Tell me more! Examining the Benefits of Adding Structured Probing in Asynchronous Video Interviews

Rahul D. Patel[1]

Deborah M. Powell[1]

Nicolas Roulin[2]

Jeffrey R. Spence[1]

[1] University of Guelph; [2] Saint Mary's University

Article "in press" in *International Journal of Selection and Assessment*

Abstract

The personnel selection field has observed the rising use of asynchronous video interviews (AVI). The current study investigates whether follow-up questions (probes) can optimize the applicant experience in AVIs. Across two experimental studies with participants recruited from Prolific, we investigated whether AVIs with probing promote applicant reactions (e.g., opportunity to perform perceptions) toward the AVI and how probing influences interview behaviors, applicant perceptions, and interview performance ratings. In Study 1, 404 participants were randomly assigned to either an AVI with probing or an AVI without probing. Results indicated that probing directly improved opportunity to perform perceptions and interview performance ratings. In addition, probing positively impacted honest IM and motivation to perform indirectly through participants' perceived opportunity to perform. However, mediation analyses suggested that the effect of probing on interview performance ratings was driven by response length. In Study 2 (n = 271), we teased apart the effects of the inherently added response time that probing affords applicants with an additional condition that matched the response time of probes. Relative to Study 1, probing only slightly improved opportunity to perform perceptions, but the effect of probing on opportunity to perform perceptions was greater when compared to an AVI with equivalent response time. In addition, probing positively impacted interview performance ratings, above and beyond their increased response time. Implications, limitations, and directions for future research are discussed.

Keywords: applicant reactions, probing, opportunity to perform, impression management, interview performance ratings, asynchronous video interviews

Practitioner Points

What is currently known about the topic of our study:

- Asynchronous video interviews (AVIs) are increasingly used as a selection tool.
- AVIs are customizable in their design, yet the effects of AVI design features on interview outcomes are understudied.
- Probes are follow-up questions intended to enhance insufficient or unfinished answers or to request more details or explanations from interviewees.
- There remains a debate on the use of probing in interviews, and the effect of probing on interview outcomes remains understudied.

What our paper adds to this:

- We introduce structured probing as a customizable design feature in AVIs.
- We demonstrate that probes enhance opportunities to perform and yield higher interview performance ratings.
- Greater perceived opportunities to perform led to more engagement in honest impression management and motivation to perform.
- Applicants tend to use probes often, relative to other AVI design features.

The implications of our study findings for practitioners:

- Structured probes can be implemented to improve applicants' perceived opportunities to perform and optimize their interview performance.
- By enhancing perceived opportunities to perform with probing, applicants tend to promote their true qualifications and are more motivated to perform their best in the AVI.
- Pre-testing and refining probes are crucial to optimize applicant experience and validity.

Tell me more! Examining the Benefits of Adding Structured Probing in Asynchronous Video Interviews

When selecting personnel, there has been increasing growth in the use of the *asynchronous video interview* (AVI; Lukacik et al., 2022), an interview format in which applicants log onto an online portal and record their responses to interview questions using their computer or mobile device. Applicant responses are electronically stored and evaluated later—by a human or AI rater (Lukacik et al., 2022). AVIs bring several benefits to organizations, such as reduced costs (e.g., Mejia & Torres, 2018) and increased standardization and customizability of design features (Lukacik et al., 2022). AVIs can also benefit applicants by providing increased flexibility and convenience, as applicants can complete interviews in their own time and space and with different devices (Lukacik et al., 2022).

Despite these potential benefits, there are some side effects of AVIs in the form of negative applicant reactions. Specifically, applicants rate AVIs low in opportunity to perform as they do not feel the interview affords them the opportunity to demonstrate their abilities, qualifications, and skills (Guchait et al., 2014). Opportunity to perform perceptions are part of an applicant's overall evaluation of how fair a selection process is (Gilliland, 1993). Such negative perceptions toward AVIs may be costly to organizations, resulting in organizational reputation damages (Cropanzana et al., 2007), losing out on qualified applicants (Gilliland, 1993), and lower applicant motivation in the interview (Bourdage et al., 2018; Levashina & Campion, 2006).

Given the benefits of AVIs, there may be ways to improve their design to mitigate some of their side effects. Interestingly, despite the design customizability benefits that AVIs provide to organizations, we know little about how different AVI design features may affect applicants'

opportunity to perform perceptions, interview behaviors, and interview performance (Lukacik et al., 2022).

One customizable design feature that has the potential to influence applicant experiences with AVIs is the use of follow-up questions, or probes, after the applicants' initial response. Typically, probes are used to enhance an insufficient or unfinished answer or to request more details or explanations from applicants (Levashina et al., 2014). Research on probing questions in structured interviews is limited (Levashina et al., 2014), and conflicting views exist on using probes. On the one hand, probing may improve the accuracy of information elicited from interviewees (Schmidt & Conaway, 1999), with some evidence suggesting that using probes with past behavior interview questions can help applicants produce more narratively rich and complete stories (Brosy et al., 2020). In addition, scholars have argued that probing allows applicants to identify and recall relevant experiences when coming up with a response (Levashina et al., 2014). On the other hand, other scholars have argued that limiting probing is necessary to maintain interview structure and consistency among applicants (Campion et al., 1997; Dipboye et al., 2004; Huffcutt & Arthur, 1994). Yet, AVIs provide a unique situation to test probes because all applicants could encounter and respond to the same probes within the same time constraints, addressing concerns about the negative effects of inconsistent probing on interview structure.

In the current study, we investigated whether follow-up questions or probing applicants' initial responses may create a better applicant experience with AVIs. This research contributes to the growing literature on AVIs in three key ways: First, it is an empirical extension of a recently proposed conceptual model of the effects of AVI design features on interviewee behavior, reactions, and outcomes (Lukacik et al., 2022). Second, it tests the feasibility of probing in AVIs,

a design feature that has yet to be examined in this context (e.g., Dunlop et al., 2022). Third, the current research addresses one of the debates in the general employment interview literature around whether probing reduces interview structure (Levashina et al., 2014); we introduce a form of structured probing, which may address the interview structure concerns while retaining the benefits for the applicants' interview experience.

Theoretical Background

Lukacik et al. (2022) recently summarized AVI research, noting a dearth of empirical research examining how different design features may influence applicants' interview experiences, behaviors, and outcomes. Drawing from theories of justice in employee selection contexts, social presence, interview anxiety, and impression management, they developed a conceptual model illustrating how different AVI design features (e.g., response preparation time, re-recording of responses) may impact interview and post-interview outcomes, such as interview behaviors, performance, and applicant reactions.

Lukacik et al.'s (2022) model excludes probing as an AVI design feature. Probing may have comparable effects to extra preparation time, affording applicants more chances at performing, affecting their behaviors by granting them more chances to use impression management tactics, increasing their motivation to perform, and generally providing them with more avenues for improving their AVI performance. Accordingly, the current study uses and extends the model developed by Lukacik and colleagues as an overarching framework to test (a) whether the availability of probing in AVIs affects applicants' opportunity to perform perceptions, use of impression management tactics, interview motivation, and interview performance ratings, and (b) the mechanisms for how probing can influence interview performance ratings.

Opportunity to Perform

A key foundation of the applicant reactions literature, incorporated into Lukacik et al.'s (2022) AVI model, is Gilliland's (1993) justice-based model of applicant reactions to employment selection systems. Gilliland proposed that a selection system's compliance with procedural and distributive justice rules increases applicants' overall perceptions of fairness. Probes may help to improve one procedural justice rule in particular—the opportunity to perform. Perceptions of opportunity to perform concern the extent to which applicants feel they have adequate opportunities to demonstrate their knowledge, skills, and other abilities in the selection situation (or the possibility of exerting control in the selection situation).

Probes may bolster opportunities to perform in unique ways compared to other AVI response formatting features. Basch, Brenner, et al. (2021) found that offering one minute of preparation time in the AVI (versus just enough time to read the question) positively influenced applicants' opportunity to perform perceptions indirectly through honest impression management. Other research has investigated whether offering applicants multiple attempts to record their answers and unlimited preparation impacted their fairness perceptions toward the AVI (Roulin, Wong, et al., 2023); unexpectedly, neither design element was related to fairness perceptions. Although response formatting features like preparation time are proposed to bolster applicants' opportunities to perform in AVIs (Lukacik et al., 2022) by allowing applicants to self-direct how they organize their responses to interview questions, probes can better guide applicants toward more relevant and detailed responses by having them elaborate on specific parts of their responses. Moreover, offering probes may be one way of alerting applicants to consider whether their responses are insufficient, providing a path for them to augment their responses if needed. Thus, probes might prompt applicants to clarify their answers or provide

more detailed information about their past experiences, allowing them to present their job-related skills and abilities better (Levashina et al., 2014). Consistent with the standardized nature of the AVI and the benefits of its standardization on applicants' fairness perceptions (Basch & Melchers, 2019), we expect that standardized probes will have positive effects on applicants' opportunity to perform perceptions.

