.216	0.15					
Topic 21	2.39	0.018	0.24					

Note. Deceptive IM is transformed using a Box-Cox Transformation (Lambda = -1). Topic 22 is the reference group.

To test H3a-c, two hierarchical multiple linear regression models were created. In each, self-reported deceptive IM was regressed onto both the H1 and H2 predictor features in two steps, in alternate orders, to determine incremental explanation of variance of each over the other by examining the F-ratio associated with the change in R^2 between the two models. Hypothesis 3a predicted that in combination, an LDA- and LIWC-based approach would predict more variance in self-reported deceptive IM than LDA topics alone. The data did not support this hypothesis $F(13,106) = 1.011$, $p = .446$, $\Delta R^2 = .091$ (Table 6). Specifically, LIWC linguistic features did not significantly add to the amount of explained variance above and beyond LDA linguistic features. Hypothesis 3b predicted that in combination, an LDA- and LIWC-based approach would predict more variance in self-reported deceptive IM than LIWC-based linguistic features alone. The data did not support this hypothesis $F(21,106) = 1.215$, $p = .254$, $\Delta R^2 = .177$ (Table 7). Although LDA topic scores did explain more variance above and beyond the LIWC linguistic features, this change was not statistically significant. Additionally, although the R^2 effect size appears impressive, this value may be inflated by the large number of predictors. Thus, a more conservative estimate of the variance explained by deceptive IM is found in the adjusted R^2 values, which are all near zero.

Table 6

Hierarchical Regression results for Hypothesis 3a

Step			β	p	F	R^2	adj R^2	ΔR^2
1	Overall				1.183	0.173	0.027	0.173
		Topic 1	0.07	.476				
		Topic 2	0.13	.283				
		Topic 3	0.15	.181				
		Topic 4	0.27	.032				
		Topic 5	0.03	.750				
		Topic 6	0.24	.060				
		Topic 7	0.13	.270				
		Topic 8	0.25	.041				
		Topic 9	0.16	.193				
		Topic 10	0.04	.676				
		Topic 11	0.00	.976				
		Topic 12	0.01	.939				
		Topic 13	0.03	.778				
		Topic 14	0.21	.035				
		Topic 15	0.05	.685				
		Topic 16	0.04	.697				
		Topic 17	0.15	.185				
		Topic 18	0.16	.259				
		Topic 19	0.15	.202				

Table 6 Continued

		Topic 20	0.15	.216				
		Topic 21	0.24	.018				
2	Overall				1.118	0.264	0.028	0.091
		Topic 1	0.00	.980				
		Topic 2	0.16	.186				
		Topic 3	0.19	.120				
		Topic 4	0.31	.018				
		Topic 5	0.10	.391				
		Topic 6	0.24	.078				
		Topic 7	0.15	.195				
		Topic 8	0.34	.011				
		Topic 9	0.22	.091				
		Topic 10	0.04	.701				
		Topic 11	0.02	.877				
		Topic 12	0.07	.642				
		Topic 13	0.10	.398				
		Topic 14	0.22	.036				
		Topic 15	0.08	.522				
		Topic 16	0.11	.350				
		Topic 17	0.23	.070				
		Topic 18	0.26	.104				
		Topic 19	0.18	.134				

Table 6 Continued

Topic 20	0.18	.149
Topic 21	0.30	.016
Word Count	0.12	.301
2rd Person Pronouns	-0.20	.093
3rd Person Singular Pronouns	-0.01	.939
3rd Person Plural Pronouns	-0.09	.376
Impersonal Pronouns	-0.10	.325
1st Person Singular Pronouns	-0.18	.150
1st Person Plural Pronouns	0.21	.027
Positive Emotion	-0.04	.690
Negative Emotion	0.00	.972
Family	0.01	.933
Achievement	-0.11	.313
Leisure	0.07	.534
Authentic	-0.04	.727

Note. Deceptive IM is transformed using a Box-Cox Transformation (Lambda = -1). Topic 22 is the reference group.

Table 7

Hierarchical Regression results for Hypothesis 3b

Step		β	p	F	R^2	adj R^2	ΔR^2
1	Overall			0.929	0.087	-0.007	0.087
	Word Count	0.12	.206				
	2rd Person Pronouns	-0.12	.255				
	3rd Person Singular Pronouns	-0.04	.682				
	3rd Person Plural Pronouns	-0.08	.387				
	Impersonal Pronouns	-0.18	.058				
	1st Person Singular Pronouns	-0.08	.497				
	1st Person Plural Pronouns	0.16	.079				
	Positive Emotion	-0.04	.616				
	Negative Emotion	0.00	.992				
	Family	0.01	.927				
	Achievement	-0.13	.210				
	Leisure	0.06	.530				
	Authentic	-0.01	.952				
2	Overall			1.118	0.264	0.028	0.117
	Word Count	0.12	.301				
	2rd Person Pronouns	-0.20	.093				
	3rd Person Singular Pronouns	-0.01	.939				
	3rd Person Plural Pronouns	-0.09	.376				
	Impersonal Pronouns	-0.10	.325				
	1st Person Singular Pronouns	-0.18	.150				
	1st Person Plural Pronouns	0.21	.027				
	Positive Emotion	-0.04	.690				
	Negative Emotion	0.00	.972				
	Family	0.01	.933				
	Achievement	-0.11	.313				

Table 7 Continued

Leisure	0.07	.534
Authentic	-0.04	.727
Topic 1	0.00	.980
Topic 2	0.16	.186
Topic 3	0.19	.120
Topic 4	0.31	.018
Topic 5	0.10	.391
Topic 6	0.24	.078
Topic 7	0.15	.195
Topic 8	0.34	.011
Topic 9	0.22	.091
Topic 10	0.04	.701
Topic 11	0.02	.877
Topic 12	0.07	.642
Topic 13	0.10	.398
Topic 14	0.22	.036
Topic 15	0.08	.522
Topic 16	0.11	.350
Topic 17	0.23	.070
Topic 18	0.26	.104
Topic 19	0.18	.134
Topic 20	0.18	.149
Topic 21	0.30	.016

Note. Deceptive IM is transformed using a Box-Cox Transformation (Lambda = -1). Topic 22 is the reference group.

