Lights, camera, persuasion: Examining the impacts of impression management tactics on predictive validity

By Harley Harwood

A Thesis Submitted to Saint Mary's University, Halifax, Nova Scotia in Partial Fulfillment of the Requirements for the Degree of MSc in Applied Psychology.

August 2024, Halifax, Nova Scotia

© Harley Harwood, 2024

Approved by: Dr. Nicolas Roulin

Supervisor

Approved by: Dr. Eden-Raye Lukacik

Examiner

Approved by: Dr. Joshua Bourdage

Examiner

Approved by: Dr. Deborah Powell

External Examiner

Date: August 7th, 2024

Lights, camera, persuasion: Examining the impacts of impression management tactics on predictive validity

By Harley Harwood

Abstract

The study explored the use of self presentation behaviours (i.e., deceptive, and honest impression management; IM) in asynchronous video interviews (AVIs). Applicants use of IM in interviews has the potential to enhance (honest IM) or detract (deceptive IM) the predictive validity of interviews. However, scant research exists on the potential impacts of IM on predictive validity of job interviews. The current study employed a novel approach by collecting data in two stages: first, participants (n = 212) completed a mock AVI; next, participants (n = 168) completed two in-basket HR tasks to capture performance. Results indicated that honest self-promotion had a positive relationship with interview and task performance, and a significant indirect effect. Other IM tactics lead to some mixed and contrary findings. Proposed moderators (experience, age, and anxiety) did not impact results. Overall, honest and deceptive IM demonstrate their importance to both interview and job performance.

Keywords: Impression management; asynchronous video interviews; criterion related validity; selection interviews

Date: August 7th, 2024

Lights, camera, persuasion: Examining the impacts of impression management tactics on predictive validity

Job interviews remain one of the most popular selection tools used by organizations (McCarthy et al., 2010). Despite the popularity in both research and practice, research surrounding the use of impression management (IM) in job interviews has lagged behind. Applicant use of IM is defined as the ability for an applicant to honestly or deceptively enhance the presentation of oneself to the interviewer to obtain a more favourable impression and rating (Levashina & Campion, 2006). Deceptive IM use by applicants is worrying as it has the potential to destabilize the integrity of signals sent during the job interview process and consequently may diminish the criterion-related validity of job interviews (Roulin et al., 2016). A recent review by Melchers et al. (2020) suggests that much of the IM literature has narrowly focused on the relationship between IM and interview performance alone, and has therefore overlooked key criterion-related variables (e.g., job performance, counterproductive workplace behaviours). Further, meta-analytical evidence has demonstrated that deceptive IM use, on average, is a "much ado about nothing" as deceptive IM has no impact on interview performance (Ho et al. 2021b). Conversely, applicants use of honest IM seems to enhance interview performance ratings (Ho et al., 2021b), thus highlighting the importance of honest IM. Despite the importance of honest IM and the need to further explore the impacts of deceptive IM, researchers have not yet examined how these factors influence job performance, leaving a significant gap in the literature. Exploring how IM impacts the predictive validity of job interviews is crucial, as this is a fundamental question in both the IM and the selection and recruitment literature.

While it is important to understand the impact of IM use across all interview contexts, one particularly significant context is asynchronous video interviews (AVIs). Recently,

technology has transformed interviews and AVIs have taken center stage as a result. As of October 2022, the largest AVI provider, HireVue, has conducted over 33 million AVIs (Press Release Desk, 2022). While this initial surge in popularity may have been due to the Covid-19 pandemic, the use of AVIs continues to grow in popularity in a post-pandemic world due to the cost savings and flexibility associated with AVIs (Dunlop et al., 2022). The benefits associated with AVIs arise from their format, which eliminates synchronous two-way communication (Basch & Melchers, 2021; Gorman et al., 2018; Lukacik et al., 2022). Applicants' audio and video responses to interview questions are recorded and assessed afterwards (Lukacik et al., 2022). Despite the popularity of AVIs, relatively little is known regarding reliability and validly (Basch et al., 2020; Lukacik et al., 2022), thus compounding on the already existing gap in the literature when considering the impacts of IM us on job performance. Though some research examining the criterion-related validity of AVIs exists, it is still relatively scant and nuanced. For example, using mock AVI performance, self-rated performance at work and organizational tenure Gorman et al. (2018) demonstrated some evidence of criterion-related validity. As well, more recent work has presented some promising evidence about the criterion-related validity of automated scoring of AVIs (e.g., Hickman et al., 2022; Liff et al., 2024). Although recent literature has focused on using machine learning algorithms to score and rank applicants (e.g., Hickman et al., 2022; Liff et al., 2024), humans are still considered better sources of ratings when compared to automated systems (Liff et al., 2024). This is evident in human raters' ability to produce the highest estimates of criterion-related validity (Liff et al., 2024). Despite some research on the criterion-related validity of AVIs, evidence is still generally limited (e.g., correlations between self-report performance at work with an unrelated mock-interview [Gorman et al., 2018], correlations of self-report personality and interview personality scores [Hickman et al., 2022]) and lacks examination of IM behaviours in tandem.

The current study aims to explore the effects of applicant's IM use during the interview on performance in both the interview and subsequent tasks. By understanding the effects of IM on both interview performance and task performance I glean critical insight into the criterion-related validity of AVIs, and the potential effects of IM. Ultimately, these findings advance our theoretical understanding of the impact of both honest and deceptive IM tactics on the overall selection process. As well, this research enhances the practical understanding of applicant behaviours that organizations must consider when designing their selection processes.

Additionally, these findings can provide insight into applicant behaviour that benefits applicants, as the findings can guide applicants about how to best present themselves.

In the following sections I review the differences between AVIs and synchronous interviews, the use of IM in interviews and how it pertains to interview performance, the role the IM in interviews and how it relates to job performance, and finally how other variables (e.g., interview anxiety, age, experience) may moderate the relationships between IM, interview performance and job performance.

Asynchronous Video Interviews Versus Synchronous Interviews

When examining AVIs it is critical to understand the context in which they are conducted. While there are many similarities between AVIs and synchronous interview formats, there are also important differences. Synchronous interviews can be done over the phone, face-to-face or online through a digital platform. Comparatively, AVIs are strictly completed online through an AVI-specific platform. As well, AVIs tend to be used almost exclusively early in the recruitment and selection process, as they seem to have the most utility early on (Basch et al., 2022; Dunlop

et al., 2022; Liff et al., 2024). Inherently, AVIs provide no synchronous two-way communication (Lukacik et al., 2022). This lack of synchronous two-way communication may be problematic as applicants may find the interview "impersonal and cold" (Guchait et al., 2014). Rizi and Roulin (2023) explored ways to enhance the experience of AVIs, despite the lack of synchronous two-way communication, and found that using video instructions and video questions (vs. text) made participants feel more socially present, less anxious, and enabled more IM use (both honest and deceptive). While AVIs do have some shortcomings (e.g., not being able to prompt a candidate), there are potential unintended benefits. For example, having a third-party rater rate candidates lessens the impact of self-presentation tactics like IM (Barrick et al., 2009).

Another key difference between synchronous interviews and AVIs seems to be preparation time. Initial research found that participants performed better in AVIs compared to synchronous videoconference interviews (Langer et al., 2017). However, Langer et al. (2017) proposed this difference was likely due to the addition of preparation time (i.e., participants in the AVI condition had 60 seconds of preparation time, where as the videoconference interview condition had none). Later research examining differences in preparation time in AVIs found that participants that were given preparation time (vs. none) performed significantly better (Basch et al. 2021). However, recent research indicated this difference only mattered when participants used the preparation time (Roulin et al., 2023a). Regardless, most AVI vendors include some preparation time in their basic setup (Dunlop et al., 2022). In contrast, during synchronous interviews, applicants have no structured preparation time and feel pressured to speak immediately to avoid awkwardness (Ho et al., 2021a).

In addition to preparation time, AVIs offer a more structured format compared to synchronous interviews (Lukacik et al., 2022). For example, AVIs include previously researched

elements of interview structure (e.g., question consistency; Chapman & Zweig, 2005) and unresearched elements of interview structure (e.g., response time limits), whereas synchronous interviews may vary. The structure, or lack thereof, attenuates the effectiveness of selfpresentation tactics (Barrick et al., 2009). Ultimately, these differences are likely to change applicant behaviours and experiences. Of particular interest to this study, applicants' use of IM will likely vary in an AVIs compared to synchronous interviews. In fact, employed individuals believed AVIs provided the lowest opportunity to use IM tactics when compared to synchronous interviews (Basch et al., 2020). However, when compared to other pre-selection tools (an arguably better comparison point) AVIs provide the greatest opportunity to use IM (Basch et al., 2022). Despite offering more limited opportunities compared to synchronous interviews, AVIs still provide a context where applicants need to use self-presentation tactics (i.e., IM) to influence an employer's decision. Theoretical support for this is demonstrated by Levashina and Campion (2006), Marcus (2009) and Roulin et al. (2016), who suggests that when the ability, opportunity, and motivation are present applicant can use self-presentation tactics (IM) to enhance their performance ratings. Despite the limitations of their context, AVIs still allow for applicants to use IM tactics (i.e., the ability and opportunity are still present) and regardless of where AVIs are placed in a selection process, performing well increases ratings and, consequently, the chances of securing resources to live (i.e., motivation is also still present). Overall, AVIs provide an important context to consider the effects of self-presentation tactics, despite the differences from synchronous interviews.

Impression Management and Interview Performance

Generally, the selection process necessitates applicants' use of self-presentation behaviours in order to maximize interview performance, and therefore, to be seen as the most

desirable candidate and be selected (Marcus, 2009). More specifically, IM may be seen as a particularly useful self-presentation behaviour in an interview as it has the potential to directly influence applicant ratings. As such, IM tactics are defined as "job candidates' attempts to control and determine the images interviewers form of them regarding their [knowledge, skills, abilities and other characteristics (KSAOs)]" (Levashina and Campion, 2006, pp. 299). IM tactics are often characterized through a more negative lens (i.e., applicant faking) and are largely seen as potential threats to validity (Ho et al., 2021b; Levashina & Campion, 2006; 2007; Peck & Levashina, 2017). However, as noted by Levashina and Campion in their original 2006 paper, IM tactics can also be characterized through the use of genuine, non-deceptive (honest) behaviours. Additionally, Bourdage et al. (2018) further highlight the important distinction between IM tactics, noting that they may either be honest (i.e., an applicant truly possesses the KSAOs and promotes them) or deceptive. Generally, honest and deceptive IM have several more specific "sub-tactics" that vary in scope and target: other-focused (complimenting the organization or interviewer), defensive (addressing negative aspects about oneself), or selffocused (promoting oneself), where deceptive self-focused IM ranges from slight exaggeration (i.e., an applicant exaggerates the extent of their KSAOs) to more extreme and blatant forms (i.e., an applicant does not have the KSAOs required but lies and makes them up).

While honest and deceptive IM are statistically related (Bourdage et al., 2018; Moon et al., 2023), they are conceptually distinct, as can be seen in their differences in antecedents and outcomes (Bill et al., 2020; Bourdage et al., 2018). Largely, the effectiveness of IM tactics, particularly deceptive tactics, garner relatively mixed results regarding interview performance (Melchers et al., 2020). However, meta-analytic evidence suggests that honest IM tactics have a small positive association with interview performance, whereas deceptive IM tactics are not

related to interview performance (Ho et al. 2021b). Moreover, Bourdage et al. (2018) observed that among a sample of real job seekers, those who engaged in more deceptive IM tactics typically advanced through the interview stage but were ultimately not selected. Therefore, applicants using deceptive IM tactics should do so with caution, as the benefits are often short-lived.

More recently, several studies have examined the use of IM in AVIs, though primarily as an ancillary factor (e.g., less focused on specific tactics), and focusing solely on its impacts on interview performance rather than job performance. For example, Basch et al. (2022) found that, when considering AVIs in the context of preselection, intentions to use IM were particularly high. However, they did not differentiate between honest and deceptive IM. Research by Roulin et al. (2023a; 2023b) that examined more specific elements of AVIs (i.e., preparation time and rerecording, and training) and their effects on IM use have found relatively contrary findings. More specifically, use of preparation time in AVIs seems to decrease the use of honest IM (Roulin et al., 2023b), whereas an AVI training had little to no effect on applicant IM use (Roulin et al., 2023a). Despite the null effects of AVI design features, Roulin et al. (2023b) found that overall honest IM did indeed impact interview performance ratings. Contrarily, Roulin et al. (2023a) found that overall IM (both honest and deceptive) had no impact on interview performance. Yet, other research has found significant positive relationships with overall honest IM use and AVI performance (Moon et al., 2023; Rizi & Roulin, 2023).

Overall, despite the potential nuances of AVIs, the use of IM tactics seems to operate largely similarly to synchronous interviews, although there may be less potential to use IM in AVIs as a whole. However, when examining honest IM at the tactic level, as opposed to a broader level, there may be some explanation for the null findings. For example, there is likely to

be less opportunity to be honestly defensive or use honest ingratiation depending on certain elements of the interview (e.g., absence of negatively framed interview questions, or lack of interviewer to ingratiate). Some evidence for this can be found in Rizi and Roulin (2023), as adding an introduction video increased participants use of other-focus honest IM. When further examining honest IM tactics, it is clear that applicants will always be able to use honest self-promotion regardless of the setting, questions, or having a person to ingratiate, as promoting ones true KSAOs is helpful across all scenarios. To a certain extent this is also true for honest ingratiation, as applicants can complement the organization, even without a live interviewer. Whereas honest defensive tactics tend to depend on interview questions and ability for interviewers to follow-up on negative aspects of an applicant's answer. Given the findings in the literature on IM and the nature of the tactics themselves, I predict the following relationships:

H1: (a) Honest self-promotion and (b) honest ingratiation will have a moderate positive relationship with overall interview performance and (c) honest defensive IM will have a small positive relationship with interview performance.

Research Question 1: Will deceptive IM tactics be related to interview performance?

Impression Management and Task Performance

Examining the influence of IM tactics on interview criterion-related validity for predicting job performance is a much-needed area of research (Melchers et al., 2020). Recent studies tend to examine the impact of IM on interview performance, which meta-analytically has demonstrated either a small impact, or a null impact for honest and deceptive IM respectively (Ho et al., 2021b). However, the purpose of an interview is to select the most qualified applicants who will perform the best *on the job*. This demonstrates a clear need to go beyond interview performance and examine the impacts that honest and deceptive IM may have on task

performance. Barrick et al. (2009) conducted a meta-analysis on self-presentation tactics and their impact on both interview performance and job performance. They found that applicant use of IM was related strongly to interview performance (r = .47) and less-strongly to job performance (r = .15). Peck and Levashina's (2017) meta-analysis had similar findings, though there was less of a difference between the relationship with IM and interview performance (r = .22) and IM with job performance (r = .24). Thus, indicating that individuals who use IM tactics in interviews are likely to perform well in the interview and at the job. However, in both the Barrick et al. (2009), and the Peck and Levashina, (2017) failed to distinguish between honest and deceptive IM and their more specific sub-tactics. Given the failure to distinguish between the more nuanced aspects of IM and the Ho et al. (2021b) meta-analysis, which indicated no relationship between deceptive IM and interview performance, it is plausible that the use of honest IM was driving the relationships seen in the Barrick et al. and Peck and Levashina meta-analyses.

The use of honest IM is defined by and reflects an individuals' true KSAOs, and fit (Bourdage et al., 2018; Levashina & Campion, 2006). Therefore, it is expected that those who honestly promote their KSAOs also have the background to perform well on the job (i.e., task). As well, those who use honest ingratiation, will demonstrate better fit with the organization and therefore perform better. Additionally, applicants who use honest IM have been characterized as more agreeable and extraverted (Roulin & Bourdage, 2017). Individuals who are both agreeable and extraverted tend to perform better at work (Barrick & Mount, 1991; Lee et al., 2019). This suggests that those who use honest IM tactics are more likely to be better performers at work, as they possess highly sought-after traits of high performers. Alternatively, examining individuals who use deceptive IM they are either outright lying about their fit with the organization, their

KSAOs, or exaggerating the extent of their KSAOs (Levashina & Campion, 2006; 2007). Therefore, individuals who use deceptive IM tactics will *lack* said fit, and KSAOs (partially, or completely) and are likely to lack the necessary experience to perform well on the job. As well, those who use deceptive IM tactics tend to be less experienced, have lower levels of conscientiousness and honest-humility, and higher extraversion (Roulin & Bourdage, 2017). While individuals with higher levels of extraversion tend to perform well on the job (Barrick & Mount, 1991; Lee et al., 2019), the remaining traits associated with deceptive IM are fairly undesirable. For instance, honesty-humility (Lee et al., 2019) and conscientiousness have been found to be on average positively related to task performance (Barrick & Mount, 1991; Lee et al., 2019). When examining honest and deceptive defensive tactics however, these tactics tend to show particularly low relationships with interview performance (Ho et al., 2021b). These weak and null relationships paired with the idea that defensive IM is mainly about addressing potential issues (that may or may not be present in the interview process itself), it is expected that defensive IM is not likely to be related to job performance whatsoever. Ultimately, those who use honest IM tactics in the interview will be applicants who both have the necessary KSAOs, and desirable traits to perform well in the interview and therefore the job. Whereas those who use deceptive IM tactics are likely to be applicants who both lack the necessary KSAOs and have undesirable personality traits, and therefore should achieve lower task performance. Given the above mentioned I predict the following:

H2: (a) Honest self-promotion, (b) honest ingratiation will be associated with greater task performance, whereas deceptive (c) slight and (d) extensive image creation, and (e) deceptive ingratiation will be associated with lower task performance.

Research Question 2: Will honest defensive and deceptive defensive IM be associated with task performance?

H3: Interview performance will mediate the relationships between (a) honest ingratiation and (b) honest self-promotion and task performance.

Applicant Influences: Interview Anxiety, Experience, and Age

In addition to the main hypotheses of the study it is important to consider potential moderators of relationships between deceptive and honest IM and interview performance. Marcus' (2009) model of applicant faking in selection settings proposes several key reasons why moderators may influence self-presentation resources and therefore the ability to use selfpresentation tactics. For example, anxiety may inhibit the necessary cognitive resources to employ IM tactics effectively. Using this framework and building on the variables Marcus (2009) proposes, some potentially important moderators include interview anxiety, job experience, and age. Interview anxiety (i.e., feelings of apprehension and nervousness during an interview) has repeatedly demonstrated a significant positive relationship with deceptive IM tactics (Powell et al., 2021; Rizi & Roulin, 2023; Roulin, Wong, et al., 2023) and a mostly negative relationship with honest IM tactics (Powell et al., 2021), though the relationship between honest IM tactics and interview anxiety is typically weaker. As well, AVIs demonstrate both unique (e.g., feelings of creepiness from staring at one's own reflection, data security issues; Langer et al., 2017; 2018) and non-unique sources of interview anxiety. Interview anxiety may lead applicants to speak out of fear of silence, leading them to start lying to start their answer (Ho et al., 2021a). Therefore, the more anxiety an applicant may have the less effective their answers may seem, as they may be identified as visible anxious. Interestingly, applicants who are seen as anxious in job interviews are likely to be wrongfully overlooked, as anxious individuals perform better than

expected once on the job (Schneider et al., 2019). Similarly, lack of experience can be a source of interview anxiety and may lead applicants to "fill in the gaps" where they do not have sufficient experience to answer the interview questions (Ho et al., 2021a). Applicant experience is positively related to honest IM and negatively related to deceptive IM (Bourdage et al., 2018). This may further harm the perceptions of the effectiveness of honest and deceptive IM, further hurting applicant scores in the interview. In addition to interview anxiety and experience, applicant age is also an important factor related to use of honest and deceptive IM. Applicant age is related to IM such that younger applicants tend to use more deceptive IM, whereas older applicants tend to use more honest IM (Bourdage et al., 2018; Melchers et al., 2020). This is likely related to acquisition of KSAOs through sheer volume of experience as one gets older. This would also make applicants use of IM seem more legitimate the older they are, regardless of the honesty in applicants' use of IM. Clearly, these applicant characteristics play an important role in the way applicants present themselves. These characteristics may attenuate or enhance the relationships between deceptive IM and honest IM with interview performance, thus acting as moderators.

Applicants who are younger, less experienced, and more anxious are more likely to lack the ability to use honest IM tactics as they lack experience and therefore may turn to deceptive IM as a defense mechanism to make up for both real and perceived shortcomings. Given the lack of experience and younger age, the use of deceptive IM is likely to be used less tactfully by the individual, making them less convincing in "selling themselves" to the interviewer/rater. Alternatively, the use of IM use may be perceived as dishonest. Although interviewers are capable of detecting honest IM (and this positively influences performance ratings), when interviewers feel as though they are being deceived (even if they are incapable of detecting it)

they will give applicants lower performance scores (Roulin et al., 2015). Given that "nervous self-presenters" are likely to be less skilled at interviewing, attempts to promote themselves honestly or deceptively may be seen as deceptive by the rater and therefore negatively impact their ratings. Thus, I expect that applicant age, experience, and anxiety will moderate the relationships between deceptive and honest IM with interview performance (and, therefore job performance). For instance, for deceptive IM tactics, the null relationship will become slightly negative for younger, less experienced, and more anxious applicants (i.e., nervous selfpresenters). Whereas for honest IM tactics, the positive relationship will become weaker for nervous self-presenters. In contrast, those who are older, more experienced, and less interviewanxious can be characterized as "confident self-presenters" and can use this to their advantage. This may only be perceptual (e.g., just because a candidate is older does not necessarily mean they will have better/more KSAOs, but they may be perceived as more believable). As Bourdage et al. (2018) found, individuals who use more deceptive IM are also more likely to use honest IM. Thus, confident self-presenters' attempts to use deceptive IM may be erroneously seen as genuine self-promotion and therefore inflate their performance ratings in the interview, despite lacking the necessary KSAOs to perform the task well. Additionally, confident self-presenters can also draw on and promote legitimate KSAOs, which will be perceived genuinely and therefore legitimately increase performance ratings. For those applicants who use deceptive IM tactfully and receive greater interview ratings, it is expected that this will weaken (and potentially reverse) the relationship between interview performance and task performance. When considering confident self-presenters' ability to use honest IM, this will enhance the overall indirect relationship between honest IM, interview performance and task performance. Additionally, the lack of accounting for these moderators may explain the "net null effect" of

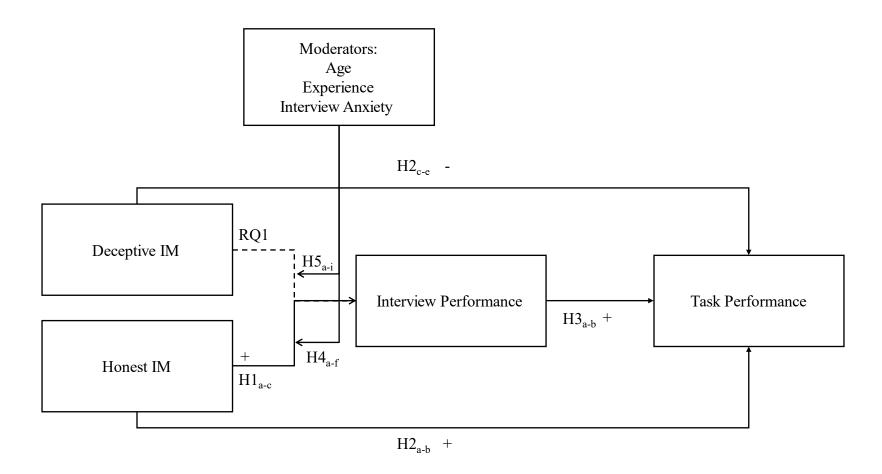
deceptive IM in Ho et al.'s (2021a) meta-analysis. Specifically, those who use deceptive IM effectively and those who use it ineffectively average out to a null effect. Overall, I predict the following:

H4: (a/b) Applicant age, (c/d) experience, and (e/f) interview anxiety will moderate the relationship between honest self-promotion/ingratiation and interview performance. This relationship will be strengthened for older, more experienced, and less anxious applicants (and weakened for younger, less experienced, and more anxious ones).

H5: (a/b/c) Applicant age, (d/e/f) experience, and (g/h/i) interview anxiety will moderate the relationship between slight image creation/extensive image creation/deceptive ingratiation and interview performance. This relationship will become positive and stronger for older, more experienced, and less anxious applicants (and weakened or reversed for younger, less experienced, and more anxious ones).

For a visual representation of all hypotheses model diagram see Figure 1 below.

Figure 1 *Proposed model for the current study.*



Note. This is a simplified model that collapses: a) the various IM tactics within each type of IM and b) the competencies being assessed in both the interview and task.