H1: The availability (versus absence) of probing in AVIs will increase applicants' opportunity to perform perceptions.

Honest Impression Management

In addition to increasing opportunity to perform perceptions, the availability of probes may also influence other applicant behaviors, such as opportunities to engage in honest impression management (IM). Honest IM is "the means by which job applicants describe their true job-related credentials during a short interaction with an interviewer" (Bourdage et al., 2018, p. 598). These tactics can include presenting one's true qualifications positively or emphasizing how one's personality and motivations align with the organization's values or goals.

Levashina and Campion (2006) proposed that applicants' use of IM (specifically deceptive IM) is a multiplicative function of their capability (e.g., social and cognitive abilities), willingness (e.g., certain personality traits), and the opportunity (e.g., unstructured versus structured interviews). Lukacik et al. (2022) argued that certain AVI design features may influence the extent to which applicants are willing, capable, and have the opportunity to engage in both honest and deceptive IM. For instance, allowing applicants more time to prepare their responses or providing multiple AVI recording attempts may afford them more opportunities to put their best foot forward. Empirical evidence suggests that AVI design features can affect honest IM. For example, Basch, Brenner, et al. (2021) found that offering applicants preparation

time (versus none) in AVIs afforded them greater opportunities to use honest IM (d = 0.56). Similarly, probes can allow applicants to augment their initial responses with additional honest IM tactics that could highlight their competence and fit with employers. Though similar features like preparation time allow applicants to organize their responses in a way that may incorporate honest IM, applicants may not recognize if they need to use honest IM. Conversely, probing (and the targeted guidance probes can provide) may signal applicants to incorporate details about their true qualifications and fit with the employer.

H2: The availability (versus absence) of probing in AVIs will increase applicants' use of honest IM.

Not only may the availability of probing directly promote honest IM, but the opportunity to perform perceptions associated with these probes may indirectly motivate applicants to engage in honest IM. This follows from Bourdage et al. (2018), who asserted that applicants should be more willing to rely on their true qualifications (i.e., honest self-promotion) and perceive a greater capacity to engage in these tactics in the face of more procedurally just interviews. Indeed, Bourdage et al. (2018) found that higher perceptions of procedural justice were positively associated with increased levels of honest IM (e.g., r = .30 for honest self-promotion). Thus, the availability of probing may not only directly lead to higher levels of honest IM but may indirectly do so via promoting applicants' opportunity to perform perceptions.

H3: The relationship between the availability (versus absence) of probing and applicants' use of honest IM will be partially mediated by applicants' opportunity to perform perceptions.

Interview Performance Ratings

The availability of probing may allow applicants to increase their interview performance ratings—employers' evaluations of applicants' responses to interview questions (Huffcutt, 2011). Behavioral interviews typically ask applicants to tell a story about a past situation. The completeness of the 'stories' told is assessed by recruiters based on the situation, task, actions, and results provided in responses—commonly known as the STAR technique (Kessler, 2006). Structured probes may elicit responses that contain more of these elements. Probes can allow applicants to enhance the narrative richness of their responses by prompting them to reflect on aspects of their experiences that they might not have considered compared to other AVI design features, such as preparation time. Indeed, Brosy et al. (2020) determined that probes can encourage applicants to add more details related to the actions and results parts of the experiences they described. In addition, Roulin, Pham, et al. (2023) found that across two studies, applicants given training on AVIs received higher interview performance ratings through more structured responses that incorporated storytelling aspects such as the situation, task, action, and results or outcomes of previous experiences. Similarly, probes may give applicants the necessary direction to formulate better responses. Thus, probes that can target aspects of storytelling may help applicants determine what may be missing in their responses and provide increased opportunities to tell more narratively rich responses to past behavior interview questions (Brosy et al., 2020).

The potential performance enhancements afforded by probes can be illustrated by comparing probes with other response formatting features. Basch, Brenner, et al. (2021) found that offering applicants one additional minute of preparation time in the AVI yielded higher interview performance ratings. Similarly, Roulin, Wong, et al. (2023) found that applicants who used more available preparation time tended to receive higher interview performance ratings.

However, probing may add greater performance benefits beyond preparation time by providing specific, structured guidance to shape and enhance the relevance and depth of applicants' responses, leading to a more positive evaluation of their job-related competencies. Although preparation time can allow applicants to reflect and organize their thoughts, probing actively assists applicants in uncovering and communicating the most pertinent aspects of their experiences, akin to receiving real-time coaching to fine-tune their narratives. Moreover, when applicants are asked past behavior questions, probes may help them better access their autobiographical memory, provide a more comprehensive account of their past behavior, and ultimately better showcase their skills and competencies (Murphy, 2020). Therefore, structured probes may increase applicants' interview performance ratings due to offering structured guidance.

H4: The availability (versus absence) of probing will lead to higher interview performance ratings.

In addition to a direct performance rating effect, this increase in interview performance may also be indirect through honest IM. Applicants' use of honest IM is typically expected and encouraged (Jansen et al., 2012). When used, honest IM conforms to evaluators' expectations and can help applicants spotlight their skills, likeability, competence, and fit (Robie et al., 2020). Studies have consistently found a positive relationship between honest IM and interview performance ratings (e.g., Amaral et al., 2019; Bourdage et al., 2018). Furthermore, a recent meta-analysis suggests that applicants' use of honest (but not deceptive) IM is positively related to their interview performance ratings, $\rho = .21$, 95% CI [.01, .41] (Ho et al., 2021). Thus, probing may increase honest IM (H2) and interview performance ratings (H4), and honest IM may be one mechanism through which probes increase interview performance ratings.

H5: Applicants' use of honest IM will partially mediate the relationship between the availability (versus absence) of probing and applicants' interview performance ratings.

Interview Motivation

Increased opportunity to perform perceptions may also increase applicants' motivation to expend effort during the interview (McCarthy et al., 2009). This potential relation follows from Gilliland (1993), who proposed that justice perceptions impact test-taking motivation. That is, being given more opportunities to perform in the AVI might make applicants feel their efforts have a higher chance of leading to successful outcomes (e.g., advancing to the next stage in the hiring process, receiving a job offer) to the extent they can capitalize on those more available opportunities and demonstrate their competence. Furthermore, Lukacik et al. (2022) proposed that fairness perceptions broadly mediate the relationship between AVI design decisions and applicant reactions, such as interview motivation. As opportunity to perform is one procedural justice component, we reason that opportunity to perform perceptions also influence motivation to perform in the AVI.

There is evidence that opportunity to perform perceptions, or broadly—fairness perceptions—motivate applicants to perform well. For example, Bell et al. (2006) found a positive relationship between procedural justice perceptions and test-taking motivation (r = .11). Bauer et al. (2006) examined applicants' reactions to online screening procedures and found that procedural justice perceptions were positively and strongly related to applicants' test-taking motivation (r = .54). Furthermore, McCarthy et al. (2009) examined police officer applicants' reactions to promotional exams and reported that procedural justice perceptions were positively related to test-taking motivation (r = .13). Overall, these studies offer preliminary support for the

idea that selection situations that bolster opportunity to perform may increase applicants' motivation to perform well.

H6: The availability (versus absence) of probing will increase applicants' interview motivation, and this relationship will be mediated by applicants' opportunity to perform perceptions.

Interview motivation may also play a role in determining interview performance ratings (Huffcutt, 2011). Applicants motivated to perform in the interview may put in more effort insofar as such effort could improve their performance. Meta-analytic work by Hausknecht et al. (2004) supports this argument with their estimate of a positive test-taking motivation—test performance relationship (r = .21, 95% CI [.19, .21]). Therefore, applicants' opportunity to perform perceptions may lead to higher interview performance ratings due to higher interview motivation.

H7: The relationship between applicants' opportunity to perform perceptions and interview performance will be mediated by interview motivation.

All hypothesized relationships are summarized in the conceptual model (see Figure 1).

In addition to examining differences in opportunity to perform perceptions and its relations with honest IM, interview motivation, and interview performance ratings using a quantitative approach, we also explored in participants' own words why (or why not) probes were (or would be) helpful to them through a qualitative investigation. Below, we highlight results from an initial pilot study to create a suitable AVI and conduct two experimental studies to test our hypotheses.

Pilot Studies

The purposes of the pilot studies were to a) ensure that participants would not encounter any confusion throughout the study, b) derive an estimate of the study length, c) determine whether participants would be motivated to complete the AVI, d) assess whether a response time of 30 seconds for probes would be adequate, and e) calibrate the interview performance rating system. First, eight graduate students completed the AVI, suggested how to clarify instructions, provided an initial estimate of the study length, and confirmed that the instructions would motivate participants. A second pilot study with n = 10 Prolific participants suggested a median completion time of 28 minutes. This pilot also revealed that participants were highly motivated to complete the AVI (M = 4.80 out of 5, SD = 0.42). In addition, they perceived the initial probe response time of 30 seconds to be too short. We thus increased it to one minute for the main study. The interview performance ratings (M = 2.62, SD = 0.92) ranged from 1.33 to 4.00 out of 5.00, suggesting that the rating system facilitated a good range of ratings.