Post-hoc and Exploratory Analyses

Post-hoc exploratory analyses were conducted to further assess how sample characteristics may have influenced the findings. These post-hoc analyses considered both data source (i.e., psychology participant pool vs. general campus recruitment) and interview experience. Participants were categorized as “high” in interview experience if they reported participating in more than four interviews, the sample median number of interviews, in their lifetime. Participants were categorized as “low” in interview experience if they reported participating in fewer than four interviews in their lifetime. In general, mean deceptive IM behaviors did not differ across these samples (Psychology pool: M = 1.51, SD = .49, n = 115; Campus recruited: M = 1.51, SD = .46, n = 26; High interview experience: M = 1.49, SD = .48, n = 56; Low interview experience: M = 1.52, SD = .47, n = 59). Additionally, the suggested number of topics extracted for each group remained relatively the same, ranging from 16-20 topics. Lastly, each of the hypothesis tests was replicated within each of the four specific samples. The results did not change when these tests were replicated in either the psychology participant pool or campus recruited sample. The results also did not change when the hypotheses were tested using the “low” interview experience sample. Some results did change, however, when the hypotheses were tested using the “high” interview experience sample. Specifically, word count (H1a) significantly predicted deceptive IM in this population $\beta = .29$, $t(55) = 2.26$, $p = .028$, $R^2 = 0.09$. Additionally, the hypothesized LIWC topics (H1f) significantly predicted deceptive IM $F(4,55) = 2.56$, $p = .049$, $R^2 = 0.17$, *adj* $R^2 = 0.102$. These findings suggest that some relationships between linguistic features and deceptive IM may be stronger when participants are more experienced with interviews.

CHAPTER 4

DISCUSSION

This study examined the use of open and closed-vocabulary natural language processing approaches for the detection of deceptive IM in mock employment interviews. In general, neither of these approaches successfully predicted deceptive IM. Using a closed-vocabulary approach, several LIWC dictionaries were hypothesized to predict deceptive IM including word count, positive and negative emotion words, pronoun usage, and certain topics (leisure, family, achievement, and authenticity). However, these dictionaries did not significantly predict deceptive IM. Using an open-vocabulary approach, LDA topic probability scores were expected to significantly predict deceptive IM. Overall, however, LDA topics from the open-vocabulary approach did not significantly predict deceptive IM. Additionally, each of the two approaches was expected to explain incremental variance in deceptive IM over each other. However, neither approach significantly predicted more variance in deceptive IM usage than the other.

These findings build on existing deceptive impression management theory from both a deductive and inductive perspective. Deceptive IM theory posits that people alter what they say by embellishing and tailoring answers, constructing answers, omitting or masking information, and by gaining favor with the interviewer, and this study sought to identify more specific linguistic cues to deceptive IM. By taking a closed-vocabulary approach to linguistic analysis, the current study assessed how well existing theory on the relationship between deception and linguistic cues extended in the context of deceptive IM behaviors in mock employment interviews. This approach, in the context of this study, was a deductive mechanism for expanding deceptive IM theory. However, these relationships were generally not supported, suggesting that the relationship between deception and linguistic cues may not replicate in the context of deceptive IM in mock employment interviews. By taking an open-vocabulary approach to linguistic analysis, the current study sought to identify new sets of linguistic features that may be associated with deceptive IM behaviors. The open-vocabulary approach, in this context, was an inductive mechanism for identifying relevant linguistic features that predict deceptive IM. However, because topic usage did not significantly predict deceptive IM, these findings do not support that the approach used to extract topics from mock employment interview text yield topics theoretically relevant to deceptive IM behaviors. Ultimately, neither the deductive or inductive approach identified specific linguistic features that are relevant to deceptive IM behaviors.

In addition to the lack of predictive power of the open-vocabulary approach, the topics extracted using the sample in this study lacked obvious interpretability, further limiting the usefulness of LDA in this context. Words that were relevant to each topic occurred in multiple topics and there was no intuitively apparent way to interpret or label each of the twenty-two topics. This could potentially be due to the specific method used in this study to extract topics, given that there are many ways to do so. For example, there is an infinite number of topics that could be extracted and it is possible that extracting a different number would have yielded more interpretable topics, similar to the challenge faced when interpreting an exploratory factor analysis. Additionally, it is possible that particularly high-frequency words used by interviewees made it difficult for the LDA model to identify those words' topics accurately. Overall, given the methods used in the present study, topic modeling was not useful for predicting deceptive IM or yielding interpretable topics. This finding is relevant when considering the usefulness of LDA as both a predictive and interpretation tool in this context. LDA was used in this study for both its predictive power and as a means for identifying relevant topics that were related to deceptive IM use. However, in the context of this study and the approach taken to extracting topics, LDA did not accomplish either of those goals.

Regardless, the findings of this study also contribute to NLP theory. First, the third set of hypotheses compare open and closed-vocabulary NLP approaches, more specifically comparing a dictionary-based approach with topic modeling. In the field of computer science, theory testing is framed as algorithmic performance testing. In other words, the goal is to understand the utility and predictive success of an algorithm. In contrast to previous findings that have shown differences in the information extracted from open and closed vocabulary approaches as well as the predictive power of each approach (Iacobelli, Gill, Nowson, and Oberlander, 2011; Schwartz et al., 2013), the findings of the current study did not show that one approach differed significantly from the other when predicting deceptive IM.

There are two broad conclusions that could be drawn from these findings. First, people may not change the number of words they use, their pronoun usage, the tone of their words, their use of words related to leisure, family, achievement, or authenticity, or the general topics they discuss in an employment interview when engaging in deceptive IM. Although impression management theory posits that people change what they say to influence others, it is possible that the specific linguistic features chosen in this study are not the linguistic features altered by those engaging in deceptive impression management. This reasoning could apply to the specific subset of linguistic features in each approach (i.e., the specific LIWC dictionaries and the specific topics extracted from the text) or more broadly to the approaches themselves, meaning that word counts based on pre-specified dictionaries (closed-vocabulary) or probability scores of topic usage from LDA do not predict self-reported deceptive IM. However, in addition to the previous theoretical research suggesting that these relationships should exist, there are several alternative reasons for why the hypotheses were not supported as will be detailed later. Thus, the conclusion that these relationships do not exist should not be heavily considered without additional research. For example, another way to test this conclusion would be by conducting a within-subject study in which participants were asked to complete an interview both honestly and deceptively. It would be possible to identify how these linguistic indicators differentially predict cases of high and low deceptive IM. If people do not change the specific linguistic features between such directed faking conditions, there should be no difference, which would further support this conclusion.