Method¹

Participants

Mote-Carlo Simulation²

The sample size was determined by a Monte-Carlo simulation as per best practice recommendations (Kelloway, 2015). A combination of the paramtest (Hugh, 2017) and lavaan packages (Rosseel, 2012) were used to build the model and simulate all paths of interest, total and indirect effects. Previous meta-analytic values were used where possible (i.e., interview performance to task performance r = .32 [Sackett et al., 2022], IM to task performance direct r =.11 [Barrick et al., 2009], honest IM tactics to interview performance: $r_{Honest \, self-promotion} = .23, r_{Honest \,$ Honest ingratiation = .15, $r_{Honest defensive IM}$ = .10 [Ho et al., 2021]). Otherwise, correlations between all types of honest and deceptive IM were averaged from three studies (Bourdage et al., 2018; Powell et al., 2021; Roulin & Bourdage, 2017) to determine the estimated coefficient values. When creating the model for the Monte-Carlo simulation, paths from deceptive IM and interview performance, as well as the moderators, were not included. This was done for one of two reasons: 1) previous research has demonstrated little to no effect (i.e., deceptive IM-interview performance relationship) or 2) a lack of research and evidence surrounding the estimation (i.e., moderators). After examining the results of the simulation, a total sample of 200 people was determined as best to get sufficient power ($\beta > .80$) for the total and indirect effects (see Appendix A for figures).

General Information

¹ This study was pre-registered on AsPredicted: https://aspredicted.org/LQG 5J4

² The power analysis did not include moderation or covariates for the sake of simplicity.

Participants were recruited via online research platforms. First, participants were recruited via Connect (n = 158) by CloudResearch and then on Prolific (n = 74) leading to a total sample of (n = 232). Participants were first screened for attention checks using three total attention check items, in which participants needed to pass two of three to be included. Two of the attention check items required participants to select the correct response and were embedded within other survey items (e.g., please select "To a moderate extent"), while the third item asked participants to select the Keanu Reeve's movie that would come first alphabetically. After screening participants for attention (n = 7), duplicates (on Prolific, n = 7) and data quality issues (n = 6) a total of 20 participants were removed from the dataset leading to a final sample of 212 participants in Phase 1. Phase 2 the sample was lower due to attrition, leaving a final sample of phase 2 data of 168 participants.⁴ Participants had to currently reside in Canada or the U.S., be fluent in English, have completed some form of education after high school, and have some hiring experience (as per the job description, see Appendix B). Participants were paid for completing both the mock AVI (~8.25 CAD) and the in-basket HR tasks (~18 CAD), as well as a performance dependant bonus in the HR task (~10 CAD). Important sample statistics (based on Phase 1) include the majority of the sample being white (65.40%; 11.80% = identified as black), with an even split of male (49.80%) and female (49.20%) with those identifying as non-binary or other being a small portion (1%), highly educated (with 51.20% holding a college/university degree, 33.20% hold a Master's, 12.30% with an associate's degree, and 3.30% holding a PhD), with an average age of 41.62 (SD = 11.39). Additionally, the sample was highly motivated to

³ Participants were recruited on both platforms due to time constraints.

⁴ T-tests were conducted with anxiety, motivation, interview performance (overall), age, work experience and all individual IM tactics (honest and deceptive) to ensure there were no differences between those who did and did not complete phase 2. Almost all t-tests were non-significant (ts ranged from -0.69 to 1.19, ps > .240), with the only exception being for honest ingratiation. Those who did not complete Phase 2 (M = 3.48) used more ingratiation compared to those who did complete Phase 2 (M = 3.09), t (211) = 2.29, p = .025.

complete the mock AVI (M = 4.39, SD = 0.55), and had lots of interview experience (M = 30.03, SD = 80.03) and work experience (M = 20.66, SD = 10.38), but lower experience with AVIs (M = 2.28, SD = 5.90) and HR-specific work experience (M = 4.93, SD = 5.78).

Procedure

Data collection was completed in two phases, similar to that of the Dunlop et al. (2020) study, which involves deceptively telling participants that performance in the first research task would determine selection for a second high paying research task. This was done to enhance the realism of the experimental setting. Phase one involved a mock AVI and phase two required participants to complete two in-basket human resources (HR) style tasks. The first phase asked participants to first read the job description for an HR position at a mock Oil and Gas organization (see Appendix B) and watch an introduction video in which a representative for organization gave instructions for the interview they were about to complete and conveyed messages about the mock organization of Western Oil Resource Group (WORG)⁵. Next participants completed a mock AVI answering six text-based interview questions (two per competency, see Appendix C). The six interview questions assessed three different competencies that are essential to HR functions (i.e., communication, decision making, and attention to detail). Each competency was assessed by two different interview questions and was rated on behaviourally anchored rating scales (BARS). Once the interview was complete participants were informed that they were deceived, and that all participants would be recruited for phase two. This was done so that participants could answer the post-AVI survey in a free manner, aiming to minimize socially desirable responding. The survey measured their IM tactics use, interview anxiety experienced and demographics (see Appendix D).

⁵ Rizi and Roulin (2023) demonstrated increasing media richness increases opportunity for IM use.

The second phase of the study involved inviting all phase 1 participants back to complete a set of high-paying tasks of HR generalist work, rewarding the top 10% performers with a bonus (see Appendix E for instructions). The tasks were consistent with the interview, using the same scenario and framing from the job interview (i.e., working for WORG, the mock Oil and Gas organization). Similar to the Dunlop et al. (2020) study, this portion of the data collection commenced three-weeks after the completion of the interview phase. During phase 2 participants were asked to complete an interview rating task for an engineer position at WORG (see Appendix F) and an article summary/policy design task (see Appendix G). The interview task required participants watch two videos from four candidates and rate their competences on a one to five scale for two dimensions (i.e., organization and resilience), rank-order the candidates, and write a recommendation. The article task required participants to read the article to formulate a summary email, and propose a policy based on the article. Once participants completed the inbasket exercises their performance was evaluated, and bonuses were awarded to the top 10% of study participants. The bonus was included to motivate participants to perform well on the task, increasing realism.

Selecting the Stimuli

All candidates used for the interview task were white males to minimize potential biases and confounds. Candidate's videos came from the Roulin et al. (2023b) study and were rated by graduate students in I/O psychology (n = 12) to ensure similarity on other potentially biased characteristics (e.g., attractiveness, warmth, etc.). The potential candidates were first identified from the dataset and picked for their "levels" of performance (i.e., high, good, okay, and bad). Once potential candidates were selected, the final four were chosen based on performance level,

the 95% confidence intervals (see Table 1), ANOVAs⁶ and the frequency for perceived ethnicity. Generally, all of the chosen candidates were perceived equally which should limit potential confounds.

Table 1Means, standard deviations, and confidence intervals for "candidates" to be used for Phase 2.

Variable	M	SD	Lower CI	Upper CI	Candidate 1	Candidate 2	Candidate 3	Candidate 4
attractiveness	3.18	1.25	2.47	3.89	3.58	2.83	3.83	2.58
competent	4.38	1.16	3.72	5.04	4.50	4.33	5.08	3.50
warm	3.87	1.18	3.21	4.54	4.58	3.25	3.67	3.92
knowledgeable	4.43	1.30	3.70	5.17	4.58	4.17	5.00	4.00
likeable	3.96	1.23	3.26	4.66	4.33	3.50	4.08	3.50
age	33.15	3.92	30.93	35.37	33.17	19.50	34.50	41.50
ethnicity	-	-	-	-	58.33%	100.00%	83.33%	83.33%

Note. $n_{raters} = 12$. Ethnicity shows percentage of "Caucasian" selected as perceived ethnicity. Candidate 1 = high performance rating (4.20), 2 = good performance rating (3.4), 3 = okay performance rating (2.8), and <math>4 = bad performance rating (1.2). These were also chosen to demonstrate a clear and objective difference in performance.

Measures

Interview Performance

Performance was measured with BARS (see Appendix C for questions and BARS). The BARS were developed alongside the interview questions and anchors for one, three and five were created. Performance was rated using custom python code using the ChatGPT 40 API. To do so, all interviews were transformed from video to audio files, so they could then be transcribed to text files using Whisper (a freely available tool to translate audio files to text). Whisper has demonstrated accuracy for English language transcripts in mental health research

⁶ ANOVAs NS (ps > .05), except for perceived age, and competence (Candidate's #2 = sig. younger vs. all other, #4 = sig. older vs. all; candidate #3 sig. more competent than candidate #4).

(Spiller et al., 2023). Once transcribed to text, the interviews were then able to be entered and rated by ChatGPT 40 API.

Initial testing of the efficacy of using ChatGPT to score interview performance involved the author manually scoring some participants' answers on the first two questions ($n_{responses} =$ 30). Additionally, prompts and interview responses were entered into the web version of ChatGPT for the same 30 responses. The prompts were entered to allow for ChatGPT to provide reasoning as to why it gave a specific rating (see Appendix H for full list of prompts). The reasoning provided by ChatGPT was examined to ensure ratings were consistent with the BARS and reasonable (see Appendix I for examples). Once all 30 responses were scored by both the author and ChatGPT correlations and intraclass correlations (ICCs) were calculated, where initial testing was both promising and reasonable enough to continue ($r_{average} = .79$, $ICC_{average} = .77$). Additionally, to ensure measures of performance are reliable and valid, 20% of participants performance was rated by two trained research assistants (RAs). Both RAs have degrees in psychology, and previous experience rating participant performance using BARS. Both raters were trained to use the BARS, this was done by scoring two participants in a meeting and discussing ratings and discrepancies when they appeared. Once RAs were trained the first 15 participant responses were scored and initial ICCs were calculated (ICC average = .72) and large discrepancies (differences in ratings of two or more) were identified and then reviewed by RAs. Once complete, final ICCs were calculated between RAs for 20% of the final sample (i.e., n participants = 40; ICC average = .78). Final ICCs were calculated between ChatGPT ratings and average RA ratings to ensure consistency ($n_{participants} = 40$; $ICC_{average} = .67$). The ICCs between ChatGPT ratings and human ratings were relatively consistent and similar estimates between

humans and machine learning scores found in Hickman et al. (2024; *ICCs* ranged from .59 to .89) and was therefore deemed acceptable to be used for the current study.

Task Performance

Task performance was rated in different ways depending on the competency being measured. As mentioned above the three competencies are decision making, communication, and attention to detail. ChatGPT was used to score several portions of task performance (see Appendix J for all prompts, and Appendix K for examples of initial testing reasoning). Similar to the interview performance ratings ChatGPT ratings were first tested using the author and the web version of ChatGPT, followed by a verification with trained RAs that scored 25% of the data (n = 42). Correlations between my ratings and the web version of ChatGPT for the first 15 responses revealed average correlations that were good ($r_{decision-making} = .85$, $r_{communication} = .81$). The average correlations exceed findings from Hickman et al. (rs range from .44 to .55; 2024) that compared human to machine learning scores (i.e., models trained specifically for the task at hand, compared to a general AI tool such as ChatGPT). Average ICCs between RAs were good (ICCc average = .78), and subsequent average ICCs between the average human ratings and ChatGPT were satisfactory (ICCs ranged from .44 to .86, ICC average = .73) and mostly similar to findings from machine learning studies (Hickman et al., 2024).

Decision Making. In the interview task, participants ranked the four candidates.

Candidates rank was based on previous scores obtained from Roulin et al. (2023b). The original scores in Roulin et al. (2023b) were completed by trained RAs who used BARS to rate the candidates. Performance scores were determined by every rank position in which the participant ranked the candidate off. Therefore, every position a candidate is ranked off of by the participant received a deduction (e.g., candidate 1 in rank 4, would be -3; candidate 2 in rank 2, would be a

0). Where participants had completed the recommendation (written task) but not the ranking, participant ranks were manually added in where possible (n = 5). For all participants who completed Phase 2 but left the rankings empty (and could not be determined by their written response) the maximum penalty was applied. Then scores for each rank were then summed and transformed by adding 10 (to eliminate negative scores). Final scores ranged from 0 to 10, with 0 meaning you ranked all candidates in the worst possible rank, and 10 meaning all candidates were ranked correctly. Decision making was also measured by the correlation between participant ratings of the candidate (on a one to five scale) and the objective scores obtained from Roulin et al. (2023b). These scores ranged from -1 to +1 where, +1 would indicate perfect ratings on both performance dimensions participants assessed of the candidates.

For the article task, decision making was rated by examining both the article summary and policy recommendation. For both, rubrics were developed to ensure participants made the correct decision to include major points in their summary and policy. These were scored similarly to the interviews, on a five-point BARS-like scale. The ChatGPT API was used to score all participant answers. The prompts given to ChatGPT included the rubrics and this process was repeated to score each component separately. Once ChatGPT had scored both the policy and summary for decision making, averages were calculated to determine the participant's overall decision-making score on the written task.

Communication. In the interview task, participants wrote a summary of each of the candidates' and their recommendation of the top candidate. Participants' answers were rated on two aspects of communication, the first being technical (e.g., grammar and structure of writing) and the second on clarity of the recommendation (e.g., providing clear, detailed writing).

Technical communication was also scored in the policy and article summary task. Each of the

tasks were scored using rubrics that have anchors for one, three, and five (similar to the decision-making rubric). The ChatGPT API was used to score all participants', and an average communication score was computed to give an overall rating of the communication competency.

Attention to detail. This was scored by using multiple-choice items for both the interview task and the article summary task. Items asked participants to recall details about the candidate interviews details regarding the article they read. Participants scores were totaled and divided to give a percent correct (i.e., correct = +1 point, incorrect = no point, and then divided by the total number of items). Each task asked participants to answer six total items (i.e., six for the interview task, and six for the article task).

Overall Task performance. To gather a measure of overall task performance all of the scores were averaged. However, to ensure that each of the scores were represented appropriately some were transformed. All of the scores on a one to five scale were subtracted by one, divided by four and then multiplied by 100 to give a percentage. The subtraction was done to ensure that percentages accurately reflected the score given (e.g., a score of 1 is best reflected as zero percent).

Impression Management – Honest and Deceptive⁷

Honest IM was measured using items from the Bourdage et al. (2018) and deceptive IM was measured with items from the Levashina and Campion (2007) measure of interview faking behaviours. However, only short versions of the honest defensive and deceptive image protection were used as they were not expected to be central to our hypotheses. Other tactics measured included honest self-promotion (14 items), honest ingratiation (8 items), deceptive slight image

⁷ I also tried to capture competency level faking from participants by asking them to recall their answers to two interview questions (i.e., that measured the same competency) and rate how (dis)honest they were when answering those questions (0 = completely dishonest, 100 = completely honest). However, these were extremely skewed (means > 88.53, with medians of 100 for all) and were therefore excluded from analyses.

creation (7 items), deceptive extensive image creation (9 items), deceptive ingratiation (6 items). Participants rated items on a on a 5-point Likert-type scale (1 = *To no extent*, 5 = *To a very great extent*) to determine the extent to which they engaged in the behaviours during the interview. The scales measure honest IM behaviours, as well as deceptive IM behaviours. An example item includes the following 'I told fictional stories prepared in advance of the interview to best present my credentials.' Additionally, adapted versions of the items were used to fit the AVI setting (similar to Roulin et al., 2023b). Cronbach's Alpha's were good (α's ranging from .84 – .95).

Interview Anxiety

Interview anxiety was measured using the Short State Interview Anxiety Scale (SSIA; Powell et al., 2024). The SSIA is a six-item measure of state interview anxiety is designed to measure general state interview anxiety. This six-item scale references general interview settings and therefore did not need to be adapted to the AVI setting of the study (e.g., "I felt anxious" or "I wish my interview did not bother me so much). Items are rated on a five-point Likert scale (1 = not at all, 5 = extremely) and the alpha was good ($\alpha = .93$).

Motivation to Perform

To ensure that participants were indeed motivated to perform during the mock interview participant motivation was measure using an adapted scale from Arvey et al. (1990). The scale includes ten items, and the questions substituted the word "test" for "interview" (e.g., "I wanted to do well on the test" was changed to "I wanted to do well on the interview"). The Cronbach's Alpha was acceptable ($\alpha = .93$).

Demographics

Participants were asked several questions about their demographics. These questions include age, gender, and job experience, AVI experience and interview experience.

Results

Validity Evidence of Measures

First, phase 2 task score variables were examined for validity, as non-valid measures of task performance would seriously threaten the research design and undermine the ability to test the hypotheses. To do so each of the competency's task scores were examined looking at means, standard deviations and correlations.

Decision-making

The results from the decision-making competency scores can be seen in Table 2. Surprisingly, both measures from the candidate evaluation component of phase 2 were problematic. Neither the candidate ranking score (r = .00, p = .953), nor the candidate rating correlation score (r = .04, p = .634) correlated with the written decision-making score, interview performance (r candidate ranking x overall interview = -.11, p = .161 and r candidate correlation x overall interview = .10, p = .208), or overall task performance (r candidate ranking x task performance = .14, p = .062, r candidate correlation x task performance = .08, p = .306). The likely reason behind these tasks not working was due to the limited variance. Specifically, 88.10% of the sample either ranked the candidates perfectly, or were one off (i.e., scores of eight or above, out of ten) and 86.30% had a rating correlation score over .60. Examining the means (M Candidate rating score = 8.67, M Participant-objective rating correlations = .76) further supports the notion that the task was too easy and, therefore, unable to discriminate between high and low performers. Evidence for validity for the written component of decision-making was satisfactory (rs >.19, ps <.05) for interview performance (both decision-making specific and overall). Thus, both the candidate rating and ranking were excluded from overall

task performance score, leaving just the average decision-making score from the writing tasks to be included in overall task performance calculation and analyses exploring decision-making task performance as an outcome.⁸

Communication

The results for the communication competency scores can be seen in Table 3. Notably, each component of both tasks correlated with each other, task performance for communication specifically, task performance overall, and interview performance (for both strictly communication and overall). However, the relationships between message clarity and the summary task score (r = .14, p = .073), and overall interview performance were exceptions (r = .15, p = .064). Despite the lack of significance, the relationship is still in the expected direction and therefore all components were deemed valid and were included in the overall task performance score.

Attention to Detail

Last, attention to detail tasks were examined (See Table 4). Overall, these results indicate that the attention to detail tasks were satisfactory mostly correlating in expected directions, with exceptions for attention to detail task from the article task and interview performance (overall r = .03, p = .694, and attention to detail specific performance r = .09, p = .256). However, both measures were included for comprehensiveness and construct validity of the tasks.

⁸ Analyses for testing hypotheses 1-3 were tested with full decision-making task performance and modified overall performance (including "excluded" portions of decision-making), results did not differ (see Appendix L for all tables).

Table 2Means, standard deviations and correlations for decision making tasks.

Variable	Scale	M	SD	1	2	3	4	5	6
1. Average written decision-making task score	1-5	2.78	0.87						
2. Candidate ranking task	0-10	8.67	2.06	.00					
3. Participant - objective ratings correlation	(-1) to (+1)	0.76	0.22	.04	.26**				
4. Task performance (decision-making)	0 to 1	0.73	0.13	.47**	.68**	.74**			
5. Interview performance (decision-making)	1-5	2.76	0.90	.19*	13	.00	.01		
6. Interview performance (overall)	1-5	2.53	0.73	.25**	11	.10	.11	.85**	
7. Task performance (overall)	0-100	51.86	15.55	.78**	.14	.08	.47**	.26**	.34**

Note. * indicates p < .05. ** indicates p < .01. Correlations were done using the apaTables Package with R (Stanley, 2021).

Table 3 *Means, standard deviations and correlations for communication tasks.*

Variable	Scale	M	SD	1	2	3	4	5	6	7
1. Policy task (communication score)	1-5	3.32	1.24							
2. Candidate recommendation task (communication score)	1-5	2.57	1.11	.24**						
3. Article summary task (communication score)	1-5	3.42	1.11	.47**	.18*					
4. Message clarity score (candidate recommendation task)	1-5	3.26	1.36	.25**	.56**	.14				
5. Task performance (communication)	1-5	3.12	0.85	.71**	.71**	.62**	.74**			
6. Interview performance (communication)	1-5	2.59	0.87	.18*	.19*	.16*	.15	.24**		
7. Interview performance (overall)	1-5	2.53	0.73	.29**	.22**	.15	.25**	.32**	.80**	
8. Task performance (overall)	0-100	51.86	15.55	.71**	.47**	.56**	.56**	.82**	.19*	.34**

Note. * indicates p < .05. ** indicates p < .01. Correlations were done using the apaTables Package with R (Stanley, 2021).

Table 4Means, standard deviations and correlations for attention to detail tasks.

Variable	Scale	М	SD	1	2	3	4	5
1. Article task % correct (attention to detail score)	0-100	55.46	23.75					
2. Candidate task % correct (attention to detail score)	0-100	60.62	27.63	.11				
3. Attention to detail (task performance)	0-100	58.04	19.19	.70**	.79**			
4. Attention to detail (interview performance)	1-5	2.25	0.86	.09	.19*	.20*		
5. Interview performance (overall)	1-5	2.53	0.73	.03	.23**	.19*	.84**	
6. Task performance (overall)	0-100	51.86	15.55	.42**	.53**	.64**	.37**	.34**

Note. * indicates p < .05. ** indicates p < .01. Correlations were done using the apaTables Package with R (Stanley, 2021).

Correlation Results

Correlations, means, and standard deviations are presented in Table 5. Correlations of note are examined to provide a general overview of study variables relationships and initial evidence of hypothesis support (or lack thereof). First, examining correlations of AVI performance, they demonstrate consistency (being that they are all related) and convergent and discriminant validity (as some overlap is expected, but none of the correlations are greater than .70; rs range from .49 to .59, ps < .01). A similar pattern can be noted for the task performance dimensions (rs range from .19 to .51, ps < .05). This demonstrates a good level of convergent and discriminant validity for both the interviews and job tasks. Next, the correlations between interview and task performance demonstrated evidence for criterion-related validity, as each competency was correlated in both the interview and the task (r decision-making = .19, p = .015, r communication = .24, p = .003, r attention to detail = .20, p = .013). Additionally, overall performance in the interview and task were significantly correlated, r = .34, p < .001.

Hypothesis 1 & RQ 1: Interview IM and Interview Performance. Examining the relationships between honest IM tactics and interview performance we can see that only honest self-promotion and attention to detail were significantly correlated in the expected direction demonstrating partial support for hypothesis 1a, r = .15, p = .032. Examining research question 1 results were somewhat contrary to previous research. More specifically, slight image creation was significantly related with attention to detail interview performance (r = .15, p = .034), decision-making interview performance (r = .18, p = .009), and overall interview performance (r = .18, p = .011), and extensive image creation was significantly related to decision-making interview performance (r = .15, p = .036). Additionally noteworthy was the lack of correlations between interview performance and honest ingratiation, interview performance and honest

defensive IM, and interview performance and deceptive image protection. Thus, hypotheses 1b and 1c were not supported, and evidence for deceptive IM (i.e., research question 1) was relatively mixed. Overall, these findings suggest some limited support for my hypotheses, and some contrary findings to previous literature.

Hypothesis 2 & RQ 2: Interview IM and Task Performance. Examining the correlations between honest IM in the interview and task performance there are some consistent findings and two unexpected findings. As predicted in hypothesis 2a, honest self-promotion was significantly positively related to communication (r = .22, p = .004) and overall task performance (r = .15, p = .046). However, honest ingratiation was negatively related to attention to detail task performance (r = -.16, p = .033) demonstrating contrary evidence to what was predicted in hypothesis 2b. Interestingly, deceptive ingratiation was significantly negatively related to attention to detail task performance (r = -.18, p = .024) and therefore demonstrating some support for hypothesis 2e. All other hypothesized relationships were non-significantly related to task performance. Thus, hypotheses 2c and 2d were not supported. Examining research question 2, honest defensive IM was significantly negatively related to attention to detail task performance (r = -.20, p = .010) but deceptive image protection was non-significant. Overall, this suggests limited support for the impact of interview IM on task performance.