Study 1

Method

Study 1 employed an experimental between-subjects (AVI without probes, AVI with probes) design with participants recruited from Prolific to complete a mock AVI. The pre-registered hypotheses, as well as measures, materials, R code, analyses, raw data, descriptions of variables in a data file, and detailed pilot study results, can be accessed in the following Open Science Framework (OSF) project: https://osf.io/hswfu/?view_only=0e1deb32da3c4275a98091a116c829b9.

Sample

Power analyses were conducted for each hypothesis in R for the direct effects and using an R Shiny App for the indirect effects (Schoemann et al., 2017). The smallest effect sizes of

interest for all hypotheses were computed by weighted averages from relevant effects reported in the literature (see Online Supplement). To detect hypothesized effects with 80% power, 430 participants were required.

The initial sample consisted of 463 participants recruited from Prolific. In anticipation of some participants experiencing technical glitches on the AVI study platform, participants were recruited in batches. Accordingly, completion times ranged from approximately 20-30 minutes, with payments ranging from £3.60-4.50. Participants who performed extremely well (e.g., rated 14/15, given that each of the three interview questions was rated out of 5) were given bonuses that amounted to double their compensation. Participants were removed based on not consenting to the study, incorrectly answering attention/comprehension checks, not wanting the researchers to use their data, experiencing technical issues, and insufficient effort responding (see Online Supplement for data cleaning steps). The final sample consisted of 404 participants (gender: 44.67% females, 51.12 % males, 4.21% other or did not respond; age: M = 39.60 years, SD = 13.30). Most participants (58.10%) were employed full-time, and 15.71% were employed part-time. On average, participants had completed 13.10 (SD = 17.30) traditional interviews (e.g., videoconference, face-to-face) and 1.13 (SD = 2.20) AVIs; 40.40% had prior experience with AVIs.

Experimental Design and Procedure

Eligible participants were directed to the AVI platform and were randomly assigned to one of two AVI conditions: an AVI with probes or a control condition (no probes). First, participants were welcomed to the study with a brief overview of the tasks. Second, they were taken to an instructions page to test their microphone and webcam. Third, participants read and completed an informed consent form. Participants were then provided with instructions for

completing the AVI, including a statement on using the STAR approach to answering interview questions.[1] Next, participants read a job description for an Assistant Brand Manager position they were asked to imagine applying for, followed by a comprehension-check question (see Online Supplement for detailed instructions).

In the practice phase, participants familiarized themselves with the AVI; they were informed about where the interview questions/probes would be displayed and how to start/stop recording responses. Then, participants responded to three behavioral interview questions (e.g., "Describe a situation where you delegated effectively"; see Table 1). Once they recorded their responses to the interview question, the following pages they viewed depended on their condition. If participants were in the condition without probes, they were directed to the next interview question. If they were in the condition with probes, they next went to a page with the first probe. The probe condition asked two probes per interview question (6 probes total). Following the AVI, participants completed a survey assessing their reactions toward the AVI, their use of honest IM in the AVI, and demographic questions.

Probes. Probes were accompanied by a discretionary note that participants may use the probes as an opportunity to add additional/clarifying information if they deemed their initial response insufficient. This probing process was repeated for each of the subsequent probes. The probes written for the current study first targeted the tasks/actions performed in their past experiences (e.g., "Can you please tell us more about what you considered in your decisions?") and then the results of their described past behavior (e.g., "Can you please tell us more about the results of your actions?"). These two aspects are less often incorporated in interview responses (Bangerter et al., 2014, 2023; Brosy et al., 2020).

[1] Providing these instructions were necessary to rule out whether outcomes of interest were affected by the specific nature of the probes offered to participants, in addition to the mere offering of probes (versus none).

Measures

Opportunity to Perform Perceptions. Opportunity to perform perceptions were measured with a modified version of the Selection Procedural Justice Scale (Bauer et al., 2001), consistent with how applicant reactions have been operationalized in previous AVI studies (e.g., Basch, Brenner, et al., 2021; Basch et al., 2020; Basch & Melchers, 2019). Perceived opportunities to perform were measured with four items (e.g., "I could really show my skills and abilities through this interview"; ω = .93). Participants indicated their responses on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*).

Impression Management. Honest IM was measured via a modified short version of the honest IM scale (Bourdage et al., 2018). The following subscales (each with four items) were administered: *honest self-promotion* (e.g., "I made sure to talk about my job credentials"; ω = .67), *honest defensive* (e.g., "I gave an honest account of why I lacked control over past negative events that came up during the interview"; ω = .71). Participants indicated their responses on a 5-point scale ranging from 1 (*To no extent*) to 5 (*To a very great extent*), with the option to select NA (Not Applicable).

Interview Motivation. Three items from the Test Attitude Survey—Motivation subscale (TAS-M; Arvey et al., 1990) were included to measure participants' motivation levels to complete the AVI. The three items were shortlisted based on the highest-loading items of the TAS-M, as evidenced by McCarthy and Goffin (2003). Participants indicated their responses on a 5-point Likert scale ranging from 1 (*Strongly disagree*) to 5 (*Strongly agree*) for each item (e.g., "I tried to do the very best I could do on this interview"; ω = .91).

Interview Performance Ratings. Three raters independently rated the interview responses. The first author was one of the raters, and two other raters were trained through

several hours of frame-of-reference training. Behavioral anchors were provided for responses in the range of the BARS, from 1 = *Novice* to 5 = *Expert*. Accordingly, each answer was rated using these BARS. Raters discussed discrepant ratings depending on whether ratings differed by two or more points or when one rater (but not others) provided a rating of 1 or 5 (see Online Supplement).[2] Interview performance ratings were computed by averaging ratings assigned for each question and available probe they answered. The intraclass correlation (two-way random effects, absolute agreement) was computed to determine the reliability of these mean interview performance ratings: ICC (2,3) = .92 for the condition without probes and ICC (2,2) = .89 for the condition with probes.[3] Overall, the ICCs were excellent, consistent with previous recommendations suggesting values above the .75-.80 range (Koo & Li, 2016; Liljequist et al., 2019).

Qualitative Questions. With the first n = 111 participants, we solicited their experiences of the AVI using open-ended qualitative questions. In the condition with probing, participants were presented with one question: "Consider the asynchronous video interview you just completed. In general, do you think that receiving follow-up questions/prompts/probes (e.g., 'Can you please tell us more about the outcome of your actions?') after responding to interview questions in this type of interview was beneficial to you? Please make your answer as detailed as possible."

In the condition without probing, participants answered a similar question but were asked to imagine that they received follow-up questions/prompts/probes and whether this would benefit them. They then were asked, "How fair did you find the asynchronous interview process that you

[2] Readers can view the BARSs in the Online Supplement.
[3] Three raters rated interviews for the condition without probes; 2 raters rated interviews for the condition with probes.

undertook? Consider aspects of the interview that made it fair or less fair in your answer. Please make your answer as detailed as possible."

Exploratory Measures

We also measured participants' global procedural justice perceptions (Bauer et al., 2001; ω = .94), the extent to which they engaged in deceptive IM (Canagasuriam & Roulin, 2021; ω = .88), and their state interview anxiety (Tross, 1999; ω = .94). Since we had no pre-registered hypotheses about these constructs, we only present a brief overview of the results in an exploratory analyses section below.

Results

Preliminary Analyses

Preliminary analyses with Welch's t-tests revealed that the conditions did not differ in gender (male vs. female), $\chi^2(1) = 2.50$, $p = .100$. Furthermore, conditions did not differ in age or prior interview experience, all ts ≤ 0.23, all ps $\geq .300$. Descriptive statistics, scale reliabilities, and correlations among study variables are in Table 2.

Analytic Process

The main hypotheses were evaluated with structural equation modeling (SEM) using the lavaan R package. We used Mardia's test to assess for multivariate normality concerning skewness and kurtosis (Mardia, 1970, 1980) in the data, with results suggesting departures from multivariate normality (multivariate skew = 63.43, $p < .001$; multivariate kurtosis = 436.99, $p < .001$). Due to multivariate non-normality in the data, we used the diagonally weighted least squares (DWLS) lavaan estimator for SEM analyses to compute accurate estimates (Mîndrilă, 2010).[4] The DWLS estimator is recommended for data with small sample sizes, large models,

[4] We tested an alternative model using Maximum Likelihood estimation, which yielded similar results (see Online Supplement).

and skewed data (Mîndrilă, 2010). To use all possible data and to reduce the likelihood of biased estimates, missing data were handled with pairwise deletion.