The second potential conclusion is that despite a relationship between the linguistic features measured in this study and deceptive IM, the effects were not detectable given the present research design or execution. This conclusion was partially supported given the results of the post-hoc analyses, which indicated that some of these relationships, specifically word count and LIWC topics, exist when the sample is relatively high in interview experience. Broadly, there are three contributing factors to this second conclusion. First, given that the sample size obtained did not meet the a priori power analysis requirement, it is possible that there was not enough statistical power to detect the effects. The sampling approach in this study allowed for consistency in interview administration and the collection of self-reported deceptive IM scores. However, this approach limited the amount of data that could be collected in a reasonable amount of time. In most NLP research, sample size is typically quite large so that there are many more participants than linguistic features, making small effect sizes are more detectable (e.g., Pennebaker, Chung, Frazee, Lavergne, and Beaver, 2014; Kulkarni et al., 2017). As a result, although the sampling approach in this study was necessary for testing the hypotheses, the hypothesized relationships may have required additional statistical power.

However, the low observed effect sizes suggest that even the a priori power analysis was targeting larger effects than may exist in the population for some hypotheses. Second, there was a low base rate of overall deceptive IM behaviors, meaning there was little variance in the dependent variable for modeling. Overall, participants did not engage in many deceptive IM behaviors and were consistent in this effort, making it more difficult to predict the behaviors due to range restriction. Therefore, if a sample in which participants varied greatly in the extent to which they engaged in deceptive IM behaviors, and engaged in a higher base rate in general, some of the hypotheses may have been supported. Third, because topics were extracted from the entire interview transcript, it is possible that the topics extracted from the interview text to some extent represent different topics associated with each interview question. However, extracting topics for each interview question would have substantially increased the overall number of topics, and therefore predictors, in the model, exacerbating the power problem further. This approach would have been appropriate given a larger sample size. Therefore, the topic modeling approach that was used may have extracted topics that were too broad to predict subtle topic differences in deceptive IM.

Limitations

There are two primary limitations in this study: the sample characteristics and the limitations surrounding the NLP techniques that were used. First, a small student sample in a mock interview setting was used. Because a student sample was used, most participants had limited work and interview experience and may not be representative of a typical applicant sample. Additionally, although approximately 90% of participants in this study did report at least one deceptive impression management behavior, there may be differences in how and the extent to which a sample of real job applicants engage in deceptive impression management behaviors.

Participants completed a mock interview in a lab setting and were therefore not competing for an actual job. Differences in motivation may lead to different amounts and types of interview faking. For example, Levashina and Campion (2007) found slight differences in base rates of deceptive impression management behaviors in actual interview samples and practice mock interview samples, which were conducted as part of an undergraduate class for a grade, such that there were generally higher rates of faking in the actual interview sample in contrast to the mock interview sample. In the present study, the average base rate of deceptive IM behaviors was relatively low (M = 1.51, SD = 0.48). This base rate in the present study was significantly lower, based on an independent-samples t-test, than any of the base rates reported by Levashina and Campion (2007), including the undergraduate practice mock interview sample (M = 1.73, SD = 0.61; t(290)=3.41, $p < .001$). However, steps were taken to ensure that the participants were taking the mock interview seriously. Participants who did not indicate agree or strongly agree to the statement “I took the mock interview as seriously as I would normally take a real interview” were removed from the study.

Although a limitation of this sampling approach is the potential lack of motivation to fake, it was important for participants to honestly report their deceptive impression management behaviors. While there may have been some degree of socially desirable responding to the deceptive impression management scale due to demand characteristics of the study, I contend that the low stakes and research-focused nature of the interview setting makes it more likely that participants would be willing to report these behaviors in contrast to an actual high stakes interview setting. In a realistic interview setting, it is probable that applicants would not want to disclose their faking behaviors to avoid harming their chances of obtaining employment. Thus, while there would likely be more prevalent deceptive impression management behaviors in a realistic applicant setting, self-reported deceptive IM behaviors would likely be downwardly biased. Ideally, the present study would be conducted using real applicants and self-reported deceptive IM behaviors would be collected anonymously for research purposes. In this scenario, there would likely be higher base rates in deceptive IM and truthful reporting of faking behaviors, making these effects more easily detectable. In comparison to such a study, the present study was a conservative test of whether deceptive impression management could be detected using NLP in an interview context.

Second, the natural language processing techniques used in this study are coarse techniques, meaning they are not nuanced enough to pick up on many features of language as understood by a human. Thus, although the study does not support the use of these coarse approaches to NLP, it does not condemn all NLP approaches. Both LIWC and LDA are bag of words approaches, meaning they only consider word counts and not the order of words or other linguistic features. Although topic modeling is a step towards semantic understanding of text, true natural language understanding cannot be achieved without modeling context (Cambria & White, 2014). Some NLP techniques have begun to explore a finer grain approach to NLP in hopes to achieve better natural language understanding (Cambria & White, 2014), but this work is in early stages. Although LIWC and LDA are both potentially useful forms of measurement, they are far from a human-level understanding of text. Lastly, based upon my own review, the machine transcription of the interviews in this study, which is a form of natural language processing, was not as accurate as human transcription would have been. There were some words that were incorrectly interpreted by the computer, which has the potential to undermine the conclusions drawn. However, the machine transcription services used in this study are the same quality as the services used in practice and therefore create the most realistic text dataset. It would likely not be practical for organizations to use human transcription services for all their asynchronous interviews because the practical benefit of automating interview analysis would be lost. However, as machine transcription services improve, so will the accuracy of these types of analyses.