Table 5 *Means, standard deviations, and correlations of study variables*

Variable	M	SD	1	2	3	4	5	6	7	8	9	10	11	12
1. Attention to detail (interview)	2.25	0.86												
2. Communication (interview)	2.59	0.87	.50**											
3. Decision making (interview)	2.76	0.90	.59**	.49**										
4. Interview performance (overall)	2.53	0.73	.84**	.80**	.85**									
5. Attention to detail (task)	58.04	19.19	.20*	.12	.13	.19*								
6. Communication (task)	3.12	0.85	.30**	.24**	.25**	.32**	.31**							
7. Decision making (written tasks only)	2.78	0.87	.33**	.09	.19*	.25**	.19*	.51**						
8. Task performance	51.86	15.55	.37**	.19*	.26**	.34**	.64**	.82**	.78**					
9. Honest defensive IM	2.64	1.03	09	11	10	12	20*	05	07	14				
10. Honest ingratiation	3.16	1.01	07	08	03	07	16*	05	07	13	.48**			
11. Honest self-promotion	3.53	0.89	.15*	.06	.12	.13	.00	.22**	.13	.15*	.23**	.42**		
12. Honest IM total	3.11	0.75	02	07	02	04	16*	.05	01	06	.76**	.84**	.69**	
13. Deceptive image protection	2.18	1.05	06	10	08	09	14	.00	.00	06	.19**	.19**	.08	.20**
14. Deceptive ingratiation	2.45	0.95	06	07	06	07	18*	03	.00	09	.31**	.58**	.32**	.53**
15. Deceptive slight image creation	1.57	0.74	.15*	.10	.18**	.18*	.06	.15	.07	.13	.03	.12	0.13	.12
16. Deceptive extensive image creation	1.39	0.71	.07	.03	.15*	.11	.09	.09	.06	.11	10	05	07	09
17. Deceptive IM total	1.90	0.64	.01	03	.04	.02	07	.06	.04	.02	.17*	.31**	.16*	.29**
18. Age	41.62	11.39	15*	16*	10	17*	05	32**	24**	27**	13	06	15*	13
19. AVI experience	2.28	5.90	.07	10	02	02	11	.01	.01	04	.11	.10	01	.12
20. Interview experience	30.03	80.03	.05	01	.05	.04	.02	02	.00	.00	.12	.07	.05	.11
21. HR experience	4.93	5.78	02	08	.04	02	13	10	07	11	.06	.19**	.01	.14*
22. Ethnicity	0.65	0.48	12	09	07	10	.22**	.07	02	.11	16*	15*	03	15*
23. Gender	0.50	0.50	07	.08	05	02	05	09	08	09	10	04	04	09
24. Anxiety	2.00	0.96	.00	.02	06	02	02	.00	04	03	.01	11	08	07
25. Motivation	4.39	0.55	.02	02	.03	.01	03	.00	05	04	.04	.26**	.20**	.20**

Table 5 Cont'd

Variable	13	14	15	16	17	18	19	20	21	22	23	24
13. Deceptive image creation												
14. Deceptive ingratiation	.53**											
15. Deceptive slight image creation	.34**	.39**										
16. Deceptive extensive image creation	.24**	.25**	.70**									
17. Deceptive IM total	.77**	.77**	.76**	.67**								
18. Age	13	17*	24**	16*	23**							
19. AVI experience	.08	.12	.10	.18*	.16*	10						
20. Interview experience	.02	.09	05	04	.02	.02	0.08					
21. HR experience	06	.06	03	05	02	.38**	.22**	.01				
22. Ethnicity	19**	21**	07	01	18**	.29**	16*	.06	.00			
23. Gender	10	03	13	14*	13	0.06	-0.13	.03	.06	.07		
24. Anxiety	.08	.03	.10	.11	.11	12	0.09	01	06	.06	.19**	
25. Motivation	09	.08	15*	24**	12	.10	-0.09	02	03	.11	.03	02

Note. * indicates p < .05. ** indicates p < .01. Correlations were done using the apaTables Package with R (Stanley, 2021). Ethnicity coded as 1 = white, 0 = non-white. Gender coded as 1 = Female and other, 0 = Male

Path analysis

To further examine hypotheses 1 and 2, as well as examine hypotheses 3-5, path analysis was conducted. Path analysis was completed using the Lavaan package in R. Path analysis provides an excellent way to test mediation model directly (Jenatabadi, 2015). In addition to testing the main base model, alternative models were tested (as per best practices; Kelloway, 2015) and included the moderation variables (i.e., anxiety, experience, age) as covariates.

Assumptions

Observed path analysis has the major assumption that measures be free from error (α > .70; Kelloway, 2015) and given that all scales used have Cronbach's alpha's greater than .70, this assumption was met. Additionally, normality was examined at both the multivariate level and the univariate level. This was done by using the MVN package in R (Korkmaz et al., 2014). More specifically, the Mardia test was used to assess skewness (Mardia Skewness = 4907.32, p < .001) and kurtosis (Mardia Kurtosis = 27.19, p < .001) at the multivariate level. At the univariate level the Anderson-Darling test indicated that only overall interview performance was normally distributed. Given these results the path model results used both bootstrapped standard errors and calculated bias corrected and accelerated bootstraps for confidence intervals (using 5000 bootstraps; Hayes & Scharkow, 2013) to correct for non-normality. SEM models were tested using maximum likelihood and full information maximum likelihood. Note that robust model estimates were used for fit statistics due to the violation of normality.

Base Model⁹

⁹ All effects remained in the same direction and significance when the model was run with pairwise deletion instead of FIML.

First, the base model (see <u>Figure 1</u>) was tested with each of the IM tactics having unconstrained paths to both interview performance and task performance. However, the base model did not include the moderators or any covariates (see Table 6 for base model statistics).

Model Fit. Examining the fit of the base model it is identified (i.e., model fit indices are "perfect" due to saturation) which is often the case for observed variable path analysis; Kelloway, 2015), χ^2 (0) = 0.00, robust CFI = 1.00, robust RMSEA = .00.

Hypothesis 1 & RQ 1. Examining parameter estimates some similar patterns emerge from the correlations. The path between honest self-promotion and interview performance was significant, $\beta = .20$, p = .011, 95% BCI = [0.03, 0.28], demonstrating further support for hypothesis 1a. However, all other paths between IM tactics and interview performance were non-significant, demonstrating a lack of support for hypothesis 1b and c. Examining research question 2, findings are constant with previous literature that deceptive IM is not significantly related to interview performance

Hypothesis 2 &RQ 2. Examining parameter estimates on task performance there are some findings of note. First, interview performance did predict task performance indicating criterion-related validity, $\beta = .27$, p = .005, 95% BCI = [1.81, 9.74]. Last, honest self-promotion predicted better task performance, providing support for hypothesis 2a, $\beta = .19$, p = .046, 95% BCI = [0.17, 6.70]. A similar pattern of non-significance was observed for all other IM tactics. These findings demonstrate support for hypotheses 2a, but do not support hypotheses 2b, 2c, 2d, and 2e. Examining research question 2 the use of honest and deceptive defensive IM in the interview does not impact later task performance.

Hypothesis 3. ¹⁰ Examining the indirect effects of the model for mediation there are some interesting findings. It is important to note again, due to violations of normality Bootstrapped confidence intervals (BCIs) are more reliable than p-values. First, interview performance did mediate the relationship between honest self-promotion and task performance, though this was not significant at p < .05 level, the BCIs did not include 0 (β indirect effect = .05, p = .070, 95% BCI = [0.17, 2.25]). Additionally, the total effect of honest self-promotion on task performance was significant, β = .24, p = .007, 95% BCI = [1.22, 7.39]. This provides strong support for hypothesis 3b. However, there was no significant indirect effect for honest ingratiation, thus hypothesis 3a was not supported. Other findings of note include the mediating role of interview performance in the relationship between deceptive slight image creation and task performance, again this was not significant at p < .05, but the BCIs did not include 0 (β indirect effect = .05, p = .070, 95% BCI = [0.13, 3.26]). This finding was unexpected as previous research has found that deceptive IM tends not to be significantly related to interview performance. All other total and indirect effects were non-significant at both p < .05 level and examining BCIs.

¹⁰ When models were tested using individual competencies results remained mostly consistent. Some exceptions include the following non-significant paths: honest self-promotion and communication interview performance (indirect effects); honest self-promotion and attention to detail task performance (indirect and total effects); and performance in the interview and task for decision-making and attention to detail.

Table 6 *Model fit statistics and estimates for base model.*

			Robust	Robus	t		
Relationship		χ^2	CFI	RMSE		AIC BIG	С
Model fit		0.00	1.00	0.00		1811.15 187	74.56
Interview performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	-0.07	0.06	.237	10	-0.18	0.05
Honest ingratiation	a2	-0.05	0.07	.507	06	-0.17	0.09
Honest self-promotion	a3	0.16	0.06	.011	.20	0.03	0.28
Deceptive image protection	a4	-0.07	0.05	.166	10	-0.17	0.03
Deceptive ingratiation	a5	-0.07	0.07	.367	09	-0.21	0.09
Deceptive slight image creation	a6	0.20	0.12	.089	.20	-0.01	0.45
Deceptive extensive image creation	a7	0.04	0.10	.727	.03	-0.19	0.21
Task performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1	-1.31	1.29	.311	09	-3.68	1.44
Honest ingratiation	c2	-1.70	1.71	.320	11	-5.03	1.62
Honest self-promotion	c3	3.32	1.67	.046	.19	0.17	6.70
Deceptive image protection	c4	-0.37	1.48	.803	03	-3.30	2.47
Deceptive ingratiation	c5	-0.91	1.93	.637	06	-4.50	3.06
Deceptive slight image creation	с6	2.24	2.31	.333	.11	-2.89	6.32
Deceptive extensive image creation	c7	0.45	2.49	.855	.02	-4.16	5.59
Interview performance	b1	5.74	2.02	.005	.27	1.81	9.74
Indirect	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1*b1	-0.39	0.39	.310	03	-1.40	0.17
Honest ingratiation	a2*b1	-0.26	0.42	.541	02	-1.31	0.43
Honest self-promotion	a3*b1	0.94	0.52	.070	.05	0.17	2.25
Deceptive image protection	a4*b1	-0.40	0.33	.229	03	-1.34	0.09
Deceptive ingratiation	a5*b1	-0.39	0.47	.415	02	-1.54	0.38
Deceptive slight image creation	a6*b1	1.13	0.75	.132	.05	0.13	3.26
Deceptive extensive image creation	a7*b1	0.20	0.60	.742	.01	-1.05	1.44
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+a1*b1	-1.70	1.34	.204	11	-4.14	1.20
Honest ingratiation	c2+a2*b1	-1.96	1.73	.259	13	-5.28	1.47
Honest self-promotion	c3+a3*b1	4.26	1.59	.007	.24	1.22	7.39
Deceptive image protection	c4+a4*b1	-0.77	1.53	.615	05	-3.82	2.21
Deceptive ingratiation	c5+a5*b1	-1.29	2.06	.530	08	-5.09	2.94
Deceptive slight image creation	c6+a6*b1	3.36	2.27	.138	.16	-1.38	7.54
Deceptive extensive image creation	c7+a7*b1	0.65	2.57	.799	.03	-4.41	5.78

Moderated Mediation

Hypotheses 4a-f and 5a-i were tested using a series of path models with each of the proposed moderators as an interactive term. A single model with all three moderators entered at once was too complex (i.e., R was unable to complete with an increase iteration count of 2500, hence divergence from pre-registration). Moderators were tested by creating a multiplicative term (e.g., honest self-promotion x interview anxiety), these terms were then added to the model as well as the respective moderator with paths *only* on interview performance. Overall, main effects remained consistent with results above (i.e., honest self-promotion generally predicts interview and task performance, with interview performance being a significant mediator), however, none of the moderators were significant. Rather, adding moderators lead to poor model fit. Additionally, when age, anxiety and experience were entered as covariates they were non-significant and similar patterns to previously discussed result held true. For full details see Tables 7-10.¹¹

¹¹ Simplified versions of the models were examined for thoroughness. This was done by testing the moderations using only honest self-promotion, or deceptive slight image creation. Results were non-significant in these simplified models as well.

Table 7 *Model fit statistics and estimates for moderated mediation model for age.*

Relationship	df	χ^2	Robust CFI	Robust R	RMSEA	AIC BI	C
Model fit	8	13.80	0.86	0.07		1800.93 18	90.79
Interview performance with	Path Label	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	-0.22	0.27	.403	32	-0.75	0.29
Honest ingratiation	a2	0.39	0.29	.183	.54	-0.16	0.97
Honest self-promotion	a3	0.23	0.25	.352	.28	-0.31	0.67
Deceptive image protection	a4	0.03	0.23	.892	.04	-0.44	0.47
Deceptive ingratiation	a5	-0.22	0.30	.470	29	-0.79	0.38
Deceptive slight image creation	a6	-0.10	0.53	.854	10	-1.06	1.02
Deceptive extensive image creation	a7	0.27	0.54	.617	.26	-0.83	1.28
Age	a8	0.01	0.02	.503	.22	-0.02	0.06
Honest defensive*Age	a9	0.00	0.01	.614	.22	-0.01	0.02
Honest ingratiation*Age	a10	-0.01	0.01	.183	76	-0.02	0.00
Honest self-promotion*Age	a11	0.00	0.01	.817	09	-0.01	0.01
Deceptive image protection*Age	a12	0.00	0.01	.657	15	-0.01	0.01
Deceptive ingratiation*Age	a13	0.00	0.01	.631	.21	-0.01	0.02
Deceptive slight image							
creation*Age	a14	0.01	0.01	.587	.29	-0.02	0.03
Deceptive extensive image							
creation*Age	a15	-0.01	0.01	.632	26	-0.03	0.02
Task performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1	-1.54	1.29	.232	10	-4.01	1.03
Honest ingratiation	c2	-1.40	1.62	.389	09	-4.68	1.72
Honest self-promotion	c3	3.45	1.68	.040	.19	0.23	6.79
Deceptive image protection	c4	-0.25	1.48	.863	02	-3.15	2.73
Deceptive ingratiation	c5	-0.99	1.92	.605	06	-4.82	2.62
Deceptive slight image creation	c6	2.49	2.20	.258	.12	-2.25	6.54
Deceptive extensive image creation	c7	0.05	2.49	.985	.00	-4.68	5.18
Interview performance	b1	5.74	2.02	.005	.27	1.81	9.74
Indirect	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	(a1+a9)*b1	-1.26	1.60	.431	03	-5.37	1.23
Honest ingratiation	(a2+a10)*b1	2.19	1.88	.244	06	-0.45	7.32
Honest self-promotion	(a3+a11)*b1	1.32	1.53	.387	.05	-1.35	4.81
Deceptive image protection	(a4+a12)*b1	0.16	1.32	.901	03	-2.66	2.80
Deceptive ingratiation	(a5+a13)*b1	-1.25	1.86	.502	02	-6.03	1.60
Deceptive slight image creation	(a6+a14)*b1	-0.53	3.23	.871	.05	-7.61	5.37
Deceptive extensive image creation	(a7+a15)*b1	1.52	3.29	.643	.00	-4.09	9.36
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+(a1+a9)*b		2.16	.195	13	-7.34	1.20
Honest ingratiation	c2+(a2+a10)*		2.56	.756	15	-3.59	6.45
Honest self-promotion	c3+(a3+a11)*		2.15	.026	.25	0.15	8.75
Deceptive image protection	c4+(a4+a12)*	b1 -0.09	2.11	.966	05	-4.05	4.35
Deceptive ingratiation	c5+(a5+a13)*		2.67	.402	08	-7.57	2.88
Deceptive slight image creation	c6+(a6+a14)*	b1 1.96	3.76	.601	.17	-6.20	8.86
Deceptive extensive image creation	c7+(a7+a15)*	b1 1.57	3.93	.690	.00	-6.37	9.07

Table 8 *Model fit statistics and estimates for moderated mediation model for anxiety.*

Relationship	df χ^2	Ro	bust CFI	Robust R	MSEA	AIC	BIC
Model fit		7.36 0.8		0.08		1819.03	1909.15
Interview performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	0.11	0.16	.494	.15	-0.19	0.42
Honest ingratiation	a2	-0.17	0.17	.322	24	-0.53	0.14
Honest self-promotion	a3	0.03	0.16	.850	.04	-0.30	0.34
Deceptive image protection	a4	-0.03	0.14	.819	05	-0.30	0.24
Deceptive ingratiation	a5	0.02	0.16	.885	.03	-0.29	0.34
Deceptive slight image creation	a6	0.03	0.30	.913	.03	-0.54	0.61
Deceptive extensive image creation	a7	0.35	0.24	.148	.34	-0.12	0.84
Anxiety	a8	0.05	0.30	.859	.07	-0.51	0.66
Honest defensive*Anxiety	a9	-0.10	0.08	.218	46	-0.27	0.05
Honest ingratiation*Anxiety	a10	0.07	0.08	.386	.33	-0.07	0.25
Honest self-promotion*Anxiety	a11	0.07	0.08	.37	.36	-0.10	0.23
Deceptive image protection*Anxiety	a12	-0.02	0.07	.760	09	-0.15	0.11
Deceptive ingratiation*Anxiety	a13	-0.05	0.08	.537	20	-0.20	0.10
Deceptive slight image creation*Anxiety	a14	0.06	0.13	.644	.18	-0.18	0.32
Deceptive extensive image creation*Anxiety	a15	-0.14	0.10	.158	43	-0.33	0.05
Task performance with	Path	b	se	p-value	В	CI-lower	CI-upper
Honest defensive	c1	-1.26	1.28	.324	08	-3.71	1.31
Honest ingratiation	c2	-1.69	1.64	.301	11	-4.87	1.55
Honest self-promotion	c3	3.23	1.62	.046	.19	0.04	6.42
Deceptive image protection	c4	-0.36	1.45	.802	02	-3.14	2.54
Deceptive ingratiation	c5	-0.95	1.89	.617	06	-4.76	2.74
Deceptive slight image creation	c6	2.24	2.22	.313	.11	-2.55	6.27
Deceptive extensive image creation	c7	0.43	2.39	.859	.02	-4.24	5.23
Interview performance	b1	5.86	2.02	.004	.27	1.86	9.83
Indirect	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	(a1+a9)*b1	0.04	0.54	.937	08	-1.00	1.21
Honest ingratiation	(a2+a10)*b1	-0.59	0.70	.395	.03	-2.51	0.39
Honest self-promotion	(a3+a11)*b1	0.60	0.63	.339	.11	-0.37	2.29
Deceptive image protection	(a4+a12)*b1	-0.30	0.52	.564	04	-1.55	0.58
Deceptive ingratiation	(a5+a13)*b1	-0.14	0.65	.833	05	-1.61	1.03
Deceptive slight image creation	(a6+a14)*b1	0.53	1.11	.634	.06	-1.71	2.89
Deceptive extensive image creation	(a7+a15)*b1	1.25	1.11	.261	02	-0.35	4.14
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+(a1+a9)*b1	-1.22	1.30	.350	17	-3.65	1.43
Honest ingratiation	c2+(a2+a10)*b		1.70	.179	08	-5.53	1.14
Honest self-promotion	c3+(a3+a11)*b		1.58	.015	.29	0.65	6.92
Deceptive image protection	c4+(a4+a12)*b	1 -0.66	1.51	.662	06	-3.52	2.39
Deceptive ingratiation	c5+(a5+a13)*b		2.04	.595	11	-5.07	2.94
Deceptive slight image creation	c6+(a6+a14)*b		2.14	.195	.16	-1.72	6.71
Deceptive extensive image creation	c7+(a7+a15)*b	1 1.67	2.53	.508	.00	-3.28	6.71

Table 9 *Model fit statistics and estimates for moderated mediation model for Job experience.*

Relationship	$df \chi^2$	Robu	st CFI	Robust R	RMSEA	AIC	BIC
Model fit	8 11.37	.93		0.05		1817.72	1907.83
Interview performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	-0.19	0.15	.204	28	-0.49	0.09
Honest ingratiation	a2	0.27	0.17	.113	.37	-0.04	0.61
Honest self-promotion	a3	0.25	0.14	.073	.31	-0.07	0.48
Deceptive image protection	a4	0.07	0.13	.589	.10	-0.19	0.31
Deceptive ingratiation	a5	-0.23	0.16	.147	31	-0.56	0.07
Deceptive slight image creation	a6	0.01	0.27	.974	.01	-0.49	0.54
Deceptive extensive image creation	a7	0.13	0.24	.572	.13	-0.35	0.60
Work experience	a8	0.03	0.03	.268	.40	-0.02	0.08
Honest defensive*work experience	a9	0.01	0.01	.452	.23	-0.01	0.02
Honest ingratiation*work experience	a10	-0.01	0.01	.086	78	-0.03	0.00
Honest self-promotion*work experience	a11	-0.00	0.01	.513	23	-0.02	0.01
Deceptive image protection*work experience	a12	-0.01	0.01	.160	29	-0.02	0.00
Deceptive ingratiation*work experience	a13	0.01	0.01	.260	.33	-0.01	0.24
Deceptive slight image creation*work experience	a14	0.01	0.01	.511	20	-0.02	0.03
Deceptive extensive image creation*work							
experience	a15	-0.01	0.01	.635	13	-0.03	0.02
Task performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1	-1.34	1.28	.297	09	-3.77	1.25
Honest ingratiation	c2	-1.55	1.63	.341	10	-4.66	1.78
Honest self-promotion	c3	3.25	1.67	.051	.19	0.04	6.58
Deceptive image protection	c4	-0.34	1.42	.812	02	-3.09	2.54
Deceptive ingratiation	c5	-0.99	1.85	.592	06	-4.50	2.71
Deceptive slight image creation	c6	2.26	2.26	.325	.11	-2.88	6.32
Deceptive extensive image creation	c7	0.43	2.46	.862	.02	-4.11	5.77
Interview performance	b1	5.97	2.04	.003	.28	1.97	10.03
Indirect	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	(a1+a9)*b1	-1.26	0.99	.256	01	-3.85	0.31
Honest ingratiation	(a2+a10)*b1	1.52	1.11	.171	11	-0.06	4.52
Honest self-promotion	(a3+a11)*b1	1.46	0.97	.133	.02	-0.09	3.88
Deceptive image protection	(a4+a12)*b1	0.37	0.77	.636	05	-1.11	2.05
Deceptive ingratiation	(a5+a13)*b1	-1.34	1.08	.213	.01	-4.22	0.20
Deceptive slight image creation	(a6+a14)*b1	0.10	1.59	.948	.06	-3.24	3.24
Deceptive extensive image creation	(a7+a15)*b1	0.76	1.43	.595	00	-1.60	4.39
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+(a1+a9)*b1	-2.47	1.70	.148	10	-5.71	1.00
Honest ingratiation	c2+(a2+a10)*b1	-0.04	2.01	.986	21	-379	4.25
Honest self-promotion	c3+(a3+a11)*b1	4.71	1.76	.007	.21	1.16	8.19
Deceptive image protection	c4+(a4+a12)*b1	0.03	1.74	.987	08	-3.32	3.46
Deceptive ingratiation	c5+(a5+a13)*b1	-2.34	2.17	.282	05	-6.55	1.92
Deceptive slight image creation	c6+(a6+a14)*b1	2.33	2.49	.350	.16	-2.85	6.98
Deceptive extensive image creation	c7+(a7+a15)*b1	1.19	2.66	.655	.02	-4.09	6.53

Table 10 *Model fit statistics and estimates for model with age, anxiety and AVI experience as covariates.*

Relationship	df	χ^2	Robust CFI	Robust	RMSE	A AIC	BIC
Model fit	3	10.29	.83	0.12		1793.72	
Interview performance		10.27		V.1.2		17,017	1000000
with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	-0.09	0.06	.122	12	-0.19	0.03
Honest ingratiation	a2	-0.03	0.07	.695	04	-0.16	0.11
Honest self-promotion	a3	0.16	0.07	.017	.20	0.023	0.29
Deceptive image protection	a4	-0.06	0.05	.225	09	-0.17	0.04
Deceptive ingratiation	a5	-0.08	0.08	.290	11	-0.2	0.07
Deceptive slight image							
creation	a6	0.18	0.12	.120	.19	-0.04	0.42
Deceptive extensive image							
creation	a7	0.02	0.10	.863	.02	-0.21	0.20
Age	a8	-0.00	0.01	.742	05	02	0.02
AVI Experience	a9	-0.01	0.01	.573	09	-0.03	0.01
Interview Anxiety	a10	-0.03	0.05	.598	04	-0.13	0.08
Task performance with	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1	-1.50	1.30	.248	10	-4.05	1.06
Honest ingratiation	c2	-1.48	1.69	.380	10	-4.75	1.96
Honest self-promotion	c3	3.48	1.66	.037	.20	0.30	6.85
Deceptive image protection	c4	-0.27	1.48	.856	02	-3.12	2.71
Deceptive ingratiation	c5	-0.95	1.98	.631	06	-4.81	2.89
Deceptive slight image							
creation	c6	2.50	2.27	.272	.12	-2.55	6.57
Deceptive extensive image							
creation	c7	0.06	2.44	.981	.00	-4.60	5.06
Interview performance	b1	5.67	2.04	.005	.26	1.73	9.69
Indirect	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1*b1	-0.49	0.39	.209	03	-1.51	.05
Honest ingratiation	a2*b1	-0.15	0.42	.720	01	-1.13	0.61
Honest self-promotion	a3*b1	0.93	0.54	.082	.05	0.16	2.34
Deceptive image protection	a4*b1	-0.36	0.34	.291	02	-1.28	0.13
Deceptive ingratiation	a5*b1	-0.46	0.48	.337	03	-1.72	0.27
Deceptive slight image	a6*b1						
creation	ao 01	1.03	0.75	.167	.05	-0.03	3.04
Deceptive extensive image	a7*b1						
creation		0.10	0.61	.869	.01	-1.12	1.38
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+a1*b		1.33	.133	13	-4.49	0.71
Honest ingratiation	c2+a2*b		1.70	.337	11	-4.84	1.84
Honest self-promotion	c3+a3*b		1.56	.005	.25	1.49	7.63
Deceptive image protection	c4+a4*b		1.53	.681	04	-3.58	2.46
Deceptive ingratiation	c5+a5*b	1 -1.41	2.08	.499	09	-5.47	2.65
Deceptive slight image		_					
creation	c6+a6*b	1 3.53	2.22	.112	.17	-1.10	7.70
Deceptive extensive image			2.40	0.40	0.1		7 00
creation	c7+a7*b	1 0.16	2.49	.949	.01	-4.66	5.03

Discussion

Findings and Theoretical Implications

The current study set out to answer a long-needed call in the recruitment and selection literature, specifically examining the effects of applicants' use of IM tactics on interview performance and subsequent task (job) performance. By implementing a novel experimental design, in which participants went through a multi-step simulated selection process, the study revealed that some IM tactics do impact both interview and job performance. Specifically, honest self-promotion and slight deceptive image creation were positively correlated with interview performance, which indirectly lead to greater performance on the job task. As well, this study adds more evidence of validity for the use of AVIs in selection contexts.