A two-step procedure was used to conduct the SEM (Anderson & Gerbing, 1988). In the first step, the measurement model (i.e., a multiple-factor model with saturated covariances among the latent variables) was specified with a confirmatory factor analysis. For the first step, we used item-level data for all constructs, apart from the availability of probing condition: the four opportunity to perform perception items, the four honest self-promotion items, the four honest defensive IM items, the three interview motivation items, and the average interview performance ratings for the three interview questions.[5] Model fit was assessed with the comparative fit index (*CFI*), Tucker-Lewis index (*TLI*), root-mean-square error of approximation (*RMSEA*), and standardized root mean square residual (*SRMR*). The experimental manipulation and objective interview performance ratings mitigated some concerns about common method variance in the study. In the second step, structural paths were estimated and evaluated based on 1,000 bias-corrected and accelerated bootstrap resampling.

The hypothesized model was compared to two alternative models: the first alternative model was defined by all latent variables loading onto one factor; the second alternative model included covariances set to zero among all latent variables (i.e., a null model).[6] According to chi-square difference tests, the hypothesized model fit the data statistically and significantly better

[5] Confirmatory factor analyses suggested that the honest self-promotion and honest defensive IM subscales should not be combined into a one-factor model (i.e., all items loaded to one honest IM factor). The two-factor solution demonstrated a significant increase in model fit relative to that of the one-factor solution, evidenced by a statistically significant decrease in the two-factor model's chi-square index as the one-factor model resulted in poorer fit to the data than the two-factor model, $\Delta\chi^2(1) = 83.21$, $p < .001$.

[6] It is plausible that being offered probes allows applicants to engage in honest IM without consciously perceiving greater opportunities to perform. Thus, an alternative model was tested in which probing (versus no probing) directly predicted opportunity to perform perceptions and honest IM. Despite relatively poor fit indices in the alternative model compared to the hypothesized model, path estimates were larger in magnitude and trended in the same direction (see Online Supplement). We would like to thank an anonymous reviewer for this suggested analysis.

than the alternative models ($\Delta\chi^2 \geq 811$, *p*s < .001; Table 3). Moreover, CFI/TLI indices were higher in the hypothesized model, and χ^2, RMSEA, and SRMR indices were lower than in the alternative models (see Table 3).[7]

Hypothesis Testing

Concerning the direct effects, Welch's t-test (d = 0.39, p < .001) and SEM results (b = 0.38, 95% BCa CI [0.19, 0.56], β = 0.12, p < .001) pointed toward the positive effects of probing on participants' opportunity to perform perceptions, supporting H1. Welch's t-test (d = 0.09, p = .402) and SEM results (b = -0.02, 95% BCa CI [-0.28, 0.23], β = -0.01, p = .870) showed that probing did not increase participants' use of honest self-promotion. However, while Welch's t-test (d = 0.25 p = .018) suggested that probing did increase participants' use of honest defensive IM to a small extent, SEM results estimated no such effect (b = 0.06, 95% BCa CI [-0.12, 0.25], β = 0.05, p = .493), lending mixed support for H2. In addition, Welch's t-test (d = 0.61 p < .001) and SEM results (b = 0.44, 95% BCa CI [0.29, 0.61], β = 0.38, p < .001) showed that probing led to higher interview performance ratings, supporting H4.

Concerning the indirect effects, participants' opportunity to perform perceptions mediated the relation between the availability of probing and participants' use of honest self-promotion (indirect effect of b = 0.18, 95% BCa CI [0.09, 0.33], β = 0.12, p = .002) and honest defensive IM (indirect effect of b = 0.07, 95% BCa CI [0.03, 0.14], β = 0.06, p = .016), supporting H3. Neither honest self-promotion (indirect effect of b = -0.00, 95% BCa CI [-0.09,

[7] Although our fit indices surpassed conventional cut-offs (e.g., Hu & Bentler, 1999), the cut-offs for these indices were developed under normal-theory maximum likelihood (ML) with continuous data, and their interpretation differs due to the current study's multivariate non-normal data and use of the DWLS estimator (Nye & Drasgow, 2011). Furthermore, we followed Nye and Drasgow's (2011) approach for evaluating RMSEA and SRMR cutoffs using the DWLS estimator, finding Type I error rates nearly equal to an expected Type I error rate of 5% for the RMSEA (obtained error rate of 5.48%) and SRMR (obtained error rate of 5.74%) cutoffs obtained from our study. Overall, our hypothesized model using the DWLS estimator provides a decent fit to the data.

0.04], β = -0.00, p = .892) nor honest defensive IM (indirect effect of b = -0.01, 95% BCa CI [-0.08, 0.02], β = -0.01, p = .576) mediated the relation between the availability of probing and interview performance ratings, lending no support to H5.

Participants' opportunity to perform perceptions mediated the relation between the availability of probing and interview motivation (indirect effect of b = 0.07, 95% CI [0.03, 0.14], β = 0.06, p = .014), supporting H6. However, participants' opportunity to perform perceptions did not indirectly predict their interview performance ratings through their interview motivation (indirect effect of b = 0.01, 95% CI [-0.01, 0.04], β = 0.02, p = .264), lending no support to H7.

Robustness Check: Effect of Response Length

Because the positive effect of probing on performance ratings could be due to applicants being rewarded for providing longer answers, exploratory path analyses were conducted to ascertain whether response length mediated the relationship between the availability (versus absence) of probing and applicants' interview performance ratings. The response length was computed by summing the response lengths for each question and applicable probes. There was no direct effect of probing on interview performance ratings (b = -0.05, 95% BCa CI [-0.25, 0.15], β = -0.04, p = .634) when response length was added as a mediator to the model. Yet, probing indirectly affected interview performance ratings via response length, b = 0.47, 95% BCa CI [0.29, 0.64], β = 0.34, p < .001. These results suggest that the effect of probing on interview performance ratings can be partially explained by the additional time that probes afford applicants.

Quantitative Content Analysis

Using the NVivo software, we conducted a manifest inductive content analysis on the responses to the qualitative questions posed to the first n = 111 subset of participants. Manifest

content analysis involves describing qualitative responses at the surface level, with close attention to what respondents say (Kleinheksel et al., 2020).

The content of the responses tended to map onto Gilliland's (1993) procedural justice rules. Accordingly, codes and categories were created for these procedural justice rules and then applied when analyzing the responses. New codes and categories were created when responses did not align with Gilliland's (1993) procedural justice rules. Lastly, categories were grouped within themes to represent the underlying meaning of the responses. These themes described respondents' behaviors, experiences, or emotions in the categories and codes that typified their responses.[8] A second rater coded these responses in addition to the first author. The coding was deemed reliable, with the average Kappa (*K*) = .91, indicating almost perfect agreement between the two coders (McHugh, 2012).[9]

A large majority (86%) of respondents in the condition without probes indicated they would favor an AVI with probes. Moreover, 78% of respondents would think so because of increased performance opportunities. Furthermore, responses mentioning potential increased opportunities to perform consisted of several codes. For instance, respondents believed that probes would allow them to elaborate on their responses (59%), overcome pre-existing limited opportunities to perform (25%), organize their thoughts (15%), or showcase their personality/demonstrate their fit (12%). Interestingly, when evaluating the fair/unfair aspects of the AVI without probes, a high proportion of respondents (71%) described the AVI as unfair, and 66% described the AVI as fair—including 41% of respondents mentioning both fair and unfair aspects of the AVI, yielding a mixed evaluation of fairness. Furthermore, 45% of

[8] The code guide and illustrative quotes for the quantitative content analysis can be viewed in the Online Supplement.
[9] Kappa values between .60 to .79 = moderate agreement, .80 to .90 = substantial agreement, and > .90 = almost perfect agreement (McHugh, 2012).

respondents noted that the AVI gave them lower opportunities to perform, underscoring the importance of job applicants' desire for more opportunities to perform in the AVI.

Similar to the condition without probes, a large proportion (83%) of respondents in the condition with probes favored the AVI with probes. Moreover, 77% of participants did so due to increased performance opportunities. For instance, they mentioned that probes allowed them to elaborate on their responses (73%), organize their thoughts (12%), overcome pre-existing limited opportunities to perform (12%), or showcase their personality/demonstrate their fit (8%). This pattern of results is consistent with the quantitative result that probes increased perceived opportunities to perform in the AVI.

Exploratory Analyses

We explored participants' usage of probes and how probing affected their global procedural justice perceptions, state interview anxiety, deceptive IM, and interview performance ratings in the condition with probing. The Online Supplement elaborates on the details of these exploratory analyses.

Frequency of Probe Use. We explored the extent to which participants used the probes in the study. 90% of participants in the AVI condition with probes responded to the two probes provided for each interview question, and 98% responded to at least one. Furthermore, 11% of participants used the maximum allotted probe response time of one minute when examining the proportion of allotted time used in the probes. On average, they used 77% of the allocated time (*SD* = 20%), with a median of 80%. Few participants (22%) used half or less of the allotted time. We conducted path analyses with the participants' average use of probes as a predictor instead of the availability of probes. We found similar results to those of the SEM analyses presented in the current study.