Practical Contributions

Practically, this study describes a prototypical approach for exploring NLP to detect deceptive IM in employment interviews but suggests that organizations should be careful in doing so given small, difficult-to-detect effects. Although currently unsupported, if future research with an alternative design found support for the hypotheses in this study, interview transcripts could be assessed using the techniques discussed in this study and used to help the interviewer detect deceptive IM. For example, when organizations screen applicants using a traditional interview, a transcript of that interview could be assessed using the techniques discussed in this study and used to help the interviewer detect deceptive IM. Additionally, as asynchronous interviews and algorithmic interview scoring become more commonplace (Chen et al 2016; Feloni, 2017), these methods could be used to flag interviewees who may be engaging in high amounts of deceptive IM. In both scenarios, interviewers or hiring decision makers would have the opportunity to take a second look at an applicant who might be scoring well in an interview but may be engaging in deceptive IM behaviors.

In either of these use cases, both closed- and open-vocabulary approaches could be used either alone or in combination. Both approaches would require a validation study in the context of the organizational sample. Given that actual applicants would be unlikely to report deceptive IM behaviors, participants in this validation study would likely need to be current employees. Using a concurrent validation design, employees would interview as if they were applying for a job in the organization and then be asked to self-report deceptive IM behaviors. To use the closed-vocabulary approach, relevant dictionaries would be identified, and proportion of word use on those dictionaries would be calculated for each current employee participant. Real applicants would be scored using the same dictionaries. To use the open-vocabulary approach, topics would be extracted from the transcript text and that LDA model would then be applied to the real applicant text. In other words, the LDA model developed during the concurrent validation study would be applied to the applicant sample, generating a series of posterior probabilities that the applicant sample used those same initial topics. The organization could then build a model to predict deceptive IM behaviors using the linguistic features extracted from the concurrent validation sample. This model would then be used to predict deceptive IM in applicant samples. However, given the lack of support for using NLP to detect deceptive IM in interviews, more research should be conducted prior to implementing these techniques in organizations.

Additionally, even if future research found support for the hypotheses in this study, there are several practical concerns that may hinder the implementation of using these NLP approaches to detect deceptive IM. First, given the small effect sizes detected in this study, any concurrent validation study would require a large sample size. This would require recruiting many current employees to participate and convincing stakeholders to buy into spending those resources. Such an approach also assumes that current employees and prospective employees would engage in the same type and degree of deceptive IM to enable detection, which may not be a safe assumption. Second, the NLP and machine learning techniques that would be necessary when building this type of system would require a degree of systems engineering and statistical programming knowledge and skill that most current I-O practitioners do not have.

Organizations may alternatively consider hiring a data scientist to create such a system. Some data scientists are familiar with NLP techniques, including more sophisticated NLP techniques, and often have a more data-driven perspective when building models. Given this approach, data scientists might be able to create a model with higher predictive power than the models in this study. However, it is important to keep in mind, especially in a selection context, that expertise is needed to ensure these predictive models are interpretable and fair. Thus, ideally, it would be most beneficial for data scientists and practitioners to work together to merge existing theory and data science techniques to build these types of systems. Organizations would also need to consider the financial return of such investments given the low demonstrated empirical support so far and thus relatively high degree of risk.

Future Directions

In addition to replicating this study to address the discussed limitations, there are numerous future directions that can build from the findings of the current study. The current study examined a small set of linguistic cues. Other linguistic cues (e.g., other NLP techniques) as well as non-linguistic cues (e.g., tone of voice, facial expressions, etc) may predict deceptive IM behaviors. For example, rather than focusing on LIWC dictionaries as predictors, a dictionary of words associated with deceptive IM could be crafted and used to predict deceptive IM. Additionally, in a large enough sample, all individual words or sets of words could be used as predictors of deceptive IM. However, given the large number of predictors in that scenario and the increased risk of overfitting models, it would likely be best to shift to modern predictive modeling methods, such as machine learning (Putka et al., 2017). Additionally, it would be useful to examine further the psychometric properties of using NLP to capture deceptive IM. For example, it would be interesting to explore how deceptive IM captured using the methodology in this study would relate to the predictive validity of constructs measured in interviews (e.g., personality, integrity, etc) or interview scores. It would also be interesting to measure convergence between using NLP to identify deceptive IM behaviors and third-party ratings of deceptive IM, similar to how Moore, Lee, Kim and Cable (2017) had raters assess the inauthenticity of candidates in an interview. Because of the extremely scant literature available on using NLP to measure psychological constructs, especially in the setting of an employment interview, the research possibilities are abundant.

Beyond the employment interview, future research is needed to understand how these NLP techniques and findings would replicate in other contexts. For example, in the employee selection and assessment context, NLP could be used to analyze resumes, cover letters, written responses to application questions, verbal or written responses to situational judgment test questions, and verbal or written responses in assessment centers. More broadly, it would be interesting to explore the differences in using these NLP techniques in written vs verbal language. The present study requires participants to verbally respond to a question without much time to prepare and without the opportunity to revise their answers. It is possible that people would differentially change their language when engaging in deceptive IM if they were writing out their answers instead or did not have the time pressure. Some previous research has supported this notion by suggesting differences in using NLP in spoken versus written language. For example, Mairesse et al. (2007) found that the personality traits extraversion, emotional stability, and conscientiousness were more easily predicted using spoken language than written language. In contrast, openness to experience was more easily predicted using written language. Thus, it is possible that deceptive IM may be more easily predicted using NLP if the response to the interview questions were written. Written responses may allow people to be more explicit in or conscious of their deception, resulting in more detectable changes in linguistic features. Future research should explore both the specific applications of NLP to different selection and assessment contexts as well as consider the broad patterns that occur when using NLP to analyze written and spoken language.

Conclusion

This study sought to understand how open and closed-vocabulary NLP approaches could be used to detect deceptive IM in mock employment interviews. The findings of this study suggest that these NLP approaches, both of which are word-count based, are weak predictors of deceptive IM and that more research is needed before using NLP to detect deceptive IM. Although it is possible that these approaches cannot be used in this context, there are other explanations for why these approaches did not work that should be ruled out first. Future research should address these alternatives by testing the hypotheses in a larger, more realistic applicant sample. Until then, however, organizations should proceed with caution when deciding to use NLP techniques to predict deceptive IM in employment interviews. Currently, self-report and observer assessment of deceptive IM appears to be the best approach to detecting deceptive IM in employment interviews. However, these NLP approaches may have promise in more large-scale, data-rich scenarios.