Validity and Beyond

Overall, performing well in the interview meant performing well on the job task, with estimates (r = .34) that exceeds recent meta-analytic estimates of interview criterion-related validity from Wingate et al. (r = .28; 2024) and comes close to the estimate from Sackett et al. (r = .42; 2022). Note that the differences in meta-analytic estimates between Wingate et al. and Sackett et al. may be due to inclusion criteria (i.e., real applicants only vs. not). This provides evidence that AVIs demonstrate similar levels of criterion-related validity as in person meta-analytic estimates. Despite the lack of human ratings and the context of a simulated selection process, these estimates are in line with previous literature suggesting that AVIs are indeed valid tools to use in a selection context (Gorman et al., 2018; Hickman et al., 2022; Liff et al., 2024). Additionally, this study demonstrates evidence and support for large language models (LLMs) to be used to score applicant interview responses. Despite moderate ICCs with the human raters, the validity estimates found exceed those in the Wingate et al. (2024) meta-analysis and come close

to the Sackett et al. (2022) meta-analysis. This finding demonstrates that LLMs can be used to assess candidates' performance in interviews, building on a study that found LLMs were able to validly and reliably assess certain personality traits based on participants' answers from an AVI (Zhang et al., 2024). It is important to note that humans have demonstrated better criterion-related validity compared to machine learning scores (Liff et al., 2024). However, this may be partially due to common method variance (i.e., human ratings for both interview and job performance), and with advances in technology, particularly LLMs, this may change in the near future. Although the current study used LLMs to score both the interview and the task, LLMs do not tend to be perfectly consistent (Jang & Lukasiewicz, 2023), nor do they have memory in the same way a human does.

Impression Management's Impact on Interview Performance

IM tactics have long been explored to understand how they may influence both interview and job performance. The current study found that in the context of AVIs IM tactics do seem to play a role. First, examining correlational results, both deceptive slight and extensive image creation signaled increased interview performance, in specific dimensions of interview performance. In addition to more basic correlational evidence in the study, path analysis was also conducted and were only partly consistent with the correlation findings regarding deceptive tactics and interview performance. Overall, this suggests that deceptive slight and extensive image creation *may* effectively work to inflate some interview performance scores, though the evidence is relatively weak. These findings are somewhat concerning as this would suggest that those who fake their KSAOs can inflate their ratings of performance on certain competencies. These findings contradict the Ho et al. (2021b) meta-analysis, which suggests that deceptive IM tactics do not impact performance ratings. This discrepancy implies that while meta-analytic

evidence is valuable, it may not capture the nuances of specific situations (e.g., deceptive IM in AVIs). For instance, certain jobs may be more susceptible to score inflation in the interview from slight and extensive image creation than others. In the case of HR practitioners, they may be able to fly under the radar more, as they have the experience of knowing what may alert their answers as obvious faking. Therefore, leading HR practitioners to engage in better exaggeration and faking on certain competencies. As well, AVIs may be more susceptible to slight and extensive image creation as there is no way to follow-up or verify the potential lies and exaggerations being woven into an applicants' answer. This may suggest that without any ability to prompt specific instances applicants may be able to successfully fake their KSAOs early in the selection process. Though, it is important to note that follow-ups may lead to more faking as applicants may use follow-ups as signals they are on the right track, leading to greater exaggeration (Levashina & Campion, 2007). Alternatively, this could signal that LLMs may be vulnerable to applicants who fabricate and exaggerate their KSAOs. Although human interviewers are not good at detecting those who used deceptive IM either (Roulin et al., 2015). Image protection and deceptive ingratiation did not impact interview performance and are in-line with both the predictions and prior evidence of Ho et al. (2021b).

Regarding findings on honest IM in AVIs, honest self-promotion had a significant impact on interview performance, as evidenced by both correlations and path analysis. However, when examining honest ingratiation and honest defensive IM, these tactics did not impact interview performance ratings, despite having an introduction video with a representative from the organization, who shared information about the interview and the company itself, in the beginning of the interview (and therefore a target other than the organization to ingratiate). Perhaps the effects of honest IM tactics on overall performance are limited by the structured

nature of the interview used in study, especially for honest ingratiation and honest defensive IM. Although participants did indeed use more honest ingratiation than typical in AVI research (M =3.16, compared to an average M = 2.49 across other AVI studies; Rizi and Roulin, Roulin et al., 2023a; Roulin et al., 2023b), this did not seem to influence overall interview performance. However, these findings are less surprising when considering the use of BARS and ChatGPT to score the applicants (both elements that add structure to the interview). When interviews are highly structured the impact of IM is decreased (Barrick et al., 2009). As well, interviewers in practice may tend to rate overall performance at a more global level, as opposed to using BARS, and may therefore applicants may be rewarded with higher ratings for using tactics such as honest ingratiation or honest defensive IM. Given that more global ratings of performance are more susceptible to the influences of IM tactics, these relationships may have stronger (and more significant) if the study had used such ratings. Additionally, the use of ChatGPT is likely to remove some of the biases that can creep into structured scoring with human raters. For example, ChatGPT scoring examines only transcripts, whereas human ratings usually involve watching of videos, and therefore leaves room for biases based on appearance to influence even structured ratings. Despite the use of structured scoring, the overall findings are in line with the Ho et al. (2021b) meta-analysis, which also found that honest self-promotion has the strongest relationship with interview performance, with honest ingratiation and honest defensive IM seeming to matter less. These findings demonstrates that when an applicant legitimately promotes their KSAOs, they are rewarded for doing so.

Impression Management and Task Performance

Deceptive and honest IM tactics did play a role in job performance, overall, in line with previous findings from Barrick et al. (2009) and Peck and Levashina, (2017). However, different

from the Barrick et al. (2009) and the Peck and Levashina (2017) meta-analyses (and the studies themselves) the current study examined IM by considering both honest and deceptive IM at the tactic level. More specifically, honest ingratiation and honest defensive IM were both negatively related to aspects of task performance (contrary to expectations), and honest self-promotion was positively associated with task performance. Whereas deceptive ingratiation was negatively associated with aspects of task performance. Honest self-promotion and deceptive ingratiation being related in this way aligns with predictions and make sense. Those who use honest selfpromotion are able to not only highlight their KASOs but also demonstrate them, and therefore it is expected that they will perform well on the task. The finding of deceptive ingratiation is also in line with predictions as those who spend time deceptively tailoring their answers are more likely to lack the necessary KSAOs and therefore perform worse on the task. This finding also aligns and extends on previous findings from Bourdage et al. (2018) finding that those who used deceptive IM may make it beyond initial "checkpoints" in the selection process but will largely be unsuccessful later on. Additionally, the overall findings both support and contradict those from Peck and Levashina (2017), as they found that self (r = .18) and other-focused (r = .25) IM were related to job performance. However, these findings also ignore the nuances of honest and deceptive IM, as well as categorize the tactics more broadly, which may be the source of contradiction.

While it is expected that honest self-promotion is related to (and predicts) task performance and deceptive ingratiation is related to task performance, it was unexpected that honest ingratiation and honest defensive IM were negatively related to task performance. Some reasons for these relationships may be explained by the tactics themselves. For instance, while one may honestly compliment the organization, this may not fit the structured question setting,

and therefore detract from an individual's answer and serve as filler, thus those who use honest ingratiation may be less likely to perform well at a task compared to someone who does not use honest ingratiation to "fill space." Whereas for individuals who use honest defensive IM they may honestly have "deficits" in their KSAOs, and these deficits may be so glaring they felt the need to share them in an interview (however, the interview did not include questions asking about deficits) and therefore, those who felt the need to explain any negative aspects honestly may also honestly perform worse on the tasks. Some evidence for this may be found in the relationship between honest defensive IM and perceived interview difficulty (Bourdage et al., 2018). Additionally, these findings may signal that participants who engaged in these tactics in the context of the current study were generally inattentive. Given that the questions were either situational or past behaviour (with no focus on mistakes in the past) an individual choosing to either use honest ingratiation, or honest defensive IM may not have been paying full attention to the interview question themselves, demonstrating their inattentiveness. Thus, participants who engaged in these tactics were likely to be inattentive while completing the job tasks and therefore scored lower on the attention to detail on the task. This finding was indeed partially supported by supplementary correlations (see Appendix M for correlation table) that demonstrated the more one used honest defensive IM in the interview, the less time they spent on task one (r = -.18, p =.021) and task two (r = -.16, p = .037). It is important to note that while time on task may indicate attention on a given task, it may not be fully reflective of the attention paid during the task.

Over and above the specific predictions in the study, these findings also reveal a more nuanced and full perspective compared to the Barrick et al. (2009) and Peck and Levashina meta-analyses. Not only did the current study examine different tactics that were both distinguished

between honest and deceptive IM, the relationships found between IM tactics and job performance were also somewhat stronger (β = .15) than what was presented in Barrick et al. (2009) with an r = .11 and an r = .11 of self-focused IM and job performance in Peck and Levashina (2017). Though this may be due to the lack of distinguishing IM tactics.

Examining the indirect effects from path analysis revealed two key mediation paths for honest self-promotion and deceptive slight image creation. While honest self-promotion is in-line with predictions, and further demonstrates that those who honestly possess and promote the required KSAOs will do better in the interview and therefore also better on the job, the finding that the relationship between deceptive slight image creation and task performance was significant was not. Additionally, the implication of this finding may provide some evidence for the argument that faking requires a certain level of cognitive ability and may therefore be indicative of an applicant's cognitive ability (Melchers et al., 2020). In sum, these findings reveal some new insights and confirm others that have implications for performance on the job. Thus, providing a new spark for interview IM researchers while answering a much-needed call to research.

Moderators or Covariates?

Finally, examining the findings regarding my hypotheses surrounding moderators revealed that despite some significant relationships (e.g., age being significantly negatively related to most study variables) they were overall unimportant and non-significant. While there is a possibility that moderated mediation effects require a much larger sample size in order to have the power to detect significance, the sheer lack of significance when including age, anxiety and experience as covariates may indicate these variables were not as important as previously thought. Furthermore, when examining the correlations of age, anxiety, and experience there

were several notable non-significant and reversed relationships. More specifically, age correlated negatively with deceptive IM, which is consistent with previous findings (e.g., Bourdage et al., 2018; Roulin et al., 2023a). However, age also correlated negatively with some honest IM tactics in the current study, which is opposite to correlations found in previous literature (e.g., Bourdage et al., Roulin et al., 2023a). These findings may be explained by the setting, such that younger participants felt more comfortable with the AVI setting and were able to use more IM in general compared to older participants who were possibly less comfortable with the AVI context.

Other correlations, or lack thereof, to note include those with honest and deceptive IM with anxiety. Examining findings with honest IM and interview anxiety, findings tend to be null for the most part (e.g., Powell et al., 2021; Rizi and Roulin, 2023; Roulin et al., 2023b), with an exception found in Powell et al. (2021) in which honest self-promotion was significantly negatively correlated with interview anxiety. Findings from the current study are consistent with these overall null findings, suggesting that interview anxiety may not play an important role in the use of honest IM. However, Powell et al. (2021) found that all deceptive IM tactics (except deceptive extensive image creation) were significantly positively correlated with interview anxiety. Similar patterns were observed by Rizi and Roulin (2023), though the correlations were stronger. Comparing findings from this study to previous literature, there were no significant correlations between deceptive IM tactics and interview anxiety. While this may be due to the measure of anxiety (i.e., most other studies tend to use a version of the Measure of Anxiety in Selection Interviews; McCarthy & Goffin, 2004), it may have more to do with the relatively low levels of anxiety (M = 2.00) observed here. While it may suggest that anxiety was not a significant influence in this particular instance, that should not overshadow previous findings on interview anxiety.

When examining experience, the study revealed that experience had no influence on IM, honest or deceptive. These findings are likely due to the nature of experimental design, or the type of individuals recruited for the study (e.g., HR experience was quite low on average [M = 4.93] considering the sample age [M = 41.62]) and should therefore be considered with caution. Overall, the proposed moderators seemed to lack significance almost across the board in both the moderations and correlations (with age being the only exception examining the correlations). Thus, age, anxiety, and experience do not moderate the relationship between honest and deceptive IM and interview performance, but they may still be important when considering IM use in job interviews.

Practical Implications

Overall, the findings from the current study may be applicable for both organizations and applicants alike. First, organizations can feel confident that AVIs are a valid tool to use for applicant selection. While there are issues with AVIs (e.g., applicants find them cold, and intrusive; Langer et al., 2017; 2018), they are valid tools to be used and offer many benefits (e.g., flexibility of time and location; Lukacik et al., 2022). As well, early research found that AVIs may be easier than synchronous interviews (Langer et al., 2017), however, recent research from a sample of real applicant interviews found contrary evidence to what was once a widely supported performance gap between face-to-face and videoconference interviews (Langer et al., 2024). Although this is not direct evidence for AVIs, a similar trend may also emerge once the novelty of AVIs wears off. For example, despite the current sample being collected in 2024 there was still overall inexperience with AVIs (completing an average of 2.80 prior to the study, SD = 5.90) compared to general interview experience (M = 30.03, SD = 80.03). Additionally, this study provides preliminary evidence that LLMs can be used to score job interviews. Second,

organizations should still be cautious when using interviews alone to make selection decisions as they are still susceptible to applicant faking. While meta-analyses show a lack of impact of deceptive IM on interview performance, there was potential concerning relationships in the current study. Deceptive IM can lead to increased ratings of performance in the interview, and these gains may shift who is selected, introducing issues with the validity of the selection process. When pairing this with the inability to detect those who fake (Roulin et al., 2015) this creates even further cause for concern. Further, problematic and new findings from the current study found that use of honest defensive and honest ingratiation negatively relate to task performance. Although Roulin et al. (2015) found that interviewers tend to be more accurate when identifying the use of honest IM tactics there still remains the issue of how distinguish honest IM at the tactic level accurately and how to handle those who have been identified as using honest IM. Additionally, there is likely important performance dimensions that were not captured in the current study that honest defensive and honest ingratiation may importantly predict (e.g., organizational citizenship behaviours, or counter-productive workplace behaviours). Finally, examining how these findings may apply to an applicant completing an AVI or any interview, in general, suggest that the best approach is to honestly promote one's skills. While it may be tempting to exaggerate one's responses (and even somewhat successful) it is best to stick to the truth. Keeping to answering the interview questions and promoting one's previous KSAO leads to better interview scores, which demonstrates a better ability to perform on the job. Additionally, for individuals who may be tempted to fake in their interview, they should avoid doing so. While faking in the interview may not lead to lower interview performance ratings, it will lead to a mismatch of skills on the job, leading to poor performance.

Limitations and Future Directions

There are a few key limitations to the current findings. First, the overall setting and sample. While the experiment was carefully designed, and demonstrates validity evidence, the setting is still an experiment. Ultimately, this may not reflect the exact relationships that exist in an applied setting. Additionally, while the sample was screened for hiring experience (i.e., what was available as a screener via Connect and Prolific), this does not necessarily equate to direct HR experience. While hiring is one component of HR roles, it is not the only function of these positions. This likely introduced some error into study, potentially weakening the findings. For example, a significant portion of the sample may have included individuals who work in factories and interview and hire individuals as a part of their job, but whose main focus is on other aspects (e.g., organizing and scheduling deliveries). This would indicate that their KSAOs are more reflective of their primary duties (e.g., troubleshooting and diagnosing issues with machinery), as opposed to HR specific duties. Second, while the second portion of the experiment asked individuals to not use generative AI like ChatGPT there is no way of ensuring participants did not use it. There is currently no reliable way to detect AI use in such settings (Harwood et al., 2024) so it is plausible that generative AI may have been used in both the interview and the job tasks given to the participants. Further, AVIs are susceptible to faking using ChatGPT (Canagasuriam & Lukacik, 2024). Despite this limitation, there was still evidence of validity for the interview and task. Third, the HR setting of the current study may limit and shape some of the relationships found. For example, while faking is largely seen as alarming, there may be contexts in which being able to effectively use slight deception may indicate performance. For example, in a sales role being able to exaggerate (i.e., akin to using slight image creation) may be instrumental in closing a deal. Last, the current research was highly structured. While this is the

ideal way to conduct interviews from a validity perspective (Sackett et al., 2022), it may not reflect the way that majority of interviews are conducted. The high degree of structure, particularly with BARS, may limit the potential influence of IM tactics and thus weaken the relationships observed. In contrast, in less structured interviews, the influence of IM tactics may become stronger and have a greater impact on applicant performance ratings.

For future research, these relationships need to be examined in other contexts including other industries, and applied settings and the use of LLMs should be considered and tested more explicitly. First, examining other contexts where the relationship may change is important to provide a better and more nuanced understanding of applicant IM tactic use. As well, exploring and testing these relationships with real applicants is also crucial to understand if these effects are under or over exaggerated. Second, research should examine these concepts using unstructured interview approaches. By exploring these relationships and variables with an unstructured interview, researchers may be able to further highlight pitfalls of using unstructured scoring when selecting for candidates. Last, exploring and testing LLMs may be important and beneficial in a few ways. One potentially important reason is exploring if validity relationships hold up if AI is used by the applicant/incumbent in the interview and on the job. There is a potential that this relationship does not hold true if use of AI is prohibited on the job, though this may be unlikely/difficult for organizations to stop or monitor. Additionally, while there was good evidence that LLMs were capable of scoring participant answers, this should be more fully explored and tested against other automated scoring options, like machine learning. Some preliminary research by Zhang et al. (2024) has tested the validity of LLM scoring, but more evidence is needed to test the efficacy of these approaches in other contexts (i.e., beyond personality interview questions) and compared to other types of automated scoring.

Conclusion

Overall, the current study set out to examine the impact that honest and deceptive IM can play in the criterion-related validity relationship for job interviews. Additionally, by combining this with novel elements such as AVIs and use of LLM scoring additional insights into validity were also explored. Honest self-promotion played a pivotal role in predicting both interview and job performance, indicating that honestly promoting one's KSOAs reveals performance beyond the interview. Additionally, other IM tactics seem to play some role on task performance, but evidence was relatively weak. In sum the current study provides valuable insights into the impact of IM tactics on the job interview and subsequent job performance, highlighting the importance of honest self-promotion and the potential pitfalls of deceptive IM tactics.

References

- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x
- 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.

 https://doi.org/10.1037/a0016532
- Basch, J. M., & Melchers, K. G. (2021). The use of technology-mediated interviews and their perception from the organization's point of view. *International Journal of Selection and Assessment*, 29(3–4), 495–502. https://doi.org/10.1111/ijsa.12339
- Basch, J. M., Melchers, K. G., & Büttner, J. C. (2022). Preselection in the digital age: A comparison of perceptions of asynchronous video interviews with online tests and online application documents in a simulation context. *International Journal of Selection and Assessment*, 30(4), 639–652. https://doi.org/10.1111/ijsa.12403
- Basch, J. M., Melchers, K. G., Kegelmann, J., & Lieb, L. (2020). Smile for the camera! The role of social presence and impression management in perceptions of technology-mediated interviews. *Journal of Managerial Psychology*, *35*(4), 285–299. https://doi.org/10.1108/JMP-09-2018-0398
- Bill, B., Melchers, K. G., Buehl, A.-K., & Wank, S. (2020). An Investigation of Situational and Dispositional Antecedents of Faking Intentions in Selection Interviews. *Frontiers in Psychology*, 11, 2034. https://doi.org/10.3389/fpsyg.2020.02034

- Bourdage, J. S., Roulin, N., & Tarraf, R. (2018). "I (might be) just that good": Honest and deceptive impression management in employment interviews. *Personnel Psychology*, 71(4), 597–632. https://doi.org/10.1111/peps.12285
- Canagasuriam, D., & Lukacik, E. (2024). ChatGPT, can you take my job interview? Examining artificial intelligence cheating in the asynchronous video interview. *International Journal of Selection and Assessment*, ijsa.12491. https://doi.org/10.1111/ijsa.12491
- Chapman, D. S., & Zweig, D. I. (2005). Developing a nomological network for interview structure: Antecedents and consequences of the structured selection interview. *Personnel Psychology*, *58*(3), 673–702. https://doi.org/10.1111/j.1744-6570.2005.00516.x
- Dunlop, P. D., Bourdage, J. S., de Vries, R. E., McNeill, I. M., Jorritsma, K., Orchard, M.,
 Austen, T., Baines, T., & Choe, W.-K. (2020). Liar! Liar! (when stakes are higher):
 Understanding how the overclaiming technique can be used to measure faking in
 personnel selection. *Journal of Applied Psychology*, 105(8), 784–799.
 https://doi.org/10.1037/apl0000463
- Dunlop, P. D., Holtrop, D., & Wee, S. (2022). How asynchronous video interviews are used in practice: A study of an Australian-based AVI vendor. *International Journal of Selection and Assessment*, 30(3), 448–455. https://doi.org/10.1111/ijsa.12372
- Gorman, C. A., Robinson, J., & Gamble, J. S. (2018). An investigation into the validity of asynchronous web-based video employment-interview ratings. *Consulting Psychology Journal: Practice and Research*, 70(2), 129–146. https://doi.org/10.1037/cpb0000102
- Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers A study to understand applicant perspective. *International*

- Journal of Hospitality Management, 36, 90–100. https://doi.org/10.1016/j.ijhm.2013.08.004
- Harwood, H., Roulin, N., & Iqbal, M. Z. (2024). "Anything you can do, I can do": Examining the use of ChatGPT in situational judgement tests for professional program admission.

 Journal of Vocational Behavior, 104013. https://doi.org/10.1016/j.jvb.2024.104013
- Hayes, A. F., & Scharkow, M. (2013). The Relative Trustworthiness of Inferential Tests of the
 Indirect Effect in Statistical Mediation Analysis: Does Method Really Matter?
 Psychological Science, 24(10), 1918–1927. https://doi.org/10.1177/0956797613480187
- Hickman, L., Bosch, N., Ng, V., Saef, R., Tay, L., & Woo, S. E. (2022). Automated Video Interview Personality Assessments: Reliability, Validity, and Generalizability
 Investigations. *Journal of Applied Psychology*, 107(8), 1323–1351.
 https://doi.org/10.1037/apl0000695
- Hickman, L., Liff, J., Rottman, C., & Calderwood, C. (2024). The Effects of the Training Sample Size, Ground Truth Reliability, and NLP Method on Language-Based Automatic Interview Scores' Psychometric Properties. *Organizational Research Methods*, 10944281241264027. https://doi.org/10.1177/10944281241264027
- Ho, J. L., Perossa, A., Fancett, R. S., & Powell, D. M. (2021). Examining the situational antecedents of interview faking behavior: A qualitative study. *International Journal of Selection and Assessment*, 29(3–4), 427–447. https://doi.org/10.1111/ijsa.12354
- Ho, J. L., Powell, D. M., & Stanley, D. J. (2021). The relation between deceptive impression management and employment interview ratings: A meta-analysis. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, 53(2), 164–174. https://doi.org/10.1037/cbs0000223

- Jang, M. E., & Lukasiewicz, T. (2023). Consistency Analysis of ChatGPT (arXiv:2303.06273). arXiv. http://arxiv.org/abs/2303.06273
- Jenatabadi, H. S. (2015). An Overview of Path Analysis: Mediation Analysis Concept in Structural Equation Modeling.
- Kelloway, E. K. (2015). *Using Mplus for structural equation modeling: A researcher's guide* (Second edition). SAGE.
- Korkmaz, S., Goksuluk, D., & Zararsiz, G. (2014). MVN: An R Package for Assessing

 Multivariate Normality. *The R Journal*, 6(2), 151. https://doi.org/10.32614/RJ-2014-031
- Langer, M., Demetriou, A., Arvanitidis, A., Vanderveken, S., & Hiemstra, A. M. F. (2024). A quasi-experimental investigation of differences between face-to-face and videoconference interviews in an actual selection process. *Applied Psychology*, apps.12558. https://doi.org/10.1111/apps.12558
- Langer, M., König, C. J., & Fitili, A. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection. *Computers in Human Behavior*, 81, 19–30. https://doi.org/10.1016/j.chb.2017.11.036
- Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, 25(4), 371–382. https://doi.org/10.1111/ijsa.12191
- Lee, Y., Berry, C. M., & Gonzalez-Mulé, E. (2019). The importance of being humble: A metaanalysis and incremental validity analysis of the relationship between honesty-humility and job performance. *Journal of Applied Psychology*, 104(12), 1535–1546. https://doi.org/10.1037/apl0000421

- 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. https://doi.org/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*, 92(6), 1638–1656. https://doi.org/10.1037/0021-9010.92.6.1638
- Liff, J., Mondragon, N., Gardner, C., Hartwell, C. J., & Bradshaw, A. (2024). Psychometric properties of automated video interview competency assessments. *Journal of Applied Psychology*. https://doi.org/10.1037/apl0001173
- Lukacik, E.-R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, 32(1), 100789.

 https://doi.org/10.1016/j.hrmr.2020.100789
- Marcus, B. (2009). 'Faking' From the Applicant's Perspective: A theory of self-presentation in personnel selection settings. *International Journal of Selection and Assessment*, 17(4), 417–430. https://doi.org/10.1111/j.1468-2389.2009.00483.x
- McCarthy, J., & Goffin, R. (2004). Measuring Job Interview Anxiety: Beyond Weak Knees and Sweaty Palms. *Personnel Psychology*, *57*(3), 607–637. https://doi.org/10.1111/j.1744-6570.2004.00002.x
- McCarthy, J. M., Van Iddekinge, C. H., & Campion, M. A. (2010). Are highly structured job interviews resistant to demographic similarity effects? *Personnel Psychology*, *63*(2), 325–359. https://doi.org/10.1111/j.1744-6570.2010.01172.x

- Melchers, K. G., Roulin, N., & Buehl, A. (2020). A review of applicant faking in selection interviews. *International Journal of Selection and Assessment*, 28(2), 123–142. https://doi.org/10.1111/ijsa.12280
- Moon, B., Law, S. J., Bourdage, J. S., Roulin, N., & Melchers, K. G. (2023). Exploring the role of interviewee cognitive capacities on impression management in face-to-face and virtual interviews. *International Journal of Selection and Assessment*, ijsa.12460. https://doi.org/10.1111/ijsa.12460
- Peck, J. A., & Levashina, J. (2017). Impression Management and Interview and Job Performance
 Ratings: A Meta-Analysis of Research Design with Tactics in Mind. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.00201
- Powell, D. M., Bourdage, J. S., & Bonaccio, S. (2021). Shake and Fake: The Role of Interview Anxiety in Deceptive Impression Management. *Journal of Business and Psychology*, 36(5), 829–840. https://doi.org/10.1007/s10869-020-09708-1
- Press Release Desk. (n.d.). HireVue hits 1 millionth interview milestone in Australia as demand for AI and Virtual interviewing grows. 2022. https://busycontinent.com/hirevue-hits-a-millionth-interview/#:~:text=On%20a%20global%20scale%2C%20HireVue,%25)%20were%20completed%20on%20mobile
- Rizi, M. S., & Roulin, N. (2023). Does media richness influence job applicants' experience in asynchronous video interviews? Examining social presence, impression management, anxiety, and performance. *International Journal of Selection and Assessment*, ijsa.12448. https://doi.org/10.1111/ijsa.12448

- Roulin, N., Bangerter, A., & Levashina, J. (2015). Honest and Deceptive Impression

 Management in the Employment Interview: Can It Be Detected and How Does It Impact

 Evaluations? *Personnel Psychology*, 68(2), 395–444. https://doi.org/10.1111/peps.12079
- Roulin, N., & Bourdage, J. S. (2017). Once an Impression Manager, Always an Impression Manager? Antecedents of Honest and Deceptive Impression Management Use and Variability across Multiple Job Interviews. *Frontiers in Psychology*, 8. https://doi.org/10.3389/fpsyg.2017.00029
- Roulin, N., Krings, F., & Binggeli, S. (2016). A dynamic model of applicant faking.