Global Procedural Justice Perceptions, Interview Anxiety, and Deceptive IM. We also found that probing led to differences in global procedural justice perceptions (p = 0.020, d = 0.23) but no differences in the extent to which participants experienced interview anxiety (p = 0.200, d = -0.14). Participants with probes did engage in incrementally more slight image creation (p = 0.020, d = 0.26) but not more/less in extensive image creation (p = 0.500, d = 0.06) or image protection (p = 0.900, d = -0.01).

Interview Performance Ratings in the Condition with Probes. Interview performance ratings for participants in the condition with probes were assigned twice: once based on the initial response to a given interview question (but before answering the probes) and then based on the entire response (including the initial response and the response to the follow-up probes). The mean interview performance ratings differed between the two kinds of ratings, $p < .001$, d = 0.52, with 77% of these participants experiencing improvements in their ratings after responding to the probes; as little as 10% of participants experienced reductions in interview performance ratings after responding to the probes. We also examined whether differences in ratings before versus after probes were consistent across the full range of the employed interview performance rating scale. Interestingly, the largest differences in interview performance ratings were observed for applicants who initially scored between 2 and 3 on a 5-point scale before (versus after) being offered probes, d = 1.19, 95% CI [0.90, 1.48]; a very large effect (Bosco et al., 2015). This effect is comparatively larger than the effects observed for participants who initially scored between 3 and 4 before (versus after) being offered probes, d = 0.94; or for participants who initially scored between 4 and 5 before (versus after) being offered probes, d = 0.23.[10]

Study 2

[10] We thank an anonymous reviewer for suggesting this exploratory analysis.

A key limitation of our initial study was that our robustness check with response length suggested that the effect of probing on interview performance ratings can be partially explained by the additional time that probing affords applicants. Accordingly, Study 2 explored whether the effects of probes on opportunity to perform perceptions and interview ratings were due to their unique nature, as opposed to the additional response time interviewees get through probes.[11]

Method

Sample

Power analyses were conducted based on the opportunity to perform (d = 0.39) and interview ratings (d = 0.61) effects obtained in Study 1. To detect these effects with 80% power, 258 participants were required. The initial sample comprised 295 Americans from Prolific who did not participate in Study 1 and were recruited in batches to address potential technical glitches with the AVI platform. Participants were compensated £3.00, and median completion times ranged from approximately 20-26 minutes. We used a similar data screening process as in Study 1, resulting in a final sample of 271 participants (47.19% females, 50.19% males, 2.62% other or preferred not to answer; age: M = 38.53 years, SD = 12.51). Most participants (57.30%) were employed full-time, and 22.89% were employed part-time. On average, participants had completed 13.11 (SD = 17.44) traditional interviews (e.g., videoconference, face-to-face) and 2.09 (SD = 4.31) AVIs; 52.08% had prior experience with AVIs.

Experimental Design and Procedure

Study 2 design and procedure were consistent with Study 1, with the following exceptions: First, participants were randomly assigned to one of three AVI conditions: a control condition without probes and two minutes to answer (89 participants), a condition where

[11] We thank an anonymous reviewer for suggesting this exploratory analysis.

responses to each question were followed by two probes with an extra one minute per probe (93 participants), or a no probing condition where the response time was configured at four minutes per question to match the cumulative response time of the condition with probing (89 participants). Second, following the AVI, participants completed measures of global procedural justice perceptions (Bauer et al., 2001), opportunity to perform perceptions (Bauer et al., 2001), honest IM (Bourdage et al., 2018), and demographic questions. Distinct from Study 1, we did not measure motivation to perform nor deceptive IM but asked participants about participants' attitudes toward the response time they were given (e.g., "We gave you 120 seconds (2 minutes) of time to respond to each interview question. Reflecting on your responses to the interview questions, how did you find this amount of time?") and preference for longer response time, probes, or neither. Interview performance was assessed by one of four trained research assistants who watched video responses and scored each response using 1-5 BARS. Inter-rater reliability was assessed by having all four raters assess the videos of 35 randomly selected participants from the condition without probes, ICC (1,4) = .96, considered excellent (Koo & Li, 2016; Liljequist et al., 2019).

Results

Preliminary analyses with one-way ANOVAs revealed that the experimental groups did not differ in gender, $\chi^2(10) = 10.46$, $p = .410$, nor in age or prior interview experience, all Fs $\leq$ 1.14, all ps $\geq .320$. Table 4 shows descriptive statistics, scale reliabilities, and correlations.

Opportunity to Perform Perceptions

We ran a one-way analysis of variance (ANOVA) to examine the extent to which the condition (AVI without probes, AVI with probes, AVI matching response time of AVI with probes) influenced opportunity to perform perceptions. Levene's test did not suggest population

variances were different, $F(2, 265) = 1.46$, $p = .234$. Results of the one-way ANOVA suggested no differences in opportunity to perform, $F(2, 265) = 1.95$, $p = .146$, $MSE = 1.05$, partial $\eta^2 = .01$, 95% [.00, .04]. Pairwise comparisons revealed effects that were generally in line with the main study, albeit weaker. For instance, opportunity to perform was not significantly higher in the AVI with probes ($M = 3.52$, $SD = 0.94$) than in the control ($M = 3.37$, $SD = 1.03$), $p = .315$, $d = 0.15$, 95% CI [-0.14, 0.44]. Interestingly, the AVI with equivalent response time ($M = 3.22$, $SD = 1.11$) was slightly lower than the control, $p = .348$, $d = -0.14$, 95% CI [-0.44, 0.15], demonstrating that not even added response time could enhance perceived opportunities to perform. In addition, we found a marginally higher opportunity to perform effect in the AVI with probes than in the AVI with equivalent response time, $p = .052$, $d = 0.29$, 95% CI [-0.00, 0.59]. Altogether, despite the weak effects of probes, these results suggest that opportunity to perform perceptions are not driven by the added response time that probes afford.

We further sought to understand participants' attitudes toward the response time they were provided across conditions. Most participants across conditions noted that the available response time was "just enough." Yet, a considerable proportion of participants in the equivalent response time condition (39.78%) thought the four minutes of response time/question was "too long," which was significantly larger than the proportion of participants who viewed the two minutes of response time/question in the condition with probes (9.68%) as "too long", $\chi^2(1) = 20.65$, $p < .001$, $\varphi = 0.34$, and the condition without probes (10.34%), $\chi^2(1) = 18.60$, $p < .001$, $\varphi = 0.34$.

We also asked participants if they would have preferred to have had probes, a longer response time, or neither of these options. Most participants preferred probes, including in the equivalent response time condition (65.91%), the probes condition (52.69%), and the control

condition (50.57%)—although these proportions did not significantly differ, $\chi^2(1) = 2.75$ and 3.62, $p = .098$ and .057, $\varphi = 0.12$ and 0.14, respectively. These results thereby provide some nuance for the slightly weaker opportunity to perform effect in the AVI with equivalent response time relative to the other conditions, as the AVI with probes was preferred over an increased response time when responding to past-behaviour questions.

Interview Performance Ratings

We ran a one-way analysis of variance (ANOVA) to examine the extent to which the condition influenced interview performance ratings. Levene's test did not suggest population variances were different, $F(2, 246) = 2.69$, $p = .070$. Results of the one-way ANOVA showed differences in interview performance ratings across conditions, $F(2, 246) = 2.69$, $p = .007$, $MSE = 0.72$, partial $\eta^2 = .04$, 95% [.01, .08]. Pairwise comparisons revealed effects in line with Study 1, albeit slightly weaker. Further supporting H4, interview performance ratings were higher in the condition with probes ($M = 2.71$, $SD = 0.81$) relative to the condition without probes ($M = 2.32$, $SD = 0.77$), $p = .002$, $d = 0.49$, 95% CI [0.18, 0.79], or the condition equivalent response time ($M = 2.38$, $SD = 0.96$), $p = .021$, $d = 0.36$, 95% CI [-0.06, 0.67]. Furthermore, performance was not different in the condition without probes and the condition with equivalent response time, $p = 0.645$, $d = 0.07$, 95% CI [-0.23, 0.38]. Altogether, these latter results suggest that improvements in interview performance were driven mainly by the unique nature of probes, above and beyond the additional response time that probes provide.

General Discussion

AVIs are being increasingly adopted in hiring processes (e.g., Jaser et al., 2022; Reynolds, 2022; Woods et al., 2020). AVIs provide multiple advantages to both applicants and organizations, such as reducing costs (Mejia & Torres, 2018b), improving standardization, and

offering greater flexibility regarding when, where, and how applicants complete the interview. However, applicants sometimes meet AVIs with negative reactions. One understudied benefit of AVIs that is of particular interest to the current study is the customizability of AVI design (Lukacik et al., 2022). Building on Lukacik et al.'s (2022) arguments that AVI design features can impact applicants' reactions, behaviors, and performance, we examined the effects of structured probing in AVIs.