REFERENCES

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple regression: Testing and interpreting interactions.* Newbury Park, CA: Sage.

Arun, R., Suresh, V., Madhavan, C. E., & Murthy, M. N. (2010). On Finding the Natural Number of Topics with Latent Dirichlet Allocation: Some Observations. *Advances in Knowledge Discovery and Data Mining Lecture Notes in Computer Science*, 391-402. doi:10.1007/978-3-642-13657-3_43

Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what you get: relationships among self-presentation tactics and ratings of interview and job performance. *Journal of Applied Psychology, 94*(6), 1394-1411. doi:10.1037/a0016532

Becker, B., & Gerhart, B. (1996). The impact of human resource management on organizational performance: Progress and prospects. *Academy of Management Journal, 39*(4), 779-801. doi:10.2307/256712

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3,* 993-1022.

Bolino, M. C., & Turnley, W. H. (1999). Measuring impression management in organizations: A scale development based on the Jones and Pittman taxonomy. *Organizational Research Methods, 2*(2), 187-206. doi:10.1177/109442819922005

Bond, C. F., Jr., & DePaulo, B. M. (2006). Accuracy of deception judgments. *Personality and Social Psychology Review, 10,* 214-234.

Bowen, D. E., & Ostroff, C. (2004). Understanding HRM–firm performance linkages: The role of the “strength” of the HRM system. *Academy of Management Review, 29*(2), 203-221.

Box, G. E., & Cox, D. R. (1964). An Analysis of Transformations. *Journal of the Royal Statistical Society*. *Series B (Methodological), 26*(2), 211-252.

Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society, 47*, 1287-1294.

Buller, D. B., & Burgoon, J. K. (1996). Interpersonal Deception Theory. *Communication theory*, *3*, 203-242.

Bulmer, M. G. (1979). *Principles of statistics.* New York: Dover.

Burgoon, J. K., Blair, J. P., Qin, T., & Nunamaker, J. F. (2003). Detecting deception through linguistic analysis. *Intelligence and Security Informatics*, *2665*, 91–101

Cambria, E., & White, B. (2014). Jumping NLP curves: A review of natural language processing research. *IEEE Computational intelligence magazine, 9*, 48-57. doi: 10.1109/MCI.2014.2307227

Campion, M. C., Campion, M. A., Campion, E. D., & Reider, M. H. (2016). Initial investigation into computer scoring of candidate essays for personnel selection. *Journal of Applied Psychology, 101*, 958-975. doi:10.1037/apl0000108

Cao, J., Xia, T., Li, J., Zhang, Y., & Tang, S. (2009). A density-based method for adaptive LDA model selection. *Neurocomputing, 72*(7-9), 1775-1781. doi:10.1016/j.neucom.2008.06.011

Carlson, J. R., Carlson, D. S., & Ferguson, M. (2011). Deceptive impression management: Does deception pay in established workplace relationships? *Journal of Business Ethics*, *100*(3), 497-514. doi:10.1007/s10551-010-0693-9

Chapman, D. S., & Webster, J. (2003). The Use of Technologies in the Recruiting, Screening, and Selection Processes for Job Candidates. *International Journal of Selection and Assessment, 11*(2-3), 113-120. doi:10.1111/1468-2389.00234

Chen, L., Feng, G., Leong, C. W., Lehman, B., Martin-Raugh, M., Kell, H., ... & Yoon, S. Y. (2016). Automated scoring of interview videos using Doc2Vec multimodal feature extraction paradigm. *In Proceedings of the 18th ACM International Conference on Multimodal Interaction – ICMI 2016.* doi:10.1145/2993148.2993203

Chen, L., Zhao, R., Leong, C. W., Lehman, B., Feng, G., & Hoque, M. E. (2017). Automated video interview judgment on a large-sized corpus collected online. *Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*. doi:10.1109/acii.2017.8273646

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple correlation/regression analysis for the behavioral sciences.* Mahwah, NJ: Lawrence Erlbaum.

Cook, R. D. (2000). Detection of Influential Observation in Linear Regression. *Technometrics, 42*(1), 65.

Culbertson, S. S., Weyhrauch, W. S., & Waples, C. J. (2016). Behavioral Cues as Indicators of Deception in Structured Employment Interviews. *International Journal of Selection and Assessment, 24*(2), 119-131. doi:10.1111/ijsa.12135

DePaulo, B. M., Lindsay, J. J., Malone, B. E., Muhlenbruck, L., Charlton, K., & Cooper, H. (2003). Cues to deception. *Psychological bulletin, 129*, 74-118.

DePaulo, B. M., Stone, J. I., & Lassiter, G. D. (1985). Telling ingratiating lies: Effects of target sex and target attractiveness on verbal and nonverbal deceptive success. *Journal of Personality and Social Psychology, 48*(5), 1191. doi:10.1037//0022-3514.48.5.1191

Deveaud, R., SanJuan, E., & Bellot, P. (2014). Accurate and effective latent concept modeling for ad hoc information retrieval. *Document Numérique, 17*(1), 61-84. doi:10.3166/dn.17.1.61-84

Donovan, J. J., Dwight, S. A., & Schneider, D. (2014). The impact of applicant faking on selection measures, hiring decisions, and employee performance. *Journal of Business and Psychology, 29*(3), 479-493. doi:10.1007/s10869-013-9318-5

Ekman, P. (2001). *Telling lies: Clues to deceit in the marketplace, politics, and marriage* (3rd ed.). New York, NY: W.W. Norton.

Ellingson, J. E., Smith, D. B., & Sackett, P. R. (2001). Investigating the influence of social desirability on personality factor structure. *Journal of Applied Psychology, 86*(1), 122-133. doi:10.1037//0021-9010.86.1.122

Ellis, A. P., West, B. J., Ryan, A. M., & DeShon, R. P. (2002). The use of impression management tactics in structured interviews: A function of question type? *Journal of Applied Psychology, 87*(6), 1200-1208. doi:10.1037//0021-9010.87.6.1200

El Mezouar, M., Zhang, F., & Zou, Y. (2018). Are tweets useful in the bug fixing process? An empirical study on Firefox and Chrome. *Empirical Software Engineering, 23*(3), 1704-1742.