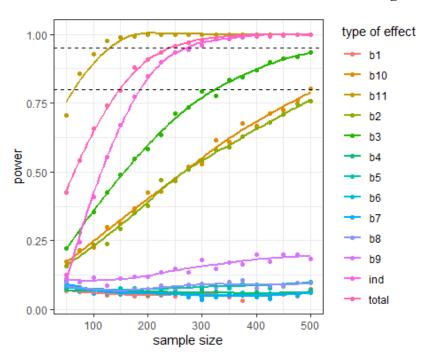
 *Organizational Psychology Review, 6(2), 145–170.

 https://doi.org/10.1177/2041386615580875
- Roulin, N., Pham, L. K. A., & Bourdage, J. S. (2023). Ready? Camera rolling... action!
 Examining interviewee training and practice opportunities in asynchronous video interviews. *Journal of Vocational Behavior*, 145, 103912.
 https://doi.org/10.1016/j.jvb.2023.103912
- Roulin, N., Wong, O., Langer, M., & Bourdage, J. S. (2023). Is more always better? How preparation time and re-recording opportunities impact fairness, anxiety, impression management, and performance in asynchronous video interviews. *European Journal of Work and Organizational Psychology*, 32(3), 333–345. https://doi.org/10.1080/1359432X.2022.2156862
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of range. *Journal of Applied Psychology*, 107(11), 2040–2068. https://doi.org/10.1037/apl0000994

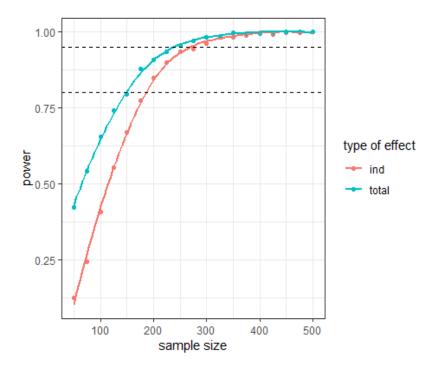
- Schneider, L., Powell, D. M., & Bonaccio, S. (2019). Does interview anxiety predict job performance and does it influence the predictive validity of interviews? *International Journal of Selection and Assessment*, 27(4), 328–336. https://doi.org/10.1111/ijsa.12263
- Spiller, T. R., Rabe, F., Ben-Zion, Z., Korem, N., Burrer, A., Homan, P., Harpaz-Rotem, I., & Duek, O. (2023). *Efficient and Accurate Transcription in Mental Health Research—A Tutorial on Using Whisper AI for Audio File Transcription*. https://doi.org/10.31219/osf.io/9fue8
- Wingate, T. G., Bourdage, J. S., & Steel, P. (2024). Evaluating interview criterion-related validity for distinct constructs: A meta-analysis. *International Journal of Selection and Assessment*, ijsa.12494. https://doi.org/10.1111/ijsa.12494
- Zhang, T., Koutsoumpis, A., Oostrom, J. K., Holtrop, D., Ghassemi, S., & Vries, R. E. D. (2024).
 Can Large Language Models Assess Personality from Asynchronous Video Interviews? A
 Comprehensive Evaluation of Validity, Reliability, Fairness, and Rating Patterns. *IEEE Transactions on Affective Computing*, 1–16.
 https://doi.org/10.1109/TAFFC.2024.3374875

Appendix A

Monte Carlo Simulation Figures



Note. all paths/ effects present.



Note. Only total and indirect effects present.

Appendix B

Job description – HR Generalist



About Western Oil Resources Group

Western Oil Resources Group is a leader in all things oil and gas.

The Role

The HR Generalist will support the business with human capital requirements. The HR Generalist will report to the Head of People and Talent. They will serve as a key business partner to develop, implement, and evaluate human resources and labour relations policies, programs and procedures and advise on human resources matters.

Key Responsibilities

- Plan, develop, and maintain HR policies and procedures for the organization.
- Assist in assessing and identifying potential talent including provide recommendations to the Head of People and Talent.
- Considers and conducts all work tasks in an ethical, honest, and fair manner.
- Collaborate closely with other key members of the People and Talent team.
- Lead new hire orientation.

About you

- Bachelor's degree or diploma/associates degree in human resources, psychology, business, or other related field.
- Excellent decision-making skills.
- Excellent communication skills.
- A keen eye for details.
- Consideration and ability to think and act with ethics and integrity.
- Proficiency with MS Office applications (Word, Excel, and Outlook).

Benefits/additional compensation for this role include:

- Full medical, dental and vision benefits
- Maternity / paternity policy
- 4 weeks paid vacation

- Competitive compensation
- Remote, work from anywhere
- Professional development opportunities, working closely with founders and senior leadership team

Seniority Level: Intermediate-Senior

Location: Remote (Canada & U.S.). We are a remote company and open to candidates from anywhere in North America.

We are proud to be an equal opportunity workplace committed to building a team culture that celebrates diversity and inclusion and as such does not discriminate on the basis of ethnicity, colour, religion, gender, national origins, age, sexual orientation, disability, or any other characteristic protected by applicable laws. Selection decisions are solely based on job related factors.

Appendix C

Interview Questions and BARS

Communication

Definition: The ability to effectively communicate a message to a target.

Could you tell us about a time when you had to communicate with key stakeholders, such as a manager you directly reported to or other leaders around you, and inform them of the importance of a new policy or program that you were trying to recommend/advocate for. What was the situation? How did you approach the stakeholders? And what was the outcome?

SCORING GUIDE: 5 1 3 The candidate... The candidate... The candidate... Provided only a vague Was able to describe a Was able to describe a or generic description specific instance where specific instance where of past interactions with they interacted with key they interacted with stakeholders stakeholders who had stakeholders who had The stakeholders were some power (e.g., a the power to influence not managers or supervisor, but not a top important policies (e.g., department heads, VPs, leaders. manager). Demonstrated a general or CEO). Did not specify how they communicated ability to communicate Demonstrated the about the policy or with the individuals, for ability to effectively program, or used a instance by providing communicate with the generic strategy without sound arguments to individuals, for instance adapting their explain the value of the by tailoring their arguments to their policy/program. arguments to their audience. Made a reasonablyaudience, and was able Made a relatively weak strong case for why to educate them on the they should support the case for why the importance of the stakeholders should policy/program, and policy/program. were able to obtain Made a compelling case support the policy/program, and some level of support. for why they should support the policy or was not able to completely convinced program, and ultimately convinced them to them.

formally do so.

have made.

Decision Making

Definition: The ability to effectively make decisions driven by the data and information available.

Could you tell us about a time when you had to make a difficult decision, such as selecting the benefits plan for your organization or choosing the best qualified applicant. What was the situation? How did you approach the decision? And what was the outcome?

SCORING GUIDE: 5 1 3 The candidate... The candidate... The candidate... Describes a scenario Outlines a decision that Outlines a decision that where they do not have was only somewhat was truly difficult to difficult to make (e.g., to make a decision make (e.g., picking picking a benefits plan between two benefits whatsoever (i.e., but there was clearly someone else made it). plans, one may be better one better option) for the employees but is Or outlines a decision Provides rationale for that was not difficult to more expensive, the their decision based on other is cheaper but make (e.g., a noncritical employee asks may offer less coverage data/information, but for time off that they they did not gather for employees) Provides rationale for were legally entitled to) extra information per se Does not provide any (i.e., using only what their decision based on rationale for their was initially provided to information/data decision. They do not gathered beyond the them). really consider Examines superficially basic information the potential costs or information/data at all. available. Does not examine the benefits of each Examines carefully the decision they could potential costs and potential costs and benefits of each have made. benefits of each decision they could decision they could

have made.

Attention to Detail

Definition: The ability to be thorough when completing a task by monitoring key components, examining even the "little" details.

Describe a project or task where you were given complex or technical data or information to analyze and interpret in order to make a recommendation. What was the situation? How did you go about processing that information? And what was the outcome?

SCORING GUIDE:							
1	3	5					
The candidate - Outlines a situation that did not require the need to be attentive to detail. - The information or data	The candidate - Outlines a situation that may require some need to be attentive to detail. - Explains in some detail	The candidate - Outlines a situation that clearly requires the need to be attentive to detail.					
that was required/ used in the situation is unclear and vague, leaving it unsure how they interpreted the information/data. - The outcome is: not described or was not affected by the persons attention to detail.	how they ensured the accuracy of the information or data (e.g., double-checking their interpretation or the data itself). However, some details may be vague/ unclear The outcome of the situation has limited success (e.g., there may still have been some errors in their recommendation). Or the recommendation described is of little importance.	 Clearly explains how they ensured the accuracy of the information or data (e.g., double-checking their interpretation and the data itself). The outcome of the situation is critical and successful (e.g., there are no errors in their recommendation, data is used effectively). 					

Communication

Definition: The ability to effectively communicate a message to a target.

You are attending a Zoom meeting with 30 managers and leaders of the company. While describing ideas for a new advertising campaign, Jason, the Vice-President of marketing makes several remarks that could been perceived as offensive by certain groups of employees. Nobody says anything about Jason's comments, probably because Jason is a senior leader in the company and outranks most attendees (you included). But you notice that a couple of individuals from that group are visibly upset. The next day, you are sending an e-mail to Jason to discuss recruiting a new employee for his team. How would you address Jason about his comments during the meeting?

SCORING GUIDE:			
1	3	5	
The candidate Does not describe to Jason that their remarks were offensive. Makes no mention in their e-mail to the executive, Jason is high on the company ladder and could get them fired. Describes getting angry/ passive aggressive with Jason, or purposely trying to humiliate the executive in their e-mail. Exclusively focuses on the new employee recruitment and nothing more to Jason in their e-mail.	The candidate - Attempts to "dance" around the fact that the remarks were offensive. - Describes being polite or professional (not both) in their e-mail. They may even describe being slightly hostile, or passive aggressive in their e-mail. - Explains why remarks made were offensive but with minimal detail. - Still includes communication about new employee recruitment in their e-mail	The candidate - States firmly that what the executive said was offensive to some employees (and could have been hurtful) - Describes a clear and concise c e-mail communication with Jason. They describe ways to ensure their email is professional and polite and does not convey angry or passive aggressiveness Clearly explains why the remarks made were offensive in detail so the executive can understand their impact Includes communication about the new employee recruitment and may even "weave" in being inclusive to help recruit new team members.	

Decision Making

Definition: The ability to effectively make decisions driven by the data and information available.

A team that you oversee must develop a leadership training program for the product management department. In addition to you, there are three other team members: two senior members who are very qualified and one junior member who is willing to work very hard but whose experience and capabilities in this area are very limited. This is a very important project, and your boss, the head of People and Talent, has made it clear that the project must be completed very quickly. How would you allocate work between the team members? To what extent would you involve the junior team member?

SCORING GUIDE:		
1	3	5
The candidate	The candidate	The candidate
 Does not involve the junior at all or gives them unimportant tasks (e.g., make copies). Puts most of the work on the senior members of the team. And, avoids taking any tasks for themselves. Or takes all the important work for themselves. 	 Involves the junior member in some of the tasks but does not provide them with the necessary support. Only pairs the member with senior members (e.g., "pawning" them off) Only takes <i>only</i> minor tasks themselves. 	 Challenges the junior member of the team to further develop. Pairs them with a senior member (or themselves) to help guide their development (i.e., some form of mentoring while ensuring that they can perform the tasks). Takes on <i>some</i> important works tasks themselves to ensure the team can efficiently meet the deadline.

Attention to Detail

Definition: The ability to be through when completing a task by monitoring key components, examining even the "little" details.

You have to design a new "work from home" policy for the company. It is an important project, which will impact many employees. You have been working on it for two months now, and you have one month left to present your final report. You have the next steps clearly laid out, and you believe that you are on track to finish it just in time if you work full-time on it. However, your supervisor informs you the timeline has changed, and the report is now due in 2 weeks. After that, you will be assigned to hiring a new engineer with specialized expertise for a project with one of the company's largest clients. How would you adjust your plans and organize your work to meet the revised deadline while starting to think about the new hiring task?

SCORING GUIDE:		
1	3	5
The candidate Only focuses on one task or the other leaving one of the tasks worse off. Fails to describe in detail the steps they would take. Describes a scenario where they offload the work to another individual.	The candidate Describes in some detail their plan of action. Mainly focuses on the "bigger" elements, which can be detrimental to some important details (e.g., cut corners to finalize the report on time). Some focus on smaller details, but they are surface level details. May touch on the hiring the engineer but nothing overly detailed	The candidate Is able to describe in detail the plan of action to complete the project on time (i.e., in 2 weeks) and starts the new task. Considers bigger picture aspects of the situation, as well as carefully considering small aspects of what would be left to do (e.g., elements of the report that could be streamlined without hurting quality). Mentions double-checking of the already clearly laid out steps. Starts initial planning of hiring the engineer.

Appendix D

Survey Items

Deception Debrief & Survey Instructions

At the beginning of the study, and as part of the recruitment message you were told that your performance in the interview you just completed would determine if you were recruited for a second high paying research task. This was done to enhance the realism of the interview, but in reality, everyone will be invited to complete the second high paying research task (and have an opportunity to earn a bonus in the second task), provided that you took the study seriously.

We apologize if this caused any discomfort or frustration, <u>this will not disadvantage you in any way.</u> If you have any immediate questions or feedback, we welcome you to email the principal investigator (<u>Harley.harwood@smu.ca</u>). If you are experiencing an immediate high level of stress or discomfort, please use the resources below.

Canada: https://www.canada.ca/en/public-health/services/mental-health-services/mental-

<u>health-get-help.html</u>

US: https://988lifeline.org/talk-to-someone-now/

Now that you have finished the interview portion, we ask you to complete the survey portion of the study. Please respond completely honestly to the questions below. Importantly, your answers to these questions will have <u>no</u> impact on your compensation for completing the present study (i.e., compensation is only contingent on response quality and completion). Your response will also <u>not</u> disqualify you for the second high-paying HR task.

Thank you so much for your participation and understanding.

Honest and Deceptive Impression Management Items

Scale: (1 = to no extent to 5 = to a very great extent)

Honest Defensive

- 1. I shared my past regrets about how I handled certain situations, and how I would improve in the future.
- 2. I gave an honest account of why I lacked control over past negative events that came up during the interview.
- 3. I recounted steps I had taken to prevent the recurrence of negative past events or occurrences in my past.
- 4. I gave reasons why I felt I benefited positively from a negative event I was responsible for.

Honest Ingratiation

- 1. I tried to find out the values or opinions of the hiring organization and I shared in common, and was vocal about these.
- 2. I talked about the values the hiring organization and I shared.
- 3. When the hiring organization had views that I shared, I focused on incorporating these into my answers.
- 4. When I agreed with the hiring organization's opinions or points, I made sure to talk about it.
- 5. I did my best to convey the values, attitudes, or beliefs that I felt me and the hiring organization shared.
- 6. I found out about values and goals that I shared with the hiring organization, and made sure to emphasize them.
- 7. I discussed interests that I shared in common with the hiring organization.
- 8. I complimented the hiring organization on accomplishments or qualities that I found impressive.

Honest Self Promotion

- 1. I made sure to talk about my job credentials.
- 2. I explained how my qualifications were well-suited for the position.
- 3. I demonstrated in genuine ways that I was a good performer in my previous job.
- 4. I talked about all the responsibilities I had on my previous jobs.
- 5. I made sure I emphasized my skills and abilities.
- 6. I described my skills and abilities in an attractive way.
- 7. I talked about how my previous work experiences were relevant to the position.
- 8. I brought up my past experience with other well-known previous employers to highlight my competence.
- 9. I talked about how I felt I could be a valuable addition to the organization.
- 10. I talked about the accomplishments I'd had at my previous job.
- 11. I made sure to recount my areas of expertise.

- 12. I looked for opportunities to discuss my success at previous jobs.
- 13. I promoted the skills and abilities that I thought most relevant to the position.
- 14. I brought up my past work experience to emphasize my competence.

Deceptive Image Protection

- 1. When asked directly, I did not mention my true reason for quitting previous jobs.
- 2. When asked directly, I did not mention some problems that I had in past jobs.
- 3. I tried to avoid discussion of job tasks that I may not be able to do.
- 4. I clearly separated myself from my past work experiences that would reflect poorly on me.

Deceptive Ingratiation

- 1. I tried to adjust my answers to the organization's values and beliefs.
- 2. I tried to incorporate the organization's views in my answers as my own.
- 3. I tried to show that I shared the organization's views and ideas even if I did not.
- 4. I tried to appear similar to the organization in terms of values, attitudes, or beliefs.
- 5. I exaggerated my positive comments about the organization.
- 6. I complimented the organization on something, however insignificant it may actually be to me.

Deceptive Image Creation (Slight)

- 1. I exaggerated my responsibilities on my previous jobs.
- 2. I exaggerated the impact of my performance in my past jobs.
- 3. I tried to use information about the company to make my answers sound like I was a better fit than I actually was.
- 4. During the interview, I distorted my answers to emphasize what the organization was looking for.
- 5. I distorted my answers based on the information about the job I obtained during the interview.
- 6. I enhanced my fit with the job in terms of attitudes, values, or beliefs.
- 7. I inflated the fit between my values and goals and values and goals of the organization.

Deceptive Image Creation (Extensive)

- 1. I told fictional stories prepared in advance of the interview to best present my credentials.
- 2. I made up stories about my work experiences that were well developed and logical.
- 3. I constructed fictional stories to explain the gaps in my work experiences.
- 4. I used made-up stories for most questions.
- 5. I misrepresented the description of an event.
- 6. I stretched the truth to give a good answer.
- 7. I invented some work situations or accomplishments that did not really occur.
- 8. When I did not have a good answer, I borrowed work experiences of other people and made them sound like my own.

9. I used other people's experiences to create answers when I did not have good experiences of my own.

Honesty per Competency Questions

The interview questions were designed to measure certain competencies. Questions 1 & 4 were designed to measure your ability to communicate. Thinking about how you answered those two questions *specifically* please rate how honest/dishonest you were on the scale. **0 = completely dishonest**, everything was fabricated, 100 = completely honest, everything was true.

```
<sliding scale from 0 to 100>
```

The interview questions were designed to measure certain competencies. Questions 2 & 5 were designed to measure your ability to make decisions. Thinking about how you answered those two questions *specifically* please rate how honest/ dishonest you were on the scale. **0 = completely dishonest**, **everything was fabricated**, **100 = completely honest**, **everything was true**.

```
<sliding scale from 0 to 100>
```

The interview questions were designed to measure certain competencies. Questions 3 & 6 were designed to measure your attention to detail. Thinking about how you answered those two questions *specifically* please rate how honest/ dishonest you were on the scale. **0 = completely dishonest**, everything was fabricated, **100 = completely honest**, everything was true.

<sliding scale from 0 to 100>

Seriousness and Attention Checks

Scale: (Strongly disagree = 1, neither agree nor disagree = 3, Strongly agree = 5)

Please indicate "To some extent" for this question.

I answered the survey questions seriously.

I eat cement everyday.

If you were to arrange the list of movies below in alphabetical order, which movie title would come first?

- a) Speed
- b) John Wick
- c) The Matrix
- d) Constantine
- e) Point Break

Interview Anxiety

(answered on a 1-5 "not at all" to 'extremely")

During my most recent interview...

- 1. I felt anxious
- 2. I felt calm (R)
- 3. I felt very tense
- 4. I had trouble controlling my nerves
- 5. I wished my interview did not bother me so much
- 6. I found myself worrying a lot

Motivation to Perform (Adapted from Arvey et al., 1990)

Scale: (1 = Strongly disagree, 5 = Strongly agree) * = Reverse coded items

- 1. Doing well on the interview is important to me.
- 2. I wanted to do well on the interview.
- 3. I tried my best in the interview.
- 4. I tried to do the very best I could to in the interview.
- 5. During the interview, I concentrated and tried to do well.
- 6. I want to be among the top performers in the interview.
- 7. I pushed myself to work hard in the interview.
- 8. I was extremely motivated to do well in the interview.
- 9. I just didn't care how I did in the interview.*
- 10. I didn't put much effort into the interview.*

Demographics

What is your age? (In whole numbers only) (Text enter numbers only)

How do you self-identify?

- a) Male
- b) Female
- c) Non-binary
- d) No Options represent me: (Please specify)

With what racial or ethnic group do you identify?

- a) White/ Caucasian
- b) Black
- c) Hispanic
- d) East Asian (e.g., Chinese, Japanese, Korean)
- e) South Asian (e.g., Indian, Pakistani, Sri Lankan)
- f) Middle Eastern
- g) Aboriginal/Indigenous
- h) Mixed
- i) No Options represent me: (Please specify)

How many years have you worked before? (Text enter numbers only)

What is the highest level of education you have completed?

- a) Less than high school
- b) High school diploma
- c) Associate degree/ diploma
- d) College/university degree
- e) Master's degree
- f) Doctoral degree

How many years of HR experience do you have? (In numbers only with a maximum of one decimal place).

(Text enter numbers only)

How many interviews they have completed in your career? (Text enter numbers only)

How many AVIs they have completed in you career? (excluding AVIs for research) (Text enter numbers only)

Appendix E

Study Instructions

You will be asked to complete <u>two</u> main HR tasks. The first involves reviewing an article from Havard Business Review and the second involves watching several candidates recorded interviews. The top 10% performers will receive a <u>bonus (6 GBP/7 USD)</u>. We ask that you carefully pay attention to all instructions.

Remember that you have been selected to work at Western Oil Resources Group (WORG) as an HR Generalist. WORG is a fully remote organization that offers engineering and design consulting services. Despite being remote we still value teamwork and collaboration.

Next you will review some materials before you move on through the tasks.

Appendix F

Candidate Ranking Task

E-mail From: Kathelene Wilson, Head of People and Talent @ WORG <kathlene@WORG.ca>

Subject: Candidate Ranking and Recommendations

Hi there again,

Thanks again for reading that article and sending the summary and policy my way! What a time save.

I am reaching out to ask that you review the video-recorded interview of our top four candidates for the engineering team. Don't worry about their technical competence, we will leave that to the team leads to rate and worry about. Please review the job description, candidate interview recordings and fill out the rating forms and send them back to me.