Main Findings and Theoretical Implications

In Study 1, we found that probing directly enhanced opportunity to perform perceptions and interview performance ratings, with the effect of probing on interview performance ratings being driven by response length. Moreover, probes positively and indirectly affected honest IM through perceived opportunities to perform. Study 2 further replicated the direct effects of probing on opportunity to perform perceptions and interview performance ratings in the same direction, albeit especially weaker for perceived opportunities to perform. In contrast, simply increasing response time had no effects on opportunity to perform perceptions and interview performance ratings.

Opportunity to Perform

In terms of perceived opportunities to perform, our findings revealed mixed effects. In Study 1, we found that probing led to moderate increases in opportunity to perform perceptions (d = 0.39). These perceptions, in turn, were associated with higher honest IM and interview motivation. However, it is unclear whether the effect of probing on perceived opportunities to perform was due to probes themselves or the fact that probes simply give interviewees more time to provide longer responses. Accordingly, we experimentally teased apart added response time from probes in Study 2. Despite finding a weaker effect of probing (versus no probing) on

opportunity to perform perceptions (d = 0.15) than in Study 1, the additional response time (versus no probing) had a similarly small effect but in the opposite direction on opportunity to perform perceptions (d = -0.14). Similarly, Tilston et al. (2024) examined how various AVI design features affect applicant reactions, finding that maximum response time was unrelated to applicants' reactions (r = -.03). Their findings suggest that AVI design features that include added response time do not necessarily have consistent effects on applicant reactions. Interestingly, the largest opportunity to perform effect in Study 2 was observed when comparing probes to equivalent response time (d = 0.29). Further, results suggested that participants found that the equivalent response time was too long compared to participants in other conditions and that all participants preferred having probes over a longer response time. Altogether, our findings suggest that probes may have unique properties besides their added response time that positively affect applicant reactions, such as perceived performance opportunities.

Our preliminary findings regarding perceived opportunities to perform are consistent with the proposition that AVI design features could increase applicants' opportunities to perform (Lukacik et al., 2022). For instance, just like granting applicants more preparation time, our findings suggest that having probes allows applicants to incorporate additional or clarifying information. Our findings that probing led to small-moderate increases in opportunity to perform perceptions may suggest greater utility than other AVI design features. For example, Basch, Brenner, et al. (2021) reported that participants experienced slightly greater opportunity to perform (d = 0.17) when given additional preparation in the AVI. However, this effect is descriptively smaller than some of the effects of probing observed in our study.

The availability of probes may have increased opportunity to perform perceptions by reducing the feelings of uncertainty or ambiguity that applicants might have about their AVI

performance (Mejia & Torres, 2018a; Muralidhar et al., 2020).[12] In face-to-face or videoconference interviews, applicants might reduce such uncertainty concerning their performance by observing interviewer signals, such as their non-verbal behavior (e.g., eye contact, slight nodding, facial expressions), which could potentially affirm applicants' perceptions of whether they have provided sufficient responses. Due to the impersonal nature of AVIs (Guchait et al., 2014), their lack of non-verbal feedback may reinforce applicants' experiences of uncertainty, driving them to seek clarity and reassurance. Consequently, applicants might respond to probes as a compensatory mechanism to replace the non-verbal cues typically available in face-to-face or videoconference interview mediums. Thus, probing may help reassure and provide applicants with implicit feedback, reducing their uncertainty and boosting their opportunities to perform in the AVI.

Consistent with the SEM results, the quantitative content analysis may shed light on the unique benefits of probing relative to other AVI design features. The results suggested that most participants in the condition without probes would favor an AVI with probes, largely because of the greater opportunities they would have to perform. Interestingly, when asked about the aspects that made the AVI fair/unfair, a substantial proportion of participants felt the AVI was not fair because of the lack of opportunities they had to perform, suggesting that probes could alleviate this concern. Indeed, participants provided with probes largely favored this AVI format because probes gave them better opportunities to demonstrate their knowledge, skills, and other abilities. Moreover, participants voiced that probes allowed them to elaborate on their responses, organize their thoughts, overcome pre-existing limited opportunities to perform, showcase their personality, or demonstrate their fit. These exploratory findings suggest that probes should be

[12] We thank an anonymous reviewer for suggesting this uncertainty-reduction effect of probing.

offered to applicants in AVIs, as they can increase applicants' perceptions of opportunity to perform—ultimately enabling them to feel like they can demonstrate the best of their knowledge, skills, and other abilities.

Honest Impression Management

The availability (versus absence) of probes led participants to engage in negligibly more honest self-promotion (d = 0.09) but slightly more honest defensive IM (d = 0.25). These findings lend weak support to theoretical arguments by Lukacik et al. (2022), who proposed that certain AVI design features (e.g., preparation time) may influence the extent to which applicants use self-focused IM. The effects of probing are also somewhat smaller than those of Basch, Brenner, et al. (2021), who reported that participants used more honest IM (d = 0.56) when given additional preparation time in the AVI. Our findings are comparable to those of Roulin, Wong et al. (2023), who found that offering participants more preparation time only led to slight increases in honest IM (d = 0.13). Yet, our research expands on those prior studies by demonstrating that participants' opportunity to perform perceptions may be a mechanism for why probes led to slight increases in honest IM. When applicants received probes, they perceived greater opportunities to demonstrate their knowledge/skills/abilities and, in turn, used more honest self-promotion and honest defensive IM. This effect is consistent with Bourdage et al. (2018), who proposed that applicants who perceive interviews as procedurally just will engage in more honest IM, as they should be more willing and feel capable of relying on their true qualifications.

Interview Motivation

Gilliland's (1993) model of applicant reactions proposed that applicants' fairness perceptions impact their test-taking motivation. Further, Lukacik et al. (2022) proposed that fairness perceptions mediate the relationship between AVI design decisions and applicants'

motivation to perform. Our findings provide empirical support to these theoretical propositions as participants' opportunity to perform perceptions was a mechanism for why the availability (versus absence) of probes was associated with higher interview motivation.

Interview Performance Ratings

The availability (versus absence) of probes also led participants to receive higher interview performance ratings (*ds* = 0.49 to 0.61) and even beyond increased response time (*d* = 0.36). Similarly, Basch, Brenner, et al. (2021) reported that participants received higher interview performance ratings (*d* = 0.68) when given additional preparation time in the AVI. Roulin, Wong et al. (2023) found that offering participants more preparation time did not improve their interview performance ratings (*d* = 0.01), albeit the use of this preparation time was slightly more positively related to performance ratings (*r* = 0.14). Our Study 2 finding that probing has performance benefits above and beyond added response time may suggest unique properties of probes. With more response time, applicants may not always use the extra time effectively to enhance the quality or relevance of their responses. Moreover, applicants might fill additional response time with redundant or tangential information that does not directly improve the evaluative content of their answers. On the other hand, probes can provide targeted prompts that encourage applicants to delve deeper into relevant competencies or provide specific examples, thereby increasing the likelihood of higher-quality responses. This guidance may help applicants structure their responses, allowing them to showcase their skills more strategically and directly address the competencies assessed, which could explain the observed performance gains beyond the addition of response time.

Exploratory Analyses

We explored participants' usage of probes and the impact of probing on interview anxiety, deceptive IM, and interview performance ratings before versus after probing. First, participants used the probes to a large extent. This usage effect is interesting, as other AVI studies reported that applicants did not use available AVI design features to such a high extent. For instance, Dunlop et al. (2022) reported that only 39.8% of job applicants having completed AVIs used re-recording attempts at least once in their interviews. In addition, Roulin, Wong, et al. (2023) reported that when provided with up to five recording attempts (i.e., one initial response and four re-recording attempts), participants, on average, used 1.48 of those attempts (i.e., 29.60% of total recording attempts used), or 0.48 of four re-recording attempts (i.e., 12.00% of re-recording attempts used). This means that most participants in their study did not use additional recording attempts. Moreover, Dunlop et al. (2022) reported that applicants would use from 34.8% to 76.7% of their allotted preparation time, and Roulin, Wong, et al. (2023) reported that when provided with unlimited preparation time, participants on average used 85.48 seconds of that time. By comparison, participants in our study used nearly all probes across the interview questions on average. Perhaps probes are used more often than re-recording attempts or preparation time because probes provide direct cues for applicants to augment their insufficient or inadequate responses. Conversely, with re-recording attempts, applicants have to weigh the benefits of potentially improving their response versus the risks of repeating already stated information and the costs of dedicating more time and cognitive effort to the AVI. In addition, spending time preparing for a response does not guarantee that one will receive a favorable evaluation.