Feloni, R. (2017, June 28). Consumer-Goods Giant Unilever has been Hiring Employees Using Brain Games and Artificial Intelligence - and it's a huge success. *Business Insider*, downloaded June 29, 2017 from http://www.businessinsider.com/unilever-artificial-intelligence-hiring- process-2017- 6.

Gilmore, D. C., Stevens, C. K., Harrell-Cook, G., & Ferris, G. R. (1999). Impression management tactics. In R. W. Eder & M. M. Harris (Eds.), *The Employment Interview Handbook*. Thousand Oaks, CA: Sage. doi:10.4135/9781452205519.n18

Golbeck, J., Robles, C., & Turner, K. (2011, May). *Predicting personality with social media.* In Proceedings of the 2011 Annual Conference on Human Factors in Computing Systems, Vancouver, BC, 253–262.

Gombos, V. A. (2006). The Cognition of Deception: The Role of Executive Processes in Producing Lies. *Genetic, Social, and General Psychology Monographs*, *132*(3), 197-214. doi:10.3200/mono.132.3.197-214

Griffiths, T. L., & Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National academy of Sciences*, *101*, 5228-5235.

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*, 267-297. doi:10.1093/pan/mps028

Hall, J. A., & Pennington, N. (2013). Self-monitoring, honesty, and cue use on Facebook: The relationship with user extraversion and conscientiousness. *Computers in Human Behavior*, *29*(4), 1556-1564. doi:10.1016/j.chb.2013.01.001

Hauch, V., Blandón-Gitlin, I., Masip, J., & Sporer, S. L. (2015). Are Computers Effective Lie Detectors? A Meta-Analysis of Linguistic Cues to Deception. *Personality and Social Psychology Review*, *19*(4), 307-342. doi:10.1177/1088868314556539

He, Q., Glas, C. A., Kosinski, M., Stillwell, D. J., & Veldkamp, B. P. (2014). Predicting self-monitoring skills using textual posts on Facebook. *Computers in Human Behavior*, *33*, 69-78. doi:10.1016/j.chb.2013.12.026

Holoien, D. S., & Fiske, S. T. (2013). Downplaying positive impressions: Compensation between warmth and competence in impression management. *Journal of Experimental Social Psychology, 49*(1), 33-41. doi:10.1016/j.jesp.2012.09.001

Holtgraves, T. (2011). Text messaging, personality, and the social context. *Journal of Research in Personality, 45*, 92–99. doi:10.1016/j.jrp.2010.11.015

Huselid, M. A. (1995). The Impact of Human Resource Management Practices on Turnover, Productivity, and Corporate Financial Performance. *Academy of Management Journal, 38*, 635-672. doi:10.2307/256741

Iacobelli, F., Gill, A. J., Nowson, S., & Oberlander, J. (2011). Large Scale Personality Classification of Bloggers. *In Affective Computing and Intelligent Interaction Lecture Notes in Computer Science*, 568-577. Springer, Berlin, Heidelberg. doi:10.1007/978-3-642-24571-8_71

Ickes, W., Reidhead, S., & Patterson, M. (1986). Machiavellianism and self-monitoring: As different as “me” and “you”. *Social Cognition, 4,* 58-74.

Kern, M. L., Eichstaedt, J. C., Schwartz, H. A., Dziurzynski, L., Ungar, L. H., Stillwell, D. J., ... & Seligman, M. E. (2014). The Online Social Self: An open vocabulary approach to personality. *Assessment, 21*, 158-169. doi:10.1177/1073191113514104

Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences, 110*(15), 5802-5805.

Kulkarni, V., Kern, M. L., Stillwell, D., Kosinski, M., Matz, S., Ungar, L., ... & Schwartz, H. A. (2017). Latent Human Traits in the Language of Social Media: An Open-Vocabulary Approach.

Landers, R. N. (2017). Crash Course in I-O Psychology: A Crash Course in Natural Language Processing. *The Industrial-Organizational Psychologist, 4*.

Levashina, J., & Campion, M. A. (2006). A model of faking likelihood in the employment interview. *International Journal of Selection and Assessment, 14*(4), 299-316. doi:10.1111/j.1468-2389.2006.00353.x

Levashina, J., & Campion, M. A. (2007). Measuring faking in the employment interview: development and validation of an interview faking behavior scale. *Journal of Applied Psychology, 9*2(6), 1638-1656. doi:10.1037/0021-9010.92.6.1638

Li, J., Cardie, C., & Li, S. (2013, August). TopicSpam: A Topic-Model based approach for spam detection. *In Proceedings of the 51st Annual Meeting of Association for Computational Linguistics, 2*, 217-221.

Liddy, E. D. (2001). Natural language processing. *Encyclopedia of Library and Information Science*, 2126-2136. New York, NY: Marcel Dekker.

Mani, I. (2001). *Automatic summarization.* Amsterdam: John Benjamins Publishing.

Moore, C., Lee, S. Y., Kim, K., & Cable, D. M. (2017). The advantage of being oneself: The role of applicant self-verification in organizational hiring decisions. *Journal of Applied Psychology, 102*, 1493.

Mairesse, F., Walker, M. A., Mehl, M. R., & Moore, R. K. (2007). Using linguistic cues for the automatic recognition of personality in conversation and text. *Journal of Artificial Intelligence Research, 30*, 457-500.

Newman, M. L., Pennebaker, J. W., Berry, D. S., & Richards, J. M. (2003). Lying words: Predicting deception from linguistic styles. *Personality and Social Psychology Bulletin, 29*(5), 665-675. doi:10.1177/0146167203029005010

Nikita, M. (2016). ldatuning: Tuning of the Latent Dirichlet Allocation Models Parameters. R package version 0.2-0, URL https://CRAN.R-project.org/package=ldatuning

O'Brien, R. M. (2007). A Caution Regarding Rules of Thumb for Variance Inflation Factors. *Quality & Quantity, 41*(5), 673-690. doi:10.1007/s11135-006-9018-6

O'Connor, B., Bamman, D., & Smith, N. A. (2011, December). Computational text analysis for social science: Model assumptions and complexity. In *Second Workshop on Computational Social Science and Wisdom of the Crowds*.