What I need from you is to watch the candidates' interviews do the following things:

- Rate and rank each of the four candidates.
- Write a clear summary of the four candidates and a recommendation of the best candidate.

If you could get that done for me in the next 15 minutes that would be excellent.

All the best,

Kathlene Wilson Head of People and Talent WORG "Lassoing Energy Challenges, One Well at a Time"



Job description to Review - Project Manager



About Western Oil Resources Group

Western Oil Resources Group is a leader in all things oil and gas.

The Role

The project manager is a member of the Technology & Innovation department. This individual ensures that cost recovery through **Scientific Research and Experimental Development** (**SRED**) claims are executed efficiently and effectively while maximizing the benefit to WORG and their clients. It also involves continuous improvement of the SRED processes for capturing technical and financial project information and work towards achieving successful audit outcomes.

Responsibilities:

- Estimate SRED value and risks for technology projects/developments
- Work with business units throughout the organization to evaluate new and continuing projects
- Coordinate, monitor and control SRED claims
- Oversee the collection/documentation of technical and cost records of projects for audit purposes
- Support the development of procedures and standards for year-round SRED claim development
- Support other grant-based funding opportunities

Qualifications:

- Excellent written and verbal communication skills
- Excellent interpersonal skills
- Highly motivated, self-managed and able to deliver on time
- Proven analytical and problem-solving skills
- Ability to handle multiple projects simultaneously
- A Degree in Engineering, Business/ Project management, or Science Degree (ideally related to oil and gas)
- Familiarity with oil sands production, Six Sigma, and project management are assets

Benefits/additional compensation for this role include:

- Full medical, dental and vision benefits
- Maternity / paternity policy
- 4 weeks paid vacation
- Competitive compensation
- Remote, work from anywhere
- Professional development opportunities, working closely with founders and senior leadership team

Seniority Level: Intermediate-Senior

Location: Remote (Canada & U.S.) with visits to sites. We are a remote company and open to candidates from anywhere in North America.

We are proud to be an equal opportunity workplace committed to building a team culture that celebrates diversity and inclusion and as such does not discriminate on the basis of ethnicity, colour, religion, gender, national origins, age, sexual orientation, disability, or any other characteristic protected by applicable laws. Selection decisions are solely based on job related factors.

Participants will then watch 2 or 3 video recordings from each of the 4 candidates.

Candidate Ratings

Now that you have watched each of the candidates, please rate them on the following scales (where 1 is poorest performance and 5 is the best performance).

Q1 Ratings for all candidates:

Initiative:

1-5 rating

Q2 Ratings for all candidates:

Persuasion:

1-5 rating

Candidate Ranking

Now please rank the candidates with the best candidate at the top of the list and worst candidate at the bottom of the list.

<ranking question on Qualtrics, where participants drag participant names in the order>

Summary and Recommendation:

Please now write your summary of the four candidates and write a recommendation for your top candidate for your manager Kathelene. Provide your reasons why you picked your top candidate.

<Text box for participant to write their summary and recommendation>

Summary and Recommendation Task

1 3	5
3	5
- Choppy and hard to follow - Major flaws with punctuation, spelling, and grammar - Is not appropriately structured (e.g., just a list of bullet points) - Language is informal - Choppy in some areas, but still relatively easy to - Had minimal errors with punctuation, spelling, and grammar - Is of bullet points) - Some language may be informal - Is of bullet points) - Is not appropriately punctuation, spelling, and grammar - Is of bullet points) - Some language may be informal	ng is Clearly logical and made sense (i.e., had a flow) Effectively Had no flaws in punctuation, grammar, and spelling Is appropriately structured (i.e., looks like a proper email) Language is formal, and appropriate given a workplace setting

Attribute: Message Clarity

1 3 5

The participant...

- Does not recommend a candidate whatsoever
- Or provides no reason for why they recommended the candidate

The participant...
- Recommends a candidate

- Provides on a simple reason for recommending a candidate (e.g., I rated them highest) The participant...

- Provides a unbiased reason as to why they selected the candidate
- Clearly states which candidate they recommend

Quiz Questions

- 1. What did the first candidate mention doing to prevent group member from procrastinating?
 - a. Sets strict individual deadlines for each group member
 - b. Provides frequent updates on each member's progress
 - c. Immediately sets up a group discussion and assigns roles
 - d. Completes most of the work personally
- 2. According to the fourth candidate, what is the expected outcome after performing the ordering process multiple times?
 - a. Mastery of the process by the third attempt
 - b. Increased difficulty with each repetition
 - c. Decreased need for planning and organization
 - d. Decreased reliance on notes and books
- 3. Which past job did the second candidate mention that required minimal planning?
 - a. Organizing poker tournaments
 - b. Driving for Uber
 - c. Working with manufacturers
 - d. Working on group projects
- 4. What is the initial approach the third candidate takes when faced with a complex task that requires significant planning?
 - a. Analyzing the cost-effectiveness of various approaches
 - b. Considering customer feedback and preferences
 - c. Starting with a macro-level brainstorm session
 - d. Performing a risk assessment to identify potential issues
- 5. Which of the candidates suggested they've received overwhelmingly positive feedback almost always?
 - a. Both candidates 1 and 2
 - b. Candidate 2 only
 - c. Candidate 3 only
 - d. Both candidates 2 and 4
- 6. What is meany by "tolerances" in the context of the manufacturing environment described by the third candidate?
 - a. The range of acceptable measurements for a give object
 - b. The number of defects allowed in a batch parts
 - c. The time frame within which a project must be completed
 - d. The budget allocated for quality control measures

Appendix G

Study Materials for Task 1: Article Task

Email From: Kathelene Wilson, Head of People and Talent @ WORG <kathlene@WORG.ca>

Subject: Article Task and Policy Reminder

Hi there,

I am reaching out to ask that you read the attached Harvard Business Review article on 1-to-1 meetings. We want to roll out some new policies and procedures to help managers across the company with their meetings (including for us and our department). A colleague of mine had recently forwarded this article by Dr. Steven Rogelberg (a top organizational scientist in the area of meetings research and distinguished Chancellor's Professor at UNC Charlotte).

What I need from you is to read through the article do the following things:

- Write me a concise summary of the article so I can understand the key take aways (sorry I don't have time to get through it today).
- Write a policies and procedures outline for holding 1-on-1 meetings at WORG. I will review it, and then send it out to all managers in all departments. So, it needs to be generally applicable to all teams/groups and ready to roll out.

If you could get that done for me in the next 15 minutes that would be excellent.

Policy Reminder: WORG does not tolerate the use of generative AI in your written work. We value your voice, and your effort. Please do not use generative AI in your work.

All the best,

Kathlene Wilson Head of People and Talent WORG "Lassoing Energy Challenges, One Well at a Time"



HBR Article and Questions

"Make the Most of Your One-on-One Meetings"12

by Steven G. Rogelberg

Turnover was high on Bill's team – higher, in fact, than on most other teams at his company. Although Bill thought of himself as a good manager, exit interviews with his departing team members suggested that they hadn't felt meaningfully engaged or fully supported in their roles and had tended to step on one another's toes with their assignments.

What, exactly, was Bill doing wrong? One area stood out when I spoke with him and his team: He held fewer regular one-on-one (1:1) meetings with his direct reports than his peers at the company. When he did meet with team members individually, the subject tended to be a critical issue he needed help with rather than their work or development.

Bill is a composite of managers I've worked with and studied, and clearly had a blind spot when it came to 1:1s. Such blind spots are not uncommon. Of 250 direct reports I surveyed, nearly half rated their 1:1 experiences as suboptimal. That's hardly surprising, given that few organizations provide strong guidance or training for managers about when and how to meet individually with their employees. But my research shows that managers who don't invest in such conversations – who view them as a burden, hold them too infrequently, or manage them poorly – risk leaving their team members disconnected, both functionally and emotionally.

I've been studying teams, leadership, engagement, and meetings at work for decades, and in the past three years I've set out specifically to learn what makes 1:1s work best by doing three studies: a global survey of 1,000 knowledge workers, a U.S. survey of 250 people who either lead or participate in 1:1s, and interviews with nearly 50 top leaders at various *Fortune* 100 companies. I've discovered that although no one-size-fits-all approach exists, there are some useful guidelines for managers. Most important is that the manager should consider the meeting a focused space for the direct report and make that explicit. The meeting should be dominated by topics relating to the needs, concerns, and hopes of the employee, who should take an active role in presenting them. As the manager, your responsibilities are to ensure that the meetings occur, actively facilitate them, encourage genuine conversation, ask good questions, offer support, and help each team member get what's needed for optimal short-term performance and long-term growth.

Before the Meetings

Setting up your 1:1s should entail more than dropping invites onto your team members' calendars. You should lay the groundwork for your conversations and plan the logistics to best fit each report's unique needs.

Communicate the initiative or your reboot of the initiative.

¹² This is a slightly abbreviated version of the original article published in Harvard Business Review.

Whether or not the practice of holding 1:1s is new to your team, announce it at a team meeting so that everyone gets the message at the same time and no one feels singled out. Tie the meetings to your organization's values and to your personal values. Also stress that these conversations are not meant to signal dissatisfaction with your team's work and are not about micromanaging; rather, they are opportunities for you and each member to get to know each other better, learn about challenges, discuss careers, and for you to give help when it's needed. This is also a good moment to tell your team members what you need from them to make the meetings successful: They should drive the agenda with key priorities, be curious, be actively engaged, communicate candidly, think deeply about problems and solutions, and be willing to ask for help and act on feedback.

Determine cadence.

My research suggests that you should adopt one of three plans for the frequency of 1:1s: (1) You meet with each of your team members *once a week* for 30 minutes or so. In my surveys, employees, regardless of job level, rated this approach the most desirable; it also correlated with the highest levels of engagement. (2) In the second-highest-rated plan you meet *every other week* for 45 to 60 minutes. (3) In a *hybrid* plan you meet with some team members weekly and others every two weeks. Whichever you choose, aim to spend roughly equivalent amounts of time with employees over the course of a month so that all team members get the same support from you. To determine the right cadence, consider:

- *Manager tenure*. If *you* are new to the team, weekly meetings are ideal for establishing relationships and alignment.
- *Remote or in person*. If your team is remote, weekly meetings can help counter a lack of spontaneous face-to-face contact.
- Team member preference. Finally, give your employees a voice in the decision.

I've seen some managers, mostly senior leaders, opt for three or four weeks between 1:1s, but investing only 60 minutes with each team member every month makes building a trusting relationship difficult. And because more-recent events are easier to recall, the longer time lapse also means that you're less likely to discuss any issues that arose several weeks prior. These meetings are most effective when you can build momentum around specific areas of the direct report's activities and growth. A monthly cadence makes that more challenging. But if your team members are seasoned and have worked with you a long time, and you are readily available for impromptu conversations, this cadence can work and is preferable to nothing. However, employees rated this option as least desirable, and it was associated with smaller gains in engagement.

Finally, avoid canceling 1:1s, which can hamper your team members' progress and make them feel that they are low on your priority list. This was one of Bill's problems: He readily canceled these meetings if he got busy. That sometimes demoralized his team members; they also found themselves duplicating efforts because they hadn't had a chance to coordinate their work through Bill. If you must cancel, reschedule the meeting right away, ideally for the same week. Another option is to reduce the length of the meeting: Some time together is better than none at all.

Set a location.

In my research, employees rated virtual 1:1s as slightly less desirable than those held in person, but they rated the ultimate value of the meetings similarly regardless of which form they'd taken. If you can meet in person, choose a location where you and your employee will feel at ease, present, and free of distractions. In my surveys the most highly rated location was the manager's office or a conference room; the lowest was the direct report's office. Support for outside locations, such as coffee shops, was uneven, so don't assume that everyone would welcome them. Talk to your team members in advance to gauge where they feel most comfortable.

Create an agenda.

Many managers assume that 1:1s are too informal to require an agenda, but my research shows that having one is a strong predictor of the effectiveness of the meeting, whether it was created in advance (which is ideal) or at the meeting itself (if necessary). Even more critical, though, is the employee's involvement in the agenda's creation: Both direct reports and managers rated meetings most highly when the reports contributed to or established the agenda themselves. Bill's habit of organizing his 1:1s around his own priorities and needs meant that his team members' concerns were usually relegated to the end of the meeting – and often went unaddressed if time ran out.

Alternatively, some managers create the agenda from broad questions, such as: What would you like to talk about today? How are things going with you and your team? What are your current priorities, and are there any problems or concerns you would like to talk through? Etc.

At the Meetings

Once you've prepared for a meeting, a fruitful discussion will depend on your ability to create a comfortable setting for your employee. A valuable 1:1 addresses both practical and personal needs — to feel respected, heard, valued, trusted, and included. To ensure that a meeting does so:

Set the tone.

First, be present. Turn off email alerts, put your phone away, and silence text notifications. Remind yourself as the meeting begins that it is fundamentally about your employee's needs, performance, and engagement.

As you go into the meeting, check your emotional state. Research shows that the mood you bring to a meeting has a <u>contagion effect</u>, so start out with energy and optimism. Reiterate your goals and hopes for the meeting and then move to some non-work-related topics, rapport building, wins, or appreciation to generate momentum and foster feelings of psychological safety. One problem for Bill was that he viewed 1:1s as merely another task on his already long list. That affected how he facilitated (or failed to), listened, collaborated, and engaged.

Listen more than you talk.

The biggest predictor of a 1:1's success, according to my research, is the employee's active participation as measured by the amount of time that person talks during the meeting. The ideal is anywhere from 50% to 90%. The agenda will have some influence on that, but you as the manager should carefully avoid talking more than your employee does.

In addition, listen actively to fully understand your direct report before you speak yourself. Display genuine interest without judgment and acknowledge the employee's viewpoint even if you disagree with it. Ask questions that clarify and constructively challenge that viewpoint. Encourage your team member to provide thoughts and potential solutions to problems. Stay vigilant about your body language and reactions to ensure that you're creating a welcoming and safe space.

Be flexible.

As you work through your established agenda, allow the conversation to move organically as needed to provide value. Focus on the items that are most critical. If some items go unaddressed, move them to the following 1:1. Let your employee know at the outset that real-time changes can be made to the agenda if a critical item emerges. Also, to best connect with each direct report, consider that person's preferences regarding communication, collaboration, and adjust your leadership approach accordingly. That will increase engagement, inclusion, deepen the relationship, and create trust.

End well.

Clarify takeaways and action items for both parties, including how you will support next steps. When both the manager and the employee document these, chances are better that the actions will be carried out. It also builds continuity between meetings and allows for needed follow-up. After Bill implemented this change, he was reminded that his 1:1s were not mere transactions to get through but, rather, represented employees' evolving stories – something to be nurtured and developed over time. Finally, show gratitude and appreciation for your direct report's time – and start and stop on schedule to demonstrate those feelings.

Steven G. Rogelberg is the Chancellor's Professor at the University of North Carolina Charlotte for distinguished national, international, and interdisciplinary contributions. He is the author of *Glad We Met: The Art and Science of 1:1 Meetings* (Oxford University Press, 2024) and *The Surprising Science of Meetings: How You Can Lead Your Team to Peak Performance* (OUP, 2019) and. He writes and speaks about leadership, teams, meetings, and engagement.

Article Summary & Policy Writing task:

Please now write a summary of the article for your manager Kathelene. Note that your manager is busy, your summary can not exceed 3000 characters but should still be appropriately structured.

<Box for participants to enter a summary>

Please now write policy guidelines for 1:1 meetings at WORG. These guidelines should be clear and easily understood for any department. You should also include some example questions that could be included in a meeting agenda (note. you should go above and beyond what was in the article, and there is likely to be some overlap with the summary).

<Box for participants to enter a summary>

Rubric for Summary Answers (to be used by Research Assistants to rate performance):

Summary Task

Attribute: Technical Communication		
1	3	5
The writing is - Illegible or incoherent - Choppy and hard to follow - Major flaws with punctuation, spelling, and grammar - Is not appropriately structured (e.g., just a list of bullet points) - Language is informal	The writing is - Generally made sense - Choppy in some areas, but still relatively easy to follow - Had minimal errors with punctuation, spelling, and grammar - Some language may be informal	The writing is - Clearly logical and made sense (i.e., had a flow) - Effectively - Had no flaws in punctuation, grammar, and spelling - Is appropriately structured (i.e., looks like a proper email) - Language is formal, and appropriate given a workplace setting
Attribute: Message Relevance (Decision of the Control of the Contr		
1	3	5
The writing - Missed the major points (less than 4) of the article - Includes several non- relevant topics	The writing - Communicates most (6-8) major points from the article (with 4-5 points = "3" score and 9-10 points = "4" score) - May include a few (1-2) irrelevant points (e.g., Points about Steven Rogelberg)	The writing - Summarized most of all (i.e., 11-12) major points in the article - Does not include any irrelevant points

Key Points in Article:

- 1. Schedule 1:1 weekly, for 30 minutes
- 2. Work with your employee to create the agenda
- 3. Come to the meeting with a positive tone
- 4. Remove barriers for team members
- 5. Let team members speak more than you do
- 6. 1:1's can be a good opportunity to get to know your subordinates
- 7. Make sure you clarify and create action items for everyone
- 8. Its okay if you don't cover everything, but move it to the next meeting
- 9. Add your perspective when appropriate
- 10. Be present during the meeting, silence alerts and focus on the 1:1
- 11. Make sure to communicate the why to employees; ensure they understand the purpose of the 1:1's
- 12. Find a time that works best for your subordinate

Policy Creation Task

Attribute: Technical Communication 5 The writing is... The writing is... The writing is... Illegible or incoherent Generally made sense Clearly logical and made Choppy and hard to follow Choppy in some areas, but sense (i.e., had a flow) Major flaws with still relatively easy to Effectively punctuation, spelling, and follow Had no flaws in Had minimal errors with punctuation, grammar, and grammar Is not structured like a spelling punctuation, spelling, and policy (e.g., exclusively a grammar Structured like a policy, list with no explanation or Has some preamble including formal language preamble, or headings/ Clearly structured making it resemble a formal policy title) Language is informal Attribute: Policy Relevance (Decision Making) 3 5 The policy... The policy... The policy... Does not suggest how Gives a general outline, Clearly includes relevant managers should conduct a but lacks details steps for managers to conduct a 1:1 meeting 1:1 meeting Includes only one reason Does not include any why 1:1s are important Includes a couple reasons reasons why 1:1s are May include an example, why 1:1s are important but it is limited Includes a few examples in important Includes no examples in Is generic, and largely their policy that clearly the policy mimics the article outline recommendations Is extremely generic and Extends on the article and limited, ignoring the fits the remote setting setting of WORG

Example of a "5/5" Policy:

As part of a new initiative at WORG we want to support our team leads with 1:1 meeting guideline. At WORG we value teamwork and mentorship. We recommend that you talk to your team about why we hold 1:1s. These reasons include valuing development, teamwork and collaboration. It is important to communicate these reasons to your team, so everyone understands the purpose of the meetings. Also, you need to fit the 1:1s with team members needs, but we recommend that team leaders schedule weekly 1:1s with all members for 30 minutes. Both managers and subordinates should make sure they take notes and create the agendas collaboratively. This can be achieved by creating a collaborative document that both the team lead and employee can drop key points and questions into to cover each meeting. Team leads should also be conscious of both the work aspects and the personal aspects and take 5 minutes of each 1:1 to chat about personal aspects, which can be as simple as asking about their favorite show, podcast, etc. (we recommend staying away from topics that are too personal if employees are not comfortable sharing). During the meeting as a team lead you need to attentive, let your team member speak, give them the floor, but also make sure you are actively listening. To ensure they feel heard you may want to repeat what the employee is saying to you, ask clarifying questions and brainstorming solutions to their problems collaboratively. We also want to emphasize that all members take these 1:1s seriously and silence any and all notification before and during these 1:1s to maximize their effectiveness and ensure employees feel heard and important. All managers need to reach out to their employees to find times that work for them, and all one-on-ones will be held virtual, but we still recommending sharing a cup of coffee (or tea) with your team members to enhance the meeting.

Questions:

Correct answers bolded

Instructions:

Please select the correct answer based on the article you have just finished reading.

- 1. Which of the following was cited as the most preferred options for 1:1 meetings?
 - a. Every other week for a half an hour
 - b. Every week for a full hour
 - c. Every other week for an hour
 - d. Every week for a half hour
- 2. What percentage of direct reports rated their one-on-one experiences as suboptimal?
 - a. One-quarter
 - b. One-third
 - c. Nearly half
 - d. Almost two-thirds
- 3. What is suggested as a way to gauge the right frequency for one-on-one meetings?
 - a. Manager's preference
 - b. Team member experience
 - c. Company policy
 - d. Employee availability
- 4. What was one of the problems Bill encountered regarding his one-on-one meetings?
 - a. Lack of team engagement
 - b. Over-communication
 - c. Micromanagement
 - d. Frequent rescheduling
- 5. Where did employees rate virtual one-on-one meetings compared to in-person meetings?
 - a. More desirable
 - b. Equally desirable
 - c. Slightly less desirable
 - d. Significantly less desirable
- 6. According to the article, what is the biggest factor for a one-on-one meeting's success?
 - a. Manager's active participation
 - b. Employee's active participation
 - c. Length of the meeting
 - d. Location of the meeting

Appendix H

ChatGPT Prompts Used

Interview Prompts:

"q1": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"Could you tell us about a time when you had to communicate with key stakeholders, such as a manager you directly reported to or other leaders around you, and inform them of the importance of a new policy or program that you were trying to recommend/advocate for. What was the situation? How did you approach the stakeholders? And what was the outcome?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

- 1 = The candidate... Provided only a vague or generic description of past interactions with stakeholders, The stakeholders were not managers or leaders, Did not specify how they communicated about the policy or program, or used a generic strategy without adapting their arguments to their audience, Made a relatively weak case for why the stakeholders should support the policy/program, and was not able to completely convinced them.
- 3 = The candidate... Was able to describe a specific instance where they interacted with stakeholders who had some power (e.g., a supervisor, but not a top manager), Demonstrated a general ability to communicate with the individuals, for instance by providing sound arguments to explain the value of the policy/program, Made a reasonably-strong case for why they should support the policy/program, and were able to obtain some level of support.
- 5 = The candidate... Was able to describe a specific instance where they interacted with key stakeholders who had the power to influence important policies (e.g., department heads, VPs, or CEO), Demonstrated the ability to effectively communicate with the individuals, for instance by tailoring their arguments to their audience, and was able to educate them on the importance of the policy/program, Made a compelling case for why they should support the policy or program, and ultimately convinced them to formally do so.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

Rating=			
The transcrip	ot is as	follows:	• • • • •

"q2": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"Could you tell us about a time when you had to make a difficult decision, such as selecting the benefits plan for your organization or choosing the best qualified applicant. What was the situation? How did you approach the decision? And what was the outcome?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

1 = The candidate... Describes a scenario where they do not have to make a decision whatsoever (i.e., someone else made it). Or outlines a decision that was not difficult to make (e.g., a non-critical employee asks for time off that they were legally entitled to). Does not provide any rationale for their decision. They do not really consider information/data at all.

Does not examine the potential costs and benefits of each decision they could have made.

- 3 = The candidate... Outlines a decision that was only somewhat difficult to make (e.g., picking a benefits plan but there was clearly one better option). Provides rationale for their decision based on data/information, but they did not gather extra information per se (i.e., using only what was initially provided to them). Examines superficially the potential costs or benefits of each decision they could have made.
- 5 = The candidate... Outlines a decision that was truly difficult to make (e.g., picking between two benefits plans, one may be better for the employees but is more expensive, the other is cheaper but may offer less coverage for employees). Provides rationale for their decision based on information/data gathered beyond the basic information available. Examines carefully the potential costs and benefits of each decision they could have made.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

Rating=	
The transcript is	as follows:"""

"q3": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"Describe a project or task where you were given complex or technical data or information to analyze and interpret in order to make a recommendation. What was the situation? How did you go about processing that information? And what was the outcome?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

- 1 = The candidate... Outlines a situation that did not require the need to be attentive to detail. The information or data that was required/ used in the situation is unclear and vague, leaving it unsure how they interpreted the information/data. The outcome is: not described or was not affected by the persons attention to detail.
- 3 = The candidate...Outlines a situation that may require some need to be attentive to detail. Explains in some detail how they ensured the accuracy of the information or data (e.g., double-checking their interpretation or the data itself). However, some details may be vague/ unclear. The outcome of the situation has limited success (e.g., there may still have been some errors in their recommendation). Or the recommendation described is of little importance.
- 5 = The candidate... Outlines a situation that clearly requires the need to be attentive to detail. Clearly explains how they ensured the accuracy of the information or data (e.g., double-checking their interpretation and the data itself). The outcome of the situation is critical and successful (e.g., there are no errors in their recommendation, data is used effectively).