In addition, probing minimally affected applicants' interview anxiety (d = -0.14) and led to small-medium differences in one kind of deceptive IM—slight image creation (d = 0.26) but

did not impact the use of extensive image creation or image protection. These findings are only partially consistent with those of Levashina and Campion (2007), who found that probing increased all types of deceptive IM in face-to-face interviews. From before to after being offered probes, participants in the probing condition experienced moderate-to-large improvements in their interview performance ratings. This effect was most pronounced for participants who initially obtained average scores (i.e., in the middle range of the BARS used to assess their performance).

Overall, while probing had minimal effects on interview anxiety and only modestly influenced slight image creation as a form of deceptive IM, its high usage by applicants and significant impact on improving interview performance ratings suggest that probing has unique effects, perhaps compared to other AVI features like re-recording attempts or preparation time.

Limitations and Future Research Directions

Although the current study has some degree of internal validity with its randomized experimental design, it has some limitations. As our studies tested the effect of semi-generic probes relative to no probing while controlling for additional response time in Study 2, we provide evidence for a rather conservative test of probing on perceived opportunities to perform and interview performance ratings. Accordingly, it may be fruitful for future research to try and tease apart the effects of the semi-generic probes we employed in our studies versus generic probes (e.g., “Do you have anything else you want to add?”) or other AVI response formatting features. For instance, an experimental design where participants are randomly assigned to a probing condition, a re-recording condition, and a probing x re-recording condition for the same interview questions would allow for this direct comparison. Thus, although probes likely have nuanced effects on the applicant experience in AVIs, future research is needed to establish their

distinct effects by experimentally comparing them with other types of probes or response formatting features.

Unexpectedly, engagement in honest IM and interview motivation did not appear to be mechanisms for why probes led to higher interview performance ratings. One alternative explanation for the performance-enhancing benefits of probing may relate to the cognitive demands of interviewing. Specifically, Behavioral Description Interviews (BDIs) can place a high cognitive demand on applicants (Bangerter et al., 2014), which may lead applicants to tell general pseudo-stories (Brosy et al., 2020) rather than detailed descriptions of specific past events. Probing may alleviate the high cognitive load typically experienced by applicants during BDIs by providing cues that assist them in structuring their responses to provide details about the tasks/actions they took and the results they obtained. Moreover, probes can aid applicants in recalling specific details from their experiences and allow them to elaborate on what they have already described. This process, in turn, could enable them to narrate more detailed stories that lead to more favorable evaluations (Brosy et al., 2020). Thus, one key difference between AVI features, such as preparation time or response time, and the probes used in the current study is that probing can help cue applicants' memory (Murphy, 2020), reducing their cognitive load. Nonetheless, future research could directly examine what cognitive mechanisms are at play (e.g., cognitive load, difficulty) when comparing the benefits of probes relative to other design features in AVIs.

Given that probing had relatively larger effects on interview performance ratings than perceived opportunities to perform, even compared to no probing and additional response time, future research could examine differential mechanisms for these effects. One possible explanation for our relatively weaker opportunity to perform effects is that our use of structured

probing may have limited applicants' sense of autonomy during the AVI. While structured probes guide applicants to enhance their answers, they may also inadvertently reduce the sense of control applicants have over how to organize and present their responses. As a result, applicants might feel they are responding to predefined prompts rather than freely demonstrating their abilities. This could explain why the perceived opportunities to perform were only moderately impacted by probing; although applicants might recognize the performance benefits of probing, they may not feel they can fully showcase their competencies on their terms. Consequently, the structured direction provided by probes may be perceived as a facilitator of better answers rather than an opportunity to demonstrate one's full potential freely.

Likewise, we found that our use of standardized probing can sometimes conflict with applicants who felt that responding to the standardized probes was unnecessary due to the quality of their initial response. Moreover, our quantitative content analyses suggested that a small proportion of participants felt the probes were confusing or redundant when they felt their initial response to the interview question was sufficient (e.g., "I'd say not all the follow-up questions made complete sense with the answers I gave but I feel like that's to be expected"—see Online Supplement). However, one benefit of the current study's use of standardized probes is that they reconcile the tension between the detriments that unique, personalized probes have on interview validity versus the added structure that standardization provides to the interview (Campion et al., 1997). Future research might consider investigating how probes can be administered in AVIs without applicants feeling like probes are redundant or confusing.

Despite finding that probing indirectly increased honest IM through opportunity to perform perceptions, research also suggests that differences in interview medium or AVI features, such as preparation time, affect opportunity to perform and even fairness perceptions

broadly through IM (Basch, Brenner, et al., 2021; Basch et al., 2020; Basch, Melchers, et al., 2021). Results from these studies suggest that IM precedes opportunity to perform and fairness perceptions, relying on the argument that the interview medium or AVI design features can promote/constrain opportunities to engage in IM, translating into perceived opportunities to perform or fairness perceptions proportionately (Blacksmith et al., 2016). For instance, Basch, Brenner, et al. (2021) found that preparation time in AVIs indirectly promoted fairness perceptions through honest IM. We relied on Bourdage et al.'s (2018) argument that applicants should be more willing to highlight their true qualifications when interview procedures are perceived as fair. Nonetheless, these previous studies and ours measured the mediator and outcome concurrently, making it difficult to establish the causality of these relations. It follows that future research that experimentally manipulates IM and fairness perceptions can help establish the causal nature of their relationships.

Our AVI had three questions, which may be considered a short interview since AVI templates, on average, have 4.48 questions (Dunlop et al., 2022). Thus, participants' high use of probes could be an artifact of the study's short AVI; perhaps applicants already felt limited by the short AVI and felt that using the probes would help maximize their performance opportunities. Yet, applicants might find responding to probes for longer interviews more effortful (or repetitive). Future research should investigate whether probes have the same effects when used in interviews with more questions.

Our sample included participants from Prolific who completed a mock job interview, limiting the ecological validity of the study. Specifically, participants may be less motivated in studies with no high-stakes job interviews. However, the participants in this study did appear motivated to perform in the AVI (M = 4.61 out of 5, SD = 0.59). We attribute this high

motivation to the incentives given to participants, which may resemble the applicant's experience in high-stakes selection situations. Nonetheless, future research can benefit from the bolstered ecological validity associated with field samples and should investigate the effects of probing in high-stakes AVIs with actual job applicants. Furthermore, we suggest that future studies explore the impact of probes in AVIs across various applicant populations and job types.

The probes in our study inquired about participants' tasks/actions and the results of their past experiences, which paired well with the BDI nature of the AVI. Although the effects of probes in the study may generalize to other AVIs of a similar format, they may not generalize to AVIs posing other types of interview questions to applicants. Future research may investigate the effects/effectiveness of probing in AVIs using situational or technical questions.

With the explosion of generative artificial intelligence (AI) and large language models in personnel selection (Campion & Campion, 2023), specific areas of applicant responses that are insufficient or need clarification can be analyzed in real-time, with improvised probes generated on the spot. One advantage of these automated probes is that they can address concerns about whether probes are relevant to applicants. AVI providers can train generative AI models with several parameters and features on datasets consisting of applicant responses to various interview questions assessing different skills and competencies. As a result, if interview responses highlight knowledge, skills, and characteristics underscoring a competency that requires clarification or more detail, AI could generate data-driven follow-up questions to probe deeper into those knowledge, skills, and characteristics. Furthermore, these data-driven and generated probes and their parameters can be adjusted to suit an organization's preferences. At the expense of generating these tailored probes, such highly automated interviews tend to be less accepted by applicants (e.g., Langer et al., 2019). Nevertheless, these individualized probes may give

organizations a more accurate picture of applicants' competency levels. Thus, future research is needed to assess the effects of automated probing in AVIs, along with their advantages and drawbacks.

Although the current study's findings suggest that standardized probing can enhance the applicant experience, future research may consider investigating how probes may impact the validity of AVIs before being implemented at a large scale. Given the superior reliability and validity of structured interviews over unstructured interviews where personalized probes are likely in the latter (Levashina et al., 2014), using standardized probes may bolster the validity of AVIs, increasing the signal-to-noise ratio in interview responses. For instance, as probes have applicants elaborate on specific parts of their responses, applicants may provide more detailed, job-relevant content that aids in accurately assessing their competency levels. Conversely, standardized probing could also restrict the variance in applicants' answers and AVI performance. Indeed, our exploratory analyses suggested that standardized probing may improve mediocre responses to a larger extent than high-quality responses. Consequently, many applicants may end up providing better-than-average responses, thereby placing a ceiling effect that diminishes an employer's ability to differentiate effectively between applicants with varying levels of true competence for the role. However, future research is needed to ascertain the effects of probing on the validity of AVIs.