O'Neill, T. A., Lee, N. M., Radan, J., Law, S. J., Lewis, R. J., & Carswell, J. J. (2013). The impact of "non-targeted traits" on personality testing faking, hiring, and workplace deviance. *Personality and Individual Differences, 55,* 162–168. doi:10.1016/j.paid.2013.02.027

Park, G., Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Kosinski, M., Stillwell, D. J., ... & Seligman, M. E. (2015). Automatic personality assessment through social media language. *Journal of Personality and Social Psychology, 108*, 934.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of Personality and Social Psychology, 46*(3), 598 – 609. doi:10.1037//0022-3514.46.3.598

Pauls, C. A., & Crost, N. W. (2004). Effects of faking on self-deception and impression management scales. *Personality and Individual Differences, 37*(6), 1137-1151. doi:10.1016/j.paid.2003.11.018

Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). *The development and psychometric properties of LIWC2015.* Austin, TX: University of Texas at Austin.

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Mahway: Lawrence Erlbaum Associates.

Pennebaker, J. W., & Stone, L. D. (2003). Words of wisdom: language use over the life span. *Journal of personality and social psychology, 85*(2), 291. doi:10.1037/0022-3514.85.2.291

Porter, S., & Yuille, J. C. (1996). The language of deceit: An investigation of the verbal clues to deception in the interrogation context. *Law and Human Behavior, 20*(4), 443-458. doi:10.1007/bf01498980

Putka, D. J., Beatty, A. S., & Reeder, M. C. (2017). Modern prediction methods: New perspectives on a common problem. *Organizational Research Methods*, *21*(3), 689-732. doi:10.1177/1094428117697041

Roulin, N., Bangerter, A., & Levashina, J. (2014). Interviewers' perceptions of impression management in employment interviews. *Journal of Managerial Psychology*, *29*(2), 141-163. doi:10.1108/jmp-10-2012-0295

Schank, R., & Abelson, R. (1977). *Scripts, plans, goals, and understanding: An inquiry into human knowledge structure.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Schlenker, B. R. (1980). *Impression management: The self-concept, social identity, and interpersonal relations.* Monterey, CA: Brooks/Cole.

Schneider, L., Powell, D. M., & Roulin, N. (2015). Cues to deception in the employment interview. *International Journal of Selection and Assessment, 23*(2), 182-190. doi:10.1111/ijsa.12106

Schwartz, H. A., Eichstaedt, J. C., Kern, M. L., Dziurzynski, L., Ramones, S. M., Agrawal, M.,... Ungar, L. H. (2013). Personality, gender, and age in the language of social media: The open vocabulary approach. *PLOS ONE, 8*(9). doi:10.1371/journal.pone.0073791

Silva, K. T. (2016). *Let's be Clear: The Effects of Interview Transparency on Applicant Reactions* (Unpublished master's thesis). University of Houston, Houston, TX.

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*(5), 587. 587-606. doi:10.1037//0021-9010.80.5.587

Tedeschi, J. T. (1981). *Impression management theory and social psychological research.* New York, NY:Academic Press.

Tsai, W. C., Chen, C. C., & Chiu, S. F. (2005). Exploring boundaries of the effects of applicant impression management tactics in job interviews. *Journal of Management, 31*(3), 108-125. doi:10.1177/0149206304271384

Video Interview Software & Recruiting Tools | HireVue Video Intelligence. Retrieved October 16, 2017, from https://www.hirevue.com/

Vrij, A. (2000). Detecting lies and deceit: The psychology of lying and implications for professional practice. London: Wiley.

Vrij, A. (2008). *Detecting lies and deceit: Pitfalls and opportunities.* Chichester, UK: Wiley.

Waung, M., McAuslan, P., DiMambro, J. M., & Mięgoć, N. (2017). Impression management use in resumes and cover letters. *Journal of Business and Psychology*, *32*(6), 727-746.

Weaver, J. D. (2017). *Predicting Employee Performance Using Text Data from Resumes* (Unpublished doctoral dissertation).Seattle Pacific University, Seattle, WA.

Zahedi, M., Ali Babar, M., & Treude, C. (2018, January). An Empirical Study of Security Issues Posted in Open Source Projects. *In Proceedings of the 51st Hawaii International Conference on System Sciences*.

Zhou, L., & Zhang, D. (2004). Can online behavior unveil deceivers? An exploratory investigation of deception in instant messaging. *In Proceedings of the 37th Hawaii International Conference on System Sciences*. doi:10.1109/ HICSS.2004.1265079

APPENDIX A

INTERVIEW QUESTIONS

1. Suppose you are in a team that is facing a problem or a conflict due to lack of communication between team members. How would you use your communication skills to solve a problem or resolve a conflict?
2. Suppose you went through a significant change (e.g., moving to a new place, working at a new job, or other major life changes). What would you do to adapt to it?
3. Suppose you have a great idea but there is an opposition to it. What would you do to persuade others to “see things your way”?
4. Suppose you are a leader of a team. How would you use your leadership skills to accomplish a team task?
5. In the highly technical work place today, we need employees who are willing to continue to learn and grow. What would you do to continue to learn and grow?

APPENDIX B

TAXONOMY OF FAKING BEHAVIORS AND THE INTERVIEW FAKING BEHAVIOR SCALE

Instructions:
Please think about the interview you just participated in. A researcher will view your interview and score your performance. What strategies from the list below have you used during your interview? Rate the extent to which you used each strategy by circling the appropriate number (1 = To no extent, 2 = To a little extent, 3 = To a moderate extent, 4 = To a considerable extent, 5 = To a very great extent). Please answer as honestly as possible.

I. SLIGHT IMAGE CREATION (to make an image of a good candidate for the job)
   i. Embellishing (to overstate or embellish answers beyond a reasonable description of the truth)
      1. I said that it would take less time to learn the job than I knew it would.
      2. I exaggerated my future goals.
      3. I exaggerated my responsibilities on my previous jobs.
      4. I exaggerated the impact of my performance in my past jobs.
   ii. Tailoring (to modify or adapt answers to fit the job)
      1. During the interview, I distorted my answers based on the comments or reactions of the interviewer/researcher.
      2. During the interview, I distorted my answers to emphasize what the interviewer/researcher was looking for.
      3. I distorted my answers based on the information about the job I obtained during the interview.
      4. I distorted my work experience to fit the interviewer's/researcher's view of the position.
      5. I distorted my qualifications to match qualifications required for the job.
      6. I tried to find out about the organization's culture and then use that information to fabricate my answers.