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

The transcript is as follows:""",

"q4": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"You are attending a Zoom meeting with 30 managers and leaders of the company. While describing ideas for a new advertising campaign, Jason, the Vice-President of marketing makes several remarks that could been perceived as offensive by certain groups of employees. Nobody says anything about Jason's comments, probably because Jason is a senior leader in the company and outranks most attendees (you included). But you notice that a couple of individuals from that group are visibly upset. The next day, you are sending an e-mail to Jason to discuss recruiting a new employee for his team. How would you address Jason about his comments during the meeting?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

- 1 = The candidate... Does not describe to Jason that their remarks were offensive. Makes no mention in their e-mail to the executive, Jason is high on the company ladder and could get them fired. Describes getting angry/ passive aggressive with Jason, or purposely trying to humiliate the executive in their e-mail. Exclusively focuses on the new employee recruitment and nothing more to Jason in their e-mail.
- 3 = The candidate... Attempts to "dance" around the fact that the remarks were offensive. Describes being polite or professional (not both) in their e-mail. They may even describe being slightly hostile, or passive aggressive in their e-mail. Explains why remarks made were offensive but with minimal detail. Still includes communication about new employee recruitment in their e-mail
- 5 = The candidate... States firmly that what the executive said was offensive to some employees (and could have been hurtful). Describes a clear and concise c e-mail communication with Jason. They describe ways to ensure their email is professional and polite and does not convey angry or passive aggressiveness. Clearly explains why the remarks made were offensive in detail so the executive can understand their impact. Includes communication about the new employee recruitment and may even "weave" in being inclusive to help recruit new team members.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

Rating=		
The transcript	is as	follows:""

"q5": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human

resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"A team that you oversee must develop a leadership training program for the product management department. In addition to you, there are three other team members: two senior members who are very qualified and one junior member who is willing to work very hard but whose experience and capabilities in this area are very limited. This is a very important project, and your boss, the head of People and Talent, has made it clear that the project must be completed very quickly. How would you allocate work between the team members? To what extent would you involve the junior team member?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

- 1 = The candidate... Does not involve the junior at all or gives them unimportant tasks (e.g., make copies). Puts most of the work on the senior members of the team. And, avoids taking any tasks for themselves. Or takes all the important work for themselves.
- 3 = The candidate... Involves the junior member in some of the tasks but does not provide them with the necessary support. Only pairs the member with senior members (e.g., "pawning" them off). Only takes only minor tasks themselves.
- 5 = The candidate... Challenges the junior member of the team to further develop. Pairs them with a senior member (or themselves) to help guide their development (i.e., some form of mentoring while ensuring that they can perform the tasks). Takes on some important works tasks themselves to ensure the team can efficiently meet the deadline.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

Rating=	
The transcript is as follows	s:""",

"q6": """You are a professional recruiter who is responsible for evaluating the responses to a one-way video interview question provided by a job candidate who has applied for a Human resources generalist job for Western oil resources group, which is an energy services firm. The candidate was asked the following question:

"You have to design a new "work from home" policy for the company. It is an important project, which will impact many employees. You have been working on it for two months now, and you have one month left to present your final report. You have the next steps clearly laid out, and you believe that you are on track to finish it just in time if you work full-time on it. However, your supervisor informs you the timeline has changed, and the report is now due in 2 weeks. After that, you will be assigned to hiring a new engineer with specialized expertise for a project with one of the company's largest clients. How would you adjust your plans and organize your work to meet the revised deadline while starting to think about the new hiring task?".

Your task is to read the transcript of the candidate's response and give that transcript a rating from 1 to 5. Use the following scale to assign your rating and keep in mind that it is possible to give the candidate ratings of 2 and 4 when they have a mixture of elements in between the 1, 3 and 5 anchors. The rating scale is a Behaviourally Anchored Rating Scale, also known as BARS. Because of this, it is very unlikely that a transcript will provide a perfect match to any of the options. Your task is to choose the rating that most closely matches the contents of the transcript.

- 1 = The candidate... Only focuses on one task or the other leaving one of the tasks worse off. Fails to describe in detail the steps they would take. Describes a scenario where they offload the work to another individual.
- 3 = The candidate... Describes in some detail their plan of action. Mainly focuses on the "bigger" elements, which can be detrimental to some important details (e.g., cut corners to finalize the report on time). Some focus on smaller details, but they are surface level details. May touch on the hiring the engineer but nothing overly detailed.
- 5 = The candidate... Is able to describe in detail the plan of action to complete the project on time (i.e., in 2 weeks) and starts the new task. Considers bigger picture aspects of the situation, as well as carefully considering small aspects of what would be left to do (e.g., elements of the report that could be streamlined without hurting quality). Mentions double-checking of the already clearly laid out steps. Starts initial planning of hiring the engineer.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format:

Rating=	
The transcript is as follows:	"""

Appendix I

After providing the prompt ChatGPT was provided with answers:

USER:

The top manager out of 10 managers in a previous position, I was given the choice of the new recruits to choose team and went through a selection process with said recruits over a period of two weeks. I was finally able to come up with my choice, who I did pick. I approached the decision by significant time spent with each of the recruits. The one I picked went on to be wildly successful, became a manager himself, became a partner of mine in business and has done very, very, very well for himself. I got him.

ChatGPT:

Rating = 2

The candidate described a situation where they were given the responsibility to select a new recruit from a group of candidates. They mentioned spending significant time with each recruit before making their decision. However, the response lacks detailed rationale for why they chose the candidate they did. There is no mention of comparing qualifications, evaluating fit for the team, or considering potential costs and benefits of each candidate. While the outcome was positive (the selected candidate became successful), the response does not provide enough depth in terms of decision-making process or criteria used, aligning with a 2 rating on the BARS scale for this question.

USER:

We had a job opening for Lieutenant position in an institution which was about supervising between 30 and 40 staff and this was a tough selection because ten people put in for the job all of which all of who met the minimum requirements for the job. So we had to select the best qualified person for the job. There were several people on the list of people who were very popular with an institution and popular with the staff. However, they weren't the best possible selection for the job. So we selected someone who was not as popular an institution and that ended up being quite challenging for everybody in the institution because this is a leadership position in leaders and necessarily a popularity contest. So ultimately it worked out fine.

ChatGPT:

Rating = 4

The candidate described a situation where they had to select a candidate for a Lieutenant position from a pool of ten applicants who all met the minimum job requirements. They highlighted the challenge of choosing the best qualified person over more popular candidates within the institution. The decision-making process involved recognizing that leadership qualities were more critical than popularity, which demonstrates a thoughtful approach to selection criteria.

The response provides a rationale for why they chose the candidate they did, emphasizing the importance of job qualifications and leadership skills. The outcome was positive, despite initial

challenges, indicating that the decision ultimately worked out well for the institution. This aligns with a 4 rating on the BARS scale for this question.

USER:

I have been on a couple of interview panels and selecting a candidate is always very difficult. What we've resorted to doing with our current company is ranks choice voting. So we will come up with a set of criteria that we all agree to before seeing any of the client of the candidates. And then when folks come in, we ask a series of questions in real time, not asynchronous. And we are able to judge based on that rubric essentially, assigning a value to each of them near the end. We each have our own values. Everybody has their own. All points get added up at the end. And then you have a fairly quantitative value that you can assign to that person. It's not the best, of course, because there are some folks who are super charming in that, maybe the personality, the communication, the value of having that team member is much higher than the value of having their skills. But in those cases, we encourage folks to build the rubric out that way. So again, when we are hiring for a position, if it's incredibly skill-based, they're not going to be talking to a lot of people. But unless how charming they are, we have to weigh their skills heavier than their communication. But if it's a role where the skills, they can be learned, they can train. But the communication skills are very key. We weigh that heavily in the rubric before we start interviewing candidates. So that is our process that we do. And our outcome has been fairly good so far with the folks that we've got on board.

ChatGPT

It sounds like your approach to candidate selection involves a structured and systematic method using ranked choice voting and predefined criteria. Here's how I would evaluate your response based on the BARS rating scale:

Rating = 5

The candidate described a detailed and structured approach to selecting candidates during interview panels. They explained the use of ranked choice voting based on predefined criteria, highlighting a quantitative assessment process. They acknowledged the complexities involved, such as balancing technical skills and interpersonal qualities, and emphasized adapting the rubric to prioritize key attributes based on the role's requirements.

This approach demonstrates a high level of decision-making complexity, considering both quantitative metrics and qualitative factors like communication skills. The candidate's explanation shows a comprehensive understanding of effective candidate evaluation methods, ensuring that the hiring decisions align closely with the company's needs and goals. Thus, it aligns well with a rating of 5 on the BARS scale for this question.

Appendix J

Summary Communication:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the summary responses. They can be no longer than 500 words. The participants will be evaluated on technical communication on a scale from 1 to 5. Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5.

Technical communication:

- 1: The writing is... Illegible or incoherent. Choppy and hard to follow. Major flaws with punctuation, spelling, and grammar. Is not appropriately structured (e.g., just a list of bullet points). Language is informal.
- 3: The writing is... Generally made sense. Choppy in some areas, but still relatively easy to follow. Is mostly free of errors with punctuation, spelling, and grammar. Some language may be informal.
- 5: The writing is... Clearly logical and made sense (i.e., had a flow). Had no flaws in punctuation, grammar, and spelling. Is appropriately structured (i.e., looks like a proper email). Language is formal, and appropriate given a workplace setting.

Be sure to present the output in the following format: Technical communication:

The transcript is as follows:

Policy Communication:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the policy responses. The participants will be evaluated on technical communication 1 to 5. Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5.

Technical communication:

- 1: The writing is... Illegible or incoherent. Choppy and hard to follow. Major flaws with punctuation, spelling, and grammar. Is not structured like a policy: has not title or headers, the policy does not make sense based on the order it is structured in. Language is informal.
- 3: The writing is... Generally made sense. Choppy in some areas, but still relatively easy to follow. Is mostly free of errors with punctuation, spelling, and grammar. May have some preamble or title making it resemble a policy. It may or may not be in a very clear order. The language may be somewhat informal.
- 5: The writing is... Clearly logical and made sense (i.e., had a flow). Had no flaws in punctuation, grammar, and spelling. Structured like a policy, including formal language, titles

and headers to separate different portions of the policy, includes context for the policy, and the order of the policy is logical. Clearly structured and is logically laid out.

Be sure to present the output in the following format: Technical communication:

The transcript is as follows:

Recommendation Communication:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the candidate summary responses. The participants will be evaluated on technical communication from 1 to 5. Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5.

"Technical Communication:

- 1: The writing is... Illegible or incoherent. Choppy and hard to follow. Major flaws with punctuation, spelling, and grammar. Is not appropriately structured (e.g., just a list of bullet points). Language is informal.
- 3: The writing is... Generally made sense. Choppy in some areas, but still relatively easy to follow. Is mostly free of errors with punctuation, spelling, and grammar. Some language may be informal.
- 5: The writing is... Clearly logical and made sense (i.e., had a flow). Had no flaws in punctuation, grammar, and spelling. Is appropriately structured (i.e., looks like a proper email). Language is formal, and appropriate given a workplace setting.

Be sure to present the output in the following format: Technical communication:

The transcript is as follows:

Summary Decision making:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the summary responses. They can be no longer than 500 words. The participants will be evaluated on message relevance on a scale from 1 to 5. Below I provide anchors for 1, 2, 3, 4 and 5. You need to examine and use the "major points" listed below.

Major points from the article:

1. Schedule 1:1 weekly, for 30 minutes. 2. Work with your employee to create the agenda. 3. Come to the meeting with a positive tone. 4. Remove barriers for team members. 5. Let team members speak more than you do. 6. 1:1's can be a good opportunity to get to know your subordinates. 7. Make sure you clarify and create action items for everyone. 8. Its okay if you don't cover everything, but move it to the next meeting. 9. Add your perspective when appropriate. 10. Be present during the meeting, silence alerts and focus on the 1:1. 11. Make sure to communicate the why to employees; ensure they understand the purpose of the 1:1's.12. Find a time that works best for your subordinate. 13. Do not cancel meetings if you can avoid. 14.

Regular meetings reduce turnover/ increase happiness. 15. End the meeting well. 16. Set an agenda for the meetings.

Message Relevance:

- 1: The writing... Missed the major points (2 or less) of the article. Includes several non-relevant topics.
- 2: The writing... Communicates some (3-4) of the major points. May include irrelevant points.
- 3: The writing... Communicates most (6-8) major points from the article. May include a few (1-2) irrelevant points (e.g., Points about Steven Rogelberg).
- 4: The writing... Communicates a majority (8-9) of the major points. It may still include a few irrelevant points.
- 5: The writing... Summarized most of the (10 or more) major points in the article. Does not include any irrelevant points.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format: Message relevance: ___

The transcript is as follows:

Policy decision making:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the policy responses. The participants will be evaluated on message relevance from 1 to 5. Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5. Below I provide the full article for context.

Full article:

"Before the Meetings

Setting up your 1:1s should entail more than dropping invites onto your team members' calendars. You should lay the groundwork for your conversations and plan the logistics to best fit each report's unique needs.

Communicate the initiative or your reboot of the initiative.

Whether or not the practice of holding 1:1s is new to your team, announce it at a team meeting so that everyone gets the message at the same time and no one feels singled out. Tie the meetings to your organization's values and to your personal values. Also stress that these conversations are not meant to signal dissatisfaction with your team's work and are not about micromanaging; rather, they are opportunities for you and each member to get to know each other better, learn about challenges, discuss careers, and for you to give help when it's needed. This is also a good moment to tell your team members what you need from them to make the meetings successful: They should drive the agenda with key priorities, be curious, be actively engaged, communicate

candidly, think deeply about problems and solutions, and be willing to ask for help and act on feedback.

Determine cadence.

My research suggests that you should adopt one of three plans for the frequency of 1:1s: (1) You meet with each of your team members once a week for 30 minutes or so. In my surveys, employees, regardless of job level, rated this approach the most desirable; it also correlated with the highest levels of engagement. (2) In the second-highest-rated plan you meet every other week for 45 to 60 minutes. (3) In a hybrid plan you meet with some team members weekly and others every two weeks. Whichever you choose, aim to spend roughly equivalent amounts of time with employees over the course of a month so that all team members get the same support from you. To determine the right cadence, consider:

Manager tenure. If you are new to the team, weekly meetings are ideal for establishing relationships and alignment.

Remote or in person. If your team is remote, weekly meetings can help counter a lack of spontaneous face-to-face contact.

Team member preference. Finally, give your employees a voice in the decision.

I've seen some managers, mostly senior leaders, opt for three or four weeks between 1:1s, but investing only 60 minutes with each team member every month makes building a trusting relationship difficult. And because more-recent events are easier to recall, the longer time lapse also means that you're less likely to discuss any issues that arose several weeks prior. These meetings are most effective when you can build momentum around specific areas of the direct report's activities and growth. A monthly cadence makes that more challenging. But if your team members are seasoned and have worked with you a long time, and you are readily available for impromptu conversations, this cadence can work and is preferable to nothing. However, employees rated this option as least desirable, and it was associated with smaller gains in engagement.

Finally, avoid canceling 1:1s, which can hamper your team members' progress and make them feel that they are low on your priority list. This was one of Bill's problems: He readily canceled these meetings if he got busy. That sometimes demoralized his team members; they also found themselves duplicating efforts because they hadn't had a chance to coordinate their work through Bill. If you must cancel, reschedule the meeting right away, ideally for the same week. Another option is to reduce the length of the meeting: Some time together is better than none at all.

Set a location.

In my research, employees rated virtual 1:1s as slightly less desirable than those held in person, but they rated the ultimate value of the meetings similarly regardless of which form they'd taken. If you can meet in person, choose a location where you and your employee will feel at ease, present, and free of distractions. In my surveys the most highly rated location was the manager's office or a conference room; the lowest was the direct report's office. Support for outside

locations, such as coffee shops, was uneven, so don't assume that everyone would welcome them. Talk to your team members in advance to gauge where they feel most comfortable.

Create an agenda.

Many managers assume that 1:1s are too informal to require an agenda, but my research shows that having one is a strong predictor of the effectiveness of the meeting, whether it was created in advance (which is ideal) or at the meeting itself (if necessary). Even more critical, though, is the employee's involvement in the agenda's creation: Both direct reports and managers rated meetings most highly when the reports contributed to or established the agenda themselves. Bill's habit of organizing his 1:1s around his own priorities and needs meant that his team members' concerns were usually relegated to the end of the meeting – and often went unaddressed if time ran out.

Alternatively, some managers create the agenda from broad questions, such as: What would you like to talk about today? How are things going with you and your team? What are your current priorities, and are there any problems or concerns you would like to talk through? Etc.

At the Meetings

Once you've prepared for a meeting, a fruitful discussion will depend on your ability to create a comfortable setting for your employee. A valuable 1:1 addresses both practical and personal needs – to feel respected, heard, valued, trusted, and included. To ensure that a meeting does so:

Set the tone.

First, be present. Turn off email alerts, put your phone away, and silence text notifications. Remind yourself as the meeting begins that it is fundamentally about your employee's needs, performance, and engagement.

As you go into the meeting, check your emotional state. Research shows that the mood you bring to a meeting has a contagion effect, so start out with energy and optimism. Reiterate your goals and hopes for the meeting and then move to some non-work-related topics, rapport building, wins, or appreciation to generate momentum and foster feelings of psychological safety. One problem for Bill was that he viewed 1:1s as merely another task on his already long list. That affected how he facilitated (or failed to), listened, collaborated, and engaged.

Listen more than you talk.

The biggest predictor of a 1:1's success, according to my research, is the employee's active participation as measured by the amount of time that person talks during the meeting. The ideal is anywhere from 50% to 90%. The agenda will have some influence on that, but you as the manager should carefully avoid talking more than your employee does.

In addition, listen actively to fully understand your direct report before you speak yourself. Display genuine interest without judgment and acknowledge the employee's viewpoint even if you disagree with it. Ask questions that clarify and constructively challenge that viewpoint. Encourage your team member to provide thoughts and potential solutions to problems. Stay

vigilant about your body language and reactions to ensure that you're creating a welcoming and safe space.

Be flexible.

As you work through your established agenda, allow the conversation to move organically as needed to provide value. Focus on the items that are most critical. If some items go unaddressed, move them to the following 1:1. Let your employee know at the outset that real-time changes can be made to the agenda if a critical item emerges. Also, to best connect with each direct report, consider that person's preferences regarding communication, collaboration, and adjust your leadership approach accordingly. That will increase engagement, inclusion, deepen the relationship, and create trust.

End well.

Clarify takeaways and action items for both parties, including how you will support next steps. When both the manager and the employee document these, chances are better that the actions will be carried out. It also builds continuity between meetings and allows for needed follow-up. After Bill implemented this change, he was reminded that his 1:1s were not mere transactions to get through but, rather, represented employees' evolving stories – something to be nurtured and developed over time. Finally, show gratitude and appreciation for your direct report's time – and start and stop on schedule to demonstrate those feelings.""

Policy relevance:

- 1: The policy... Does not suggest how managers should conduct a 1:1 meeting. Does not include any reasons why 1:1s are important. Includes no examples in the policy. Is extremely generic and limited, ignoring the setting of WORG.
- 3: The policy... Gives a general outline but lacks details. Includes only one reason why 1:1s are important. May include an example, but it is limited. Is generic, and largely mimics the article.
- 5: The policy... Clearly includes relevant steps for managers to conduct a 1:1 meeting. Includes a couple reasons why 1:1s are important. Includes a few examples in their policy that clearly outline recommendations. Extends on the article and fits the remote setting.

Be sure to present the output in the	ne following format:	Policy relevance:
--------------------------------------	----------------------	-------------------

The transcript is as follows:

Recommandation Communication score 2 (Message Clarity):

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the candidate summary responses. The participants will be evaluated on message clarity from 1 to 5. Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5. Each participant watched candidates interview for a job and were asked to write a recommendation. Below are the candidate answers for context.

Candidate responses in text:

"Candidate 1 Question 1 Answer:

So the first thing I like to do when completing any kind of assignment, whether it's first school, which in my scenario it is, is that I like to give roles to people. I like to set roles. I like to split up work. The second most important thing for me is to not procrastinate. So obviously the very first day we're given an assignment with a new group of individuals. What I like to do is set up a group discussion, whether that is on WhatsApp or another communicating platform. I like to set that up first so we can clear that out of the way. The second thing I like to get straight is if no one else in the group is communicating and sharing ideas and wanting to get started on the project or the assignment, the first thing I do is I will be the first one to talk and I will say, okay guys, we have an assignment that is due on this date. We have to do this, this, this. I will get started on this and that's basically what I tell them or tell my group members. That basically helps just setting out the layout. I also want to ask with not procrastinating because I've been put with individuals before that procrastinate and don't put in any work until the very end of the assignment. So that is definitely something I try to remind them and try to push them towards to finish. Of course we had high quality results. I made sure everything was in line with the format that the professor was requesting. As long as all the work is split up, it meets the deadline of course by not procrastinating. And yeah, that's basically it. The very important thing is not to procrastinate. Make sure that instructions are read. Other than that, by taking initiative, we were able to complete the assignment on time and everything went well.

Candidate 1 Question 2 Answer:

So in my situation, once again, of being not actually being put into a work situation, but only in school, there have been times where my professor wasn't exactly happy about what kind of work I outputted on various assignments and projects. But one time I did get negative feedback, I didn't get a mark that I liked, but what I did was, and this was one of the first assignments, I was just getting to know that professor, but in the real business world, you could say, okay, you're just getting to know this customer, try to match their expectations. So I was just getting to know this professor, I was trying to match their expectations, I read over everything you're looking for, and I tried to write the assignment in the specific way that they're looking for. So I did that, I thought I did a pretty good job, I spent a lot of time on the assignment, I handed it in, but I got negative feedback, something that I wasn't exactly expecting, but what I did was I read over that feedback and adjusted for the next assignment or the next project. So I adjusted everything, took in all that feedback, made it better, worked on it more, and obviously I got than a mark or feedback that was better than the first time, and obviously taking in that feedback into consideration, it helped me improve my work. So what was the reason for the valuation? Obviously it was a new professor, I didn't exactly know what they're looking for, after I had a greater in-depth answer to what they're looking for, then I was able to adjust for that and improve the next time, and so forth. So that was just the first time I was getting to know the professor, I was adjusting to what they're looking for, and that worked out great for me.

Candidate 2 Question 1 Answer:

Well, let's go with a great deal of organization and planning. See, that's a past assignment. This is a tough one because all of my jobs have been so simple like we used to drive Uber. So that doesn't require much planning or organization. But just getting through the day, I believe I was born with cystic fibrosis. Unfortunately, I literally have 100% of the symptoms. Well, I looked it up and they called that an auto diagnosis. And just getting through the day is, I mean, I'm not too organized as it is. But just getting through the day is, to me, it's like the day is fly by and you need to meet the deadline very quick. So everything is, I don't even know what to pick here. But this is also asking that required a great deal, which tells me it's kind of like building the tree fort where you assist with the other people. It doesn't say other people specifically. I'm sorry. This is a tough one for me. Because I'm thinking of what has nothing had a specific deadline. Nothing required a great deal of organization. So this one might not apply to me. But I always try to do my best and provide the highest quality results regardless of what I'm doing. Even in them poker attorneys, I was talking about I just, I was a natural at it, I guess. I didn't really think I would win. I didn't know the rules. And I guess that's what I might be meant to do. But if it wasn't for these health issues, I would be, or money issues, I would probably be out there playing right now.

Candidate 2 Question 2 Answer:

Yes, I used to develop systems for roulette and I had 99 plus percent positive feedback, but I did get one negative feedback from one customer because they blame you when they lose. Obviously, this person wasn't gambling at it and we had a disclaimer. This is for entertainment purposes only. It's not like a real thing. It's just something that people buy and collect. It was a big niche back then. This was 20 years ago and I'll never forget it. It was very unfair and I just knew to do it. The reason he gave me the negative feedback was we had more than one. We had 99 percent positive. And the way I handled it was I disputed it. I don't accept false negative feedback from people. And the evaluation was as far as I know, there was only one negative feedback that actually didn't get removed from my eBay account. That's all I did was email them and just resolve it and tell them, show them. This guy is obviously a gambling addict. I showed him my disclaimer. That's it.

Candidate 3 Question 1 Answer:

So, we actually had a customer maybe about a year and a half, two years ago, who came to us. This is somebody that we had never worked with and they had a really big job for us. Without going into too much detail or giving away too many things, essentially a company that manufactures guards that go over top of your gutter on your home. They have a lot of different specifications, not only in terms of size, length, width, material, but also in terms of color, price point. And so there was just a lot of different variables that had to be involved in actually not only figuring out how we were going to do all of this, but how we were going to price all of this, who was actually going to be doing this within our facility, how we could actually go about doing this in an efficient manner, how we could do this at a costly manner, and just making sure that the customer is happy because customer satisfaction is probably our number one bullet point on our mission statement. And so when I had this very complex task, of course, because I'm the shipping and receiving manager, I am part of the group that actually figured out all of this. Now,

I am not the only one, but I am one that kind of thinks about these things in a little bit more of an analytical kind of logical way than some of the other people on my team. And because of that, the easiest way for me to figure out anything that needs a great deal of organization and planning is to actually sit there and whiteboard it. I like to sit there and write down every single thing that has to be thought about. So like I already mentioned, we're talking about all of the different sizes, all of the different colors, all of the different people that need to be involved, the cost of materials, how long things take, we had to do time studies, all of these different things. And so when we figure out all of the variables, then we can sit there and take each individual variable and figure out exactly what has to happen within that. So for me, it's kind of like this, initially, this really big macro level kind of brainstorm. And then I take each individual part and I break it into its relevant sections and I go into detail with them. This actually allows me to figure out where our pain points are, where obstacles are, where there are go or no go situations. And so having all of that pre-planning to me is probably the most effective way to deal with something that requires high levels of planning. Because if we get all the pre-planning out of the way, everything kind of falls in line. You don't end up in a situation where you're halfway through and you think, oh wow, we completely forgot about this. And now we have to go back and change everything. So I try to look at it in a macro view, down to a micro view, and I do it in a very linear fashion from the things that need to happen at the very beginning, before it even starts, all the way to the very end and how we're going to actually accomplish it.