II. EXTENSIVE IMAGE CREATION (to invent an image of a good candidate for the job)
   i. Constructing (to build stories by combining or arranging work experiences to provide better answers)
      1. I told fictional stories prepared in advance of the interview to best present my credentials.
      2. I fabricated examples to show my fit with the organization.
      3. I made up stories about my work experiences that were well developed and logical.
      4. I constructed fictional stories to explain the gaps in my work experiences.
      5. I told stories that contained both real and fictional work experiences.
      6. I combined, modified and distorted my work experiences in my answers.
      7. I used made-up stories for most questions.

ii. Inventing (to cook up better answers)
    1. I claimed that I have skills that I do not have.
    2. I made up measurable outcomes of performed tasks.
    3. I promised that I could meet all job requirements (e.g., working late or on weekends), even though I probably could not.
    4. I misrepresented the description of an event.
    5. I stretched the truth to give a good answer.
    6. I invented some work situations or accomplishments that did not really occur.
    7. I told some "little white lies" in the interview.

iii. Borrowing (to answer based on the experiences or accomplishments of others)
    1. My answers were based on examples of job performance of other employees.
    2. When I did not have a good answer, I borrowed work experiences of other people and made them sound like my own.
    3. I used other people's experiences to create answers when I did not have good experiences of my own.

III. IMAGE PROTECTION (to defend an image of a good candidate for the job)

i. Omitting (to not mention some things in order to improve answers)
    1. When asked directly, I tried to say nothing about my real job-related weaknesses.
    2. I tried to avoid discussion of job tasks that I may not be able to do.
    3. I tried to avoid discussing my lack of skills or experiences.
    4. When asked directly, I did not mention my true reason for quitting previous job.

ii. Masking (to disguise or conceal aspects of background to create better answers)
    1. I did not reveal my true career intentions about working with the hiring organization.
    2. When asked directly, I did not mention some problems that I had in past jobs.
    3. I did not reveal requested information that might hurt my chances of getting a job.
    4. I covered up some "skeletons in my closet."

iii. Distancing (to improve answers by separating from negative events or experiences)
    1. I tried to suppress my connection to negative events in my work history.
    2. I clearly separated myself from my past work experiences that would reflect poorly on me.
    3. I tried to convince the interviewer/researcher that factors outside of my control were responsible for some negative outcomes even though it was my responsibility.

IV. INGRATIATION (to gain favor with the interviewer to improve the appearance of a good candidate for the job)

i. Opinion Conforming (to express beliefs, values, or attitudes held by the interviewer or organization)
    1. I tried to adjust my answers to the interviewer's/researcher's values and beliefs.
    2. I tried to agree with interviewer outwardly even when I disagree inwardly.
    3. I tried to find out interviewer's/researcher's views and incorporate them in my answers as my own.
    4. I tried to express the same opinions and attitudes as the interviewer/researcher.
    5. I tried to appear similar to the interviewer/researcher in terms of values, attitudes, or beliefs.
    6. I tried to express enthusiasm or interest in anything the interviewer/researcher appeared to like even if I did not like it.
    7. I did not express my opinions when they contradicted the interviewer's/researcher's opinions.
    8. I tried to show that I shared the interviewer's/researcher's views and ideas even if I did not.

APPENDIX C

ATTENTION AND REALNESS QUESTIONS

1. I have lived in Antarctica

2. Please select “To no extent”

3. Please select “To a very great extent”

4. I took the mock interview as seriously as I would normally take a real interview.

APPENDIX D

INTERVIEW EXPERIENCE

1. How many employment interviews have you completed in the past year?

2. How many employment interviews have you completed in your lifetime?

3. Will you be looking for a job in the next year?

4. Have you ever attended an interview workshop?

VITA

Elena Margaret Lawrence Auer
Old Dominion University
Department of Psychology
Mills Godwin Building, Room 250
Norfolk, VA 23529

Education
Bachelor of Arts, Psychology and Organizational Sciences, The George Washington University, May 2016

Publications

Landers, R. N., **Auer, E. M.**, Collmus, A. B., & Armstrong, M. B. (2018). Gamification science, its history and future: Definitions and a research agenda. *Simulation & Gaming.*

Landers, R. N., **Auer, E. M**., Collmus, A. B. & Marin, S (in press). Data Science as a New Foundation for Insightful, Reproducible, and Trustworthy Psychology. In Landers, R. N. (Ed.), *The Cambridge Handbook of Technology and Employee Behavior* (Cambridge Handbooks in Psychology). Cambridge: Cambridge University Press.

Landers, R. N., **Auer, E. M**., Helms, A. B., Marin, S, & Armstrong, M. B. (in press). Gamification for Employee Training and Development. In Landers, R. N. (Ed.), *The Cambridge Handbook of Technology and Employee Behavior* (Cambridge Handbooks in Psychology). Cambridge: Cambridge University Press.

Presentations

Auer, E. M., Behrend, T. S., Collmus, A. B., & Landers, R. N. (2017, April). How pay affects performance and retention in longitudinal crowdsourced research. Poster presented at the 32nd Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.

Auer, E. M., Landers, R. N., (2018, April). Measuring Deceptive Impression Management using Natural Language Processing. In Armstrong, M. B. (Chair) & Landers, R. N. (Chair). (2018, April). Using natural language processing to measure psychological constructs. Symposium presented at the 33rd Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Auer, E. M., Landers, R. N., Gore, R. (2018, April). *What Do Your Tweets Say About You? Measuring Trait Sentiment.* Poster presented at the 33rd Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL.

Cavanaugh, K. J., **Auer, E. M.**, Landers, R. N., Petor, J. R., & Kinney, T. B. (2017, April). Increases in applicant pool diversity attributable to unproctored internet-based testing. Poster presented at the 32nd Annual Conference of the Society for Industrial and Organizational Psychology, Orlando, FL.