Candidate 3 Question 2 Answer:

So it actually sounds sort of terrible to say, but working in a manufacturing environment, especially with metal products, there are a lot of times where we have clients and customers that have very strict what we call tolerances. So if an object needs to measure 5.05 inches, there may be a tolerance of 1-100th of an inch, which means that on either side of that 5.05, you have 1-100th of an inch. And if it's out of that, any bigger or any smaller than that, you're out of spec and essentially the part is no good. So we had a situation a few years ago where one of our biggest clients had requested a large amount of parts. We're talking somewhere in the 400,000 range, which is a lot of parts. Essentially what had happened was one of the people that were running the production, they actually ended up packaging the thing wrong. Not only did they package it wrong by using the wrong bags and wrong boxes, but they actually packaged it incorrectly in terms of there was supposed to be like two parts in a bag, they would either put like one in or sometimes they would put three in. Now this wasn't my responsibility, but I was the one that had to deal with it because essentially the shipment came back to me as incorrect. And so I actually got a hold of their, first of all, their quality control managers so we could figure out exactly what the problem was. So not only did they tell us that there was a few issues with the actual manufacturing of the parts, but it was actually something that was very easily detected when you looked at it. The tolerance on this made something very visibly off-center. And so the next thing that we did was we actually went to the quality control person and their manager and they actually kind of gave us a little bit of information about the actual packaging and shipping that was incorrect. So essentially all we did was kind of figure out what the actual issue was. We had like a pretty big meeting about it and we kind of collaborated with them where essentially we worked out a compromise where they would send us back certain amount of parts and we

would send, fix the issue and send it back to them. So this was kind of a compromise because they didn't send all 400,000 back to us and we didn't have to do all 4,000, 100,000 at one time, but we did have to do it. So we just kind of worked out this situation that worked good for the both of us so that we could both get what we want in the time and as efficient as we wanted without wasting too much money. And so this was a situation where I didn't necessarily have a problem with it. Of course, I was frustrated that it actually happened to begin with but I'm pretty adaptable and this was something that was easy to do.

Candidate 4 Question 1 Answer:

Past assignments, again, drawing from the center job on ordering days. At this job, we're quite chaotic and required consistent organization and planning to meet the deadline of the day. You always had to get the result right to get the product into the store. So again, a couple steps in there to make sure that you were understanding how to do the process properly, the fluidity of it, making sure you're getting on the phone at the proper time. So in general, having the notes and the books to check yourself off, go through the motions. And once you do it once, you get used to it, once you do it twice, you feel better at it by the third time. Well, you should be an expert at it. So, yeah, that's kind of the best I can draw on that in this moment.

Candidate 4 Question 2 Answer:

No, actually I can't. I always pride myself on doing a good job and you know I don't often, if I ever receive negative feedback as a very perfectionist and if I were to handle or get bad negative feedback I would handle it and strive. It's part of life and these things happen. So you add it into your rip twire and you learn from it and experience from it but personally yeah I don't ever very often get negative feedback so maybe that makes me a little bit of a go but I pride myself on doing a good job and always maintaining value."

Message clarity:

- 1: The participant... Does not recommend a candidate whatsoever. Or provides no reason for why they recommended the candidate. Or they provide a clearly biased (e.g., race/ethnicity, disability) reason for candidates.
- 3: The participant... Recommends a candidate. Provides on a simple reason for recommending a candidate (e.g., I rated them highest). May provide a summary for all candidates but it may be somewhat confusing or does not provide a summary for each candidate. Provided minimal evidence for their choice.
- 5: The participant... Provides a unbiased reason as to why they selected the candidate. Clearly states which candidate they recommend. Provides a summary of each candidate helping provide understand why they picked the top candidate. Provides a high amount of detail and therefore has evidence/reasoning for their top choice.

We do not need a justification for your rating or your guesses. Instead, be sure to present the output in the following format: Message clarity: ___

The transcript is as follows:

Appendix K

Transcript Example 1:

USER:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the candidate summary responses. The participants will be evaluated on message clarity from 1 to 5.

Below I provide anchors for 1, 3 and 5, and note that 2 and 4 are still plausible and would be a mixture of 1 and 3 or 3 and 5.

Each participant watched candidates interview for a job and were asked to write a recommendation. Below are the candidate answers for context.

Candidate responses in text:

"Candidate 1 Question 1 Answer:

So the first thing I like to do when completing any kind of assignment, whether it's first school, which in my scenario it is, is that I like to give roles to people. I like to set roles. I like to split up work. The second most important thing for me is to not procrastinate. So obviously the very first day we're given an assignment with a new group of individuals. What I like to do is set up a group discussion, whether that is on WhatsApp or another communicating platform. I like to set that up first so we can clear that out of the way. The second thing I like to get straight is if no one else in the group is communicating and sharing ideas and wanting to get started on the project or the assignment, the first thing I do is I will be the first one to talk and I will say, okay guys, we have an assignment that is due on this date. We have to do this, this, this. I will get started on this and that's basically what I tell them or tell my group members. That basically helps just setting out the layout. I also want to ask with not procrastinating because I've been put with individuals before that procrastinate and don't put in any work until the very end of the assignment. So that is definitely something I try to remind them and try to push them towards to finish. Of course we had high quality results. I made sure everything was in line with the format that the professor was requesting. As long as all the work is split up, it meets the deadline of course by not procrastinating. And yeah, that's basically it. The very important thing is not to procrastinate. Make sure that instructions are read. Other than that, by taking initiative, we were able to complete the assignment on time and everything went well.

Candidate 1 Question 2 Answer:

So in my situation, once again, of being not actually being put into a work situation, but only in school, there have been times where my professor wasn't exactly happy about what kind of work I outputted on various assignments and projects. But one time I did get negative feedback, I didn't get a mark that I liked, but what I did was, and this was one of the first assignments, I was just getting to know that professor, but in the real business world, you could say, okay, you're just getting to know this customer, try to match their expectations. So I was just getting to know this professor, I was trying to match their expectations, I read over everything you're looking for, and I tried to write the assignment in the specific way that they're looking for. So I did that, I thought I did a pretty good job, I spent a lot of time on the assignment, I handed it in, but I got negative feedback, something that I wasn't exactly expecting, but what I did was I read over that feedback and adjusted for the next assignment or the next project. So I adjusted everything, took in all that feedback, made it better, worked on it more, and obviously I got than a mark or feedback that was better than the first time, and obviously taking in that feedback into

consideration, it helped me improve my work. So what was the reason for the valuation? Obviously it was a new professor, I didn't exactly know what they're looking for, after I had a greater in-depth answer to what they're looking for, then I was able to adjust for that and improve the next time, and so forth. So that was just the first time I was getting to know the professor, I was adjusting to what they're looking for, and that worked out great for me. Candidate 2 Question 1 Answer:

Well, let's go with a great deal of organization and planning. See, that's a past assignment. This is a tough one because all of my jobs have been so simple like we used to drive Uber. So that doesn't require much planning or organization. But just getting through the day, I believe I was born with cystic fibrosis. Unfortunately, I literally have 100% of the symptoms. Well, I looked it up and they called that an auto diagnosis. And just getting through the day is, I mean, I'm not too organized as it is. But just getting through the day is, to me, it's like the day is fly by and you need to meet the deadline very quick. So everything is, I don't even know what to pick here. But this is also asking that required a great deal, which tells me it's kind of like building the tree fort where you assist with the other people. It doesn't say other people specifically. I'm sorry. This is a tough one for me. Because I'm thinking of what has nothing had a specific deadline. Nothing required a great deal of organization. So this one might not apply to me. But I always try to do my best and provide the highest quality results regardless of what I'm doing. Even in them poker attorneys, I was talking about I just, I was a natural at it, I guess. I didn't really think I would win. I didn't know the rules. And I guess that's what I might be meant to do. But if it wasn't for these health issues, I would be, or money issues, I would probably be out there playing right now.

Candidate 2 Question 2 Answer:

Yes, I used to develop systems for roulette and I had 99 plus percent positive feedback, but I did get one negative feedback from one customer because they blame you when they lose. Obviously, this person wasn't gambling at it and we had a disclaimer. This is for entertainment purposes only. It's not like a real thing. It's just something that people buy and collect. It was a big niche back then. This was 20 years ago and I'll never forget it. It was very unfair and I just knew to do it. The reason he gave me the negative feedback was we had more than one. We had 99 percent positive. And the way I handled it was I disputed it. I don't accept false negative feedback from people. And the evaluation was as far as I know, there was only one negative feedback that actually didn't get removed from my eBay account. That's all I did was email them and just resolve it and tell them, show them. This guy is obviously a gambling addict. I showed him my disclaimer. That's it.

Candidate 3 Question 1 Answer:

So, we actually had a customer maybe about a year and a half, two years ago, who came to us. This is somebody that we had never worked with and they had a really big job for us. Without going into too much detail or giving away too many things, essentially a company that manufactures guards that go over top of your gutter on your home. They have a lot of different specifications, not only in terms of size, length, width, material, but also in terms of color, price point. And so there was just a lot of different variables that had to be involved in actually not only figuring out how we were going to do all of this, but how we were going to price all of this, who was actually going to be doing this within our facility, how we could actually go about doing this in an efficient manner, how we could do this at a costly manner, and just making sure that the customer is happy because customer satisfaction is probably our number one bullet point on our mission statement. And so when I had this very complex task, of course, because I'm the

shipping and receiving manager, I am part of the group that actually figured out all of this. Now, I am not the only one, but I am one that kind of thinks about these things in a little bit more of an analytical kind of logical way than some of the other people on my team. And because of that, the easiest way for me to figure out anything that needs a great deal of organization and planning is to actually sit there and whiteboard it. I like to sit there and write down every single thing that has to be thought about. So like I already mentioned, we're talking about all of the different sizes, all of the different colors, all of the different people that need to be involved, the cost of materials, how long things take, we had to do time studies, all of these different things. And so when we figure out all of the variables, then we can sit there and take each individual variable and figure out exactly what has to happen within that. So for me, it's kind of like this, initially, this really big macro level kind of brainstorm. And then I take each individual part and I break it into its relevant sections and I go into detail with them. This actually allows me to figure out where our pain points are, where obstacles are, where there are go or no go situations. And so having all of that pre-planning to me is probably the most effective way to deal with something that requires high levels of planning. Because if we get all the pre-planning out of the way, everything kind of falls in line. You don't end up in a situation where you're halfway through and you think, oh wow, we completely forgot about this. And now we have to go back and change everything. So I try to look at it in a macro view, down to a micro view, and I do it in a very linear fashion from the things that need to happen at the very beginning, before it even starts, all the way to the very end and how we're going to actually accomplish it.

Candidate 3 Question 2 Answer:

So it actually sounds sort of terrible to say, but working in a manufacturing environment, especially with metal products, there are a lot of times where we have clients and customers that have very strict what we call tolerances. So if an object needs to measure 5.05 inches, there may be a tolerance of 1-100th of an inch, which means that on either side of that 5.05, you have 1-100th of an inch. And if it's out of that, any bigger or any smaller than that, you're out of spec and essentially the part is no good. So we had a situation a few years ago where one of our biggest clients had requested a large amount of parts. We're talking somewhere in the 400,000 range, which is a lot of parts. Essentially what had happened was one of the people that were running the production, they actually ended up packaging the thing wrong. Not only did they package it wrong by using the wrong bags and wrong boxes, but they actually packaged it incorrectly in terms of there was supposed to be like two parts in a bag, they would either put like one in or sometimes they would put three in. Now this wasn't my responsibility, but I was the one that had to deal with it because essentially the shipment came back to me as incorrect. And so I actually got a hold of their, first of all, their quality control managers so we could figure out exactly what the problem was. So not only did they tell us that there was a few issues with the actual manufacturing of the parts, but it was actually something that was very easily detected when you looked at it. The tolerance on this made something very visibly off-center. And so the next thing that we did was we actually went to the quality control person and their manager and they actually kind of gave us a little bit of information about the actual packaging and shipping that was incorrect. So essentially all we did was kind of figure out what the actual issue was. We had like a pretty big meeting about it and we kind of collaborated with them where essentially we worked out a compromise where they would send us back certain amount of parts and we would send, fix the issue and send it back to them. So this was kind of a compromise because they didn't send all 400,000 back to us and we didn't have to do all 4,000, 100,000 at one time, but we did have to do it. So we just kind of worked out this situation that worked good for the

both of us so that we could both get what we want in the time and as efficient as we wanted without wasting too much money. And so this was a situation where I didn't necessarily have a problem with it. Of course, I was frustrated that it actually happened to begin with but I'm pretty adaptable and this was something that was easy to do.

Candidate 4 Question 1 Answer:

Past assignments, again, drawing from the center job on ordering days. At this job, we're quite chaotic and required consistent organization and planning to meet the deadline of the day. You always had to get the result right to get the product into the store. So again, a couple steps in there to make sure that you were understanding how to do the process properly, the fluidity of it, making sure you're getting on the phone at the proper time. So in general, having the notes and the books to check yourself off, go through the motions. And once you do it once, you get used to it, once you do it twice, you feel better at it by the third time. Well, you should be an expert at it. So, yeah, that's kind of the best I can draw on that in this moment.

Candidate 4 Question 2 Answer:

No, actually I can't. I always pride myself on doing a good job and you know I don't often, if I ever receive negative feedback as a very perfectionist and if I were to handle or get bad negative feedback I would handle it and strive. It's part of life and these things happen. So you add it into your rip twire and you learn from it and experience from it but personally yeah I don't ever very often get negative feedback so maybe that makes me a little bit of a go but I pride myself on doing a good job and always maintaining value."

Message clarity:

- 1: The participant... Does not recommend a candidate whatsoever. Or provides no reason for why they recommended the candidate. Or they provide a clearly biased (e.g., race/ethnicity, disability) reason for candidates.
- 3: The participant... Recommends a candidate. Provides on a simple reason for recommending a candidate (e.g., I rated them highest). May provide a summary for all candidates but it may be somewhat confusing. Provided minimal evidence for their choice.
- 5: The participant... Provides a unbiased reason as to why they selected the candidate. Clearly states which candidate they recommend. Provides a summary of each candidate helping provide understand why they picked the top candidate. Provides a high amount of detail and therefore has evidence/reasoning for their top choice. Be sure to present the output in the following format:

Mes	sage cl	larity:			
The	transci	ript is	as	folloy	VS

Hi Kathelene, I'm writing to confidently recommend candidate #3 as the top candidate for the position of Candidate 3 came off as much more professional and experienced than the other candidates, and seems like me will be a great fit at WORG. Specifically, for the question around organization, he provided a good work-related example, where he showed he was analytical and organized. He spoke about whiteboarding as a technique for organizing complex information. With regard to receiving negative feedback, he showed the ability to take accountability and achieve compromise in manufacturing setting. While I don't think the other candidates were particularly close to being at the same level as #3, here is a brief overview of their interview answers: Candidate 1: Indicated that he doesn't procrastinate, encourages others to talk, but didn't speak to a specific work experience. Showed inability to draw on relevant experience. For the question on negative feedback, he gave a school example, rather than a work example. He

seemed to be inexperienced and had trouble following simple instructions. Candidate 2: Had trouble answering basic question around organization, bringing up irrelevant facts. He actually said "I'm not too organized" and getting through the day is an accomplishment. He also made excuses about receiving negative feedback, saying it was customer's fault that he was receiving negative feedback. Said it was unfair. Candidate 4: He was not articulate. Didn't give specific example. He claimed to never get negative feedback. Didn't come off as credible. Everyone should be able to think of an example.

ChatGPT:

Message clarity: 5

The participant provides a clear, unbiased recommendation for Candidate 3, thoroughly explaining why this candidate is the top choice. They provide a detailed summary of Candidate 3's strengths, citing specific examples from the candidate's interview to support their recommendation. Additionally, the participant offers a brief but coherent overview of the other candidates' responses, highlighting why they are not as strong as Candidate 3. The evaluation includes sufficient evidence and reasoning, making it easy to understand the basis for their decision. This detailed and well-supported recommendation aligns with the criteria for a score of 5.

Transcript Example 2:

USER:

You are a research assistant who is responsible for evaluating the responses for a job task. What I need you to do is evaluate the summary responses. They can be no longer than 500 words. The participants will be evaluated on technical communication on a scale from 1 to 5. Below I provide anchors for 1, 2, 3, 4 and 5. You need to examine and use the "major points" listed below.

Major points from the article:

1. Schedule 1:1 weekly, for 30 minutes. 2. Work with your employee to create the agenda. 3. Come to the meeting with a positive tone. 4. Remove barriers for team members. 5. Let team members speak more than you do. 6. 1:1's can be a good opportunity to get to know your subordinates. 7. Make sure you clarify and create action items for everyone. 8. Its okay if you don't cover everything, but move it to the next meeting. 9. Add your perspective when appropriate. 10. Be present during the meeting, silence alerts and focus on the 1:1. 11. Make sure to communicate the why to employees; ensure they understand the purpose of the 1:1's.12. Find a time that works best for your subordinate. 13. Do not cancel meetings if you can avoid. 14. Regular meetings reduce turnover/increase happiness. 15. End the meeting well. 16. Set an agenda for the meetings.

Message Relevance:

- 1: The writing... Missed the major points (less than 3) of the article. Includes several non-relevant topics.
- 2: The writing... Communicates some (3-4) of the major points. May include irrelevant points.
- 3: The writing... Communicates most (6-8) major points from the article. May include a few (1-
- 2) irrelevant points (e.g., Points about Steven Rogelberg).

- 4: The writing... Communicates a majority (8-9) of the major points. It may still include a few irrelevant points.
- 5: The writing... Summarized most of the (10 or more) major points in the article. Does not include any irrelevant points.

Be sure to present the output in the following format: Message relevance: ___

The transcript is as follows:

This article summaries the importance of one on one meetings with managers and the people that report to them. To make the most of one on one meetings managers should communicate the initiative, determine cadence, set a location and create an agenda. Managers should also set the tone, listen more than you talk and be flexible. Managers should end meetings by summarizing important takeaways and items that need action. They should thank employees for their time. Effectively implementing one on one meetings can lead to better performance and lower employee turnover.

ChatGPT:

Message relevance: 3

The response effectively communicates most of the major points from the article but falls short in a few areas:

- 1. Communicated Major Points:
 - o Set an agenda for the meetings: "create an agenda"
 - o Let team members speak more than you do: "listen more than you talk"
 - o Come to the meeting with a positive tone: "set the tone"
 - Make sure to clarify and create action items for everyone: "summarizing important takeaways and items that need action"
 - Regular meetings reduce turnover/increase happiness: "better performance and lower employee turnover"
 - End the meeting well: "end meetings by summarizing...thank employees for their time"
- 2. Missing Major Points:
 - o Schedule 1:1 weekly, for 30 minutes
 - o Work with your employee to create the agenda
 - o Remove barriers for team members
 - o 1:1's can be a good opportunity to get to know your subordinates
 - o Its okay if you don't cover everything, but move it to the next meeting
 - o Add your perspective when appropriate
 - o Be present during the meeting, silence alerts and focus on the 1:1
 - o Make sure to communicate the why to employees; ensure they understand the purpose of the 1:1's
 - o Find a time that works best for your subordinate
 - o Do not cancel meetings if you can avoid
- 3. Irrelevant Points:

o No irrelevant points were included.

The response covers some essential points and captures the essence of the article but misses a few crucial details that could enhance the summary's completeness and effectiveness.

Appendix L *Tables for analysis with all components of decision making included*

Table L1Study variables means, standard deviations, and new correlations with revised decision making task performance and revised overall performance

Variable	M	SD	1	2
1. Decision making (task)	0.69	0.14		
2. Task performance (overall)	37.19	10.94	.39**	
Attention to detail (interview)	2.25	0.86	.22**	.31**
Communication (interview)	2.59	0.87	.08	.23**
Decision making (interview)	2.76	0.9	.03	.24**
Interview performance (overall)	2.53	0.73	.13	.32**
Honest defensive	2.64	1.03	08	14
Honest ingratiation	3.16	1.01	15	13
Honest self promotion	3.53	0.89	.05	.16*
Honest IM total	3.11	0.75	08	05
Deceptive image protection	2.18	1.05	02	09
Deceptive ingratiation	2.45	0.95	.00	12
Deceptive slight image creation	1.57	0.74	.05	.14
Deceptive extensive image				
creation	1.39	0.71	.00	.11
Deceptive IM total	1.9	0.64	.01	.00
Age	41.62	11.39	22**	25**
AVI experience	2.28	5.90	03	06
Interview experience	30.03	80.03	02	.00
HR experience	4.93	5.78	08	17*
Ethnicity	0.65	0.48	13	.17*
Gender	0.50	0.50	.08	09
Anxiety	2.00	0.96	.00	.00
Motivation	4.39	0.55	10	02

Note. ** p < .01, * p < .05. Ethnicity coded as 1 = white, 0 = non-white.

Gender coded as 1 = Female and other, 0 = Male.

Table L2Path analysis table with modified overall performance.

Relationship χ ²		Robus	st CFI		ıst RMSEA		BIC
Model fit 0.0	00	1.00		0.00		1684.84	1748.25
Interview performance	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1	-0.07	0.06	.252	10	-0.18	0.05
Honest ingratiation	a2	-0.04	0.07	.530	06	-0.19	0.09
Honest self promotion	a3	0.16	0.06	.012	.20	0.03	0.28
Deceptive image protection	a4	-0.07	0.05	.161	10	-0.17	0.03
Deceptive ingratiation	a5	-0.07	0.07	.353	09	-0.21	0.08
Deceptive slight image creation	a6	0.20	0.12	.094	.20	-0.02	0.44
Deceptive extensive image							
creation	a7	0.03	0.10	.731	.03	-0.19	0.21
Task performance	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1	-0.87	0.83	.293	08	-2.48	0.78
Honest ingratiation	c2	-1.05	1.13	.353	10	-3.16	1.20
Honest self promotion	c3	2.60	1.10	.018	.21	0.44	4.80
Deceptive image protection	c4	-0.46	0.96	.633	04	-2.36	1.41
Deceptive ingratiation	c5	-1.30	1.33	.331	11	-4.04	1.22
Deceptive slight image creation	c6	2.36	1.53	.123	.16	-0.77	5.21
Deceptive extensive image							
creation	c7	0.13	1.62	.935	.01	-2.95	3.56
Interview performance	b1	3.53	1.33	.008	.23	0.90	6.16
Indirect effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	a1*b1	-0.23	0.24	.330	02	-0.85	0.11
Honest ingratiation	a2*b1	-0.15	0.27	.571	01	-0.88	0.27
Honest self promotion	a3*b1	0.57	0.33	.083	.05	0.09	1.44
Deceptive image protection	a4*b1	-0.25	0.21	.239	02	-0.82	0.05
Deceptive ingratiation	a5*b1	-0.24	0.30	.410	0.02	-0.99	0.22
Deceptive slight image creation	a6*b1	0.69	0.47	.146	.05	0.04	2.01
Deceptive extensive image	a7*b1						
creation	a7 01	0.12	0.38	.748	.01	-0.64	0.92
Total effects	Path	b	se	p-value	β	CI-lower	CI-upper
Honest defensive	c1+a1*b1	-1.10	0.87	.204	10	-2.77	0.62
Honest ingratiation	c2+a2*b1	-1.20	1.13	.287	11	-3.31	1.09
Honest self promotion	c3+a3*b1	3.17	1.08	.003	.26	0.96	5.29
Deceptive image protection	c4+a4*b1	-0.71	1.00	.476	07	-2.69	1.20
Deceptive ingratiation	c5+a5*b1	-1.54	1.41	.274	13	-4.38	1.15
Deceptive slight image creation	c6+a6*b1	3.04	1.54	.048	.20	-0.09	6.01
Deceptive extensive image							
creation	c7+a7*b1	0.25	1.68	.880	.02	-3.10	3.53

Note. Bolded values indicate significance.

Appendix M

Follow-up correlation table examining honest defensive IM and honest ingratiation and time spent.

Variable	M	SD	1	2	3
1. Honest defensive IM	2.64	1.03	-		
2. Honest Ingratiation	3.14	1.00	.48**	-	
3. Time (secs) spent on task 1					
(article task)	773.86	183.53	18*	03	-
4. Time (secs) spent on task 2					
(candidate rating)	948.21	267.34	16*	06	.41**

Note. * indicates p < .05. ** indicates p < .01.