Practical Implications

The current study has several important practical implications. Our preliminary findings suggest that offering probes in AVIs can improve applicants' AVI experiences. For instance, applicants can perceive more opportunities to demonstrate their knowledge/skills/abilities. In turn, they may be more motivated to perform their best and be more likely to use honest IM to

communicate their true qualifications. In addition, probing in AVIs seems to have minimal effects on deceptive IM. Organizations should concern themselves with applicants' AVI experiences because perceptions of fairness can affect applicants' attraction to organizations (Lukacik et al., 2022; McCarthy et al., 2017). Applicants with a negative experience during selection may take their views to social media and make the public aware of their poor experiences (e.g., Molton, 2021), which may lead other potential applicants to decline invitations to complete the AVI (Lukacik et al., 2022). Overall, organizations might consider implementing probes in AVIs based on this design feature's positive effects on the applicant experience but should be cautious of their implementation due to the unknown effects of probes on the predictive validity of AVIs.

Implementing probes in AVIs can be as simple as creating follow-up prompts for each interview question. In other words, just as organizations can create interview questions in the AVI, follow-up questions/probes can be created for each interview question of interest. Given that organizations tend to design AVIs that closely resemble the default options set up by AVI vendors (Dunlop et al., 2022), AVI vendors can create default probes in the interview templates that organizations leverage. As the current study administered a structured interview with past behavioral questions, general probes were developed using a storytelling framework. The probes in the study were designed to encourage applicants to discuss more of the tasks/actions performed and results of previous experiences they described, which are often neglected in interview responses to questions that inquire about one's past behavior (Brosy et al., 2020). When using past-behavioral interview questions, organizations might consider administering AVIs with probes that expand on applicants' past experiences. Furthermore, organizations using AVIs could benefit from pre-testing their preliminary probes to evaluate their advantages and

drawbacks. Moreover, before their wide-scale implementation, we recommend that practitioners refine probes according to the nature of the questions they are nested within.

Conclusion

The personnel selection field has observed the increasing growth and use of AVIs (Lukacik et al., 2022). Despite the potential benefits AVIs bring applicants, past research has found that applicants sometimes react negatively to AVIs. The present study advances the recent conceptual work by Lukacik et al. (2022), who proposed that certain AVI design features can impact interview and post-interview outcomes through applicants' reactions and behaviors. The present study introduces probing as a novel design feature in AVIs and is one of the first to evaluate the effects of probing on the applicant experience. Our preliminary findings revealed that AVIs with probes benefit applicants by slightly to moderately improving their reactions toward AVIs, indirectly increasing honest IM and interview motivation and directly increasing their interview performance ratings to a moderate extent. Organizations might consider administering AVIs with probes, especially when such AVIs pose past-behavioral questions; likewise, applicants should take advantage of these opportunities to perform their best. Nonetheless, future research is warranted to investigate the effects of probes on the validity of AVIs before their implementation.

References

Amaral, A. A., Powell, D. M., & Ho, J. L. (2019). Why does impression management positively influence interview ratings? The mediating role of competence and warmth. *International Journal of Selection and Assessment*, *27*(4), 315–327. https://doi.org/10.1111/ijsa.12260

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-step approach. *Psychological Bulletin*, *103*(3), 411–423. https://doi.org/https://doi.org/10.1037/0033-2909.103.3.411

Bangerter, A., Corvalan, P., & Cavin, C. (2014). Storytelling in the selection interview? How applicants respond to past behavior questions. *Journal of Business and Psychology*, *29*(4), 593–604. https://doi.org/10.1007/s10869-014-9350-0

Bangerter, A., Mayor, E., Muralidhar, S., Kleinlogel, E. P., Gatica-Perez, D., & Schmid Mast, M. (2023). Automatic identification of storytelling responses to past-behavior interview questions via machine learning. *International Journal of Selection and Assessment*, *31*. https://doi.org/10.1111/ijsa.12428

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., Brenner, F., Melchers, K. G., Krumm, S., Dräger, L., Herzer, H., & Schuwerk, E. (2021). A good thing takes time: The role of preparation time in asynchronous video interviews. *International Journal of Selection and Assessment*, *29*(3–4), 378–392. https://doi.org/10.1111/ijsa.12341

Basch, J. M., & Melchers, K. G. (2019). Fair and flexible?! Explanations can improve applicant reactions toward asynchronous video interviews. *Personnel Assessment and Decisions, 5*(3), 1–11. https://doi.org/10.25035/pad.2019.03.002

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

Basch, J. M., Melchers, K. G., Kurz, A., Krieger, M., & Miller, L. (2021). It takes more than a good camera: Which factors contribute to differences between face-to-face interviews and videoconference interviews regarding performance ratings and interviewee perceptions? *Journal of Business and Psychology*, *36*(5), 921–940. https://doi.org/10.1007/s10869-020-09714-3

Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001). Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, *54*(2), 387–419. https://doi.org/10.1111/j.1744-6570.2001.tb00097.x

Bauer, T. N., Truxillo, D. M., Tucker, J. S., Weathers, V., Bertolino, M., Erdogan, B., & Campion, M. A. (2006). Selection in the information age: The impact of privacy concerns and computer experience on applicant reactions. *Journal of Management*, *32*(5), 601–621. https://doi.org/10.1177/0149206306289829

Bell, B. S., Wiechmann, D., & Ryan, A. M. (2006). Consequences of organizational justice expectations in a selection system. *Journal of Applied Psychology, 91*(2), 455–466. https://doi.org/10.1037/0021-9010.91.2.455

Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the Employment Interview: A Meta-Analysis and Future Research Agenda. *Personnel Assessment and Decisions*, *2*(1), 2. https://doi.org/10.25035/pad.2016.002

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

Brosy, J., Bangerter, A., & Ribeiro, S. (2020). Encouraging the production of narrative responses to past-behaviour interview questions: effects of probing and information. *European Journal of Work and Organizational Psychology*, *29*(3), 330–343. https://doi.org/10.1080/1359432X.2019.1704265

Campion, M. A., & Campion, E. D. (2023). Machine learning applications to personnel selection: Current illustrations, lessons learned, and future research. *Personnel Psychology*. https://doi.org/10.1111/peps.12621

Campion, M. A., Palmer, D. K., & Campion, J. E. (1997). A review of structure in the selection interview. *Personnel Psychology*, *50*(3), 655–702. https://doi.org/10.1111/J.1744-6570.1997.TB00709.X

Canagasuriam, D., & Roulin, N. (2021). The effect of organizational culture on faking in the job interview. *Personnel Assessment and Decisions*, *7*(1), 83–93. https://doi.org/10.25035/pad.2021.01.008

Cropanzana, Bowen, & Gilliland. (2007). The Management of Organizational Justice. In *Source: Academy of Management Perspectives* (Vol. 21, Issue 4).

Dipboye, R., Wooten, K., & Halverson, S. (2004). Behavioral and situational interviews. In *Comprehensive handbook of psychological assessment: Industrial and organizational assessment* (Vol. 4, pp. 297–316). Wiley.

Gilliland, S. W. (1993). The perceived fairness of selection systems: an organizational justice perspective. *Academy of Management Review*, *18*(4), 694–734. https://doi.org/10.2307/258595

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

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, *57*(3), 639–683. https://doi.org/10.1111/J.1744-6570.2004.00003.X

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*, *53*(2), 164–174. https://doi.org/10.1037/cbs0000223

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huffcutt, A. I. (2011). An empirical review of the employment interview construct literature. *International Journal of Selection and Assessment*, *19*(1), 62–81. https://doi.org/10.1111/J.1468-2389.2010.00535.X

Huffcutt, A. I., & Arthur, W. (1994). Hunter and Hunter (1984) revisited: Interview validity for entry-level jobs. *Journal of Applied Psychology, 79*(2), 184–190. https://doi.org/10.1037/0021-9010.79.2.184

Jansen, A., König, C. J., Stadelmann, E. H., & Kleinmann, M. (2012). Applicants' self-presentational behavior: What do recruiters expect and what do they get? *Journal of Personnel Psychology, 11*(2), 77. https://doi.org/10.1027/1866-5888/a000046

Jaser, Z., Petrakaki, D., Starr, R., & Oyarbide-Magaña, E. (2022, January 22). Where automated job interviews fall short. *Harvard Business Review*. https://hbr.org/2022/01/where-automated-job-interviews-fall-short

Kessler, R. (2006). *Competency-based interviews*. Career Press.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Langer, König, & Papathanasiou. (2019). Highly automated job interviews: Acceptance under the influence of stakes. *International Journal of Selection and Assessment*, *27*(3), 217–234. https://doi.org/10.1111/IJSA.12246

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., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature. *Personnel Psychology, 67*(1), 241–293. https://doi.org/10.1111/peps.12052

Liljequist, D., Elfving, B., & Roaldsen, K. S. (2019). Intraclass correlation – A discussion and demonstration of basic features. *PLoS ONE*, *14*(7). https://doi.org/10.1371/journal.pone.0219854

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). https://doi.org/10.1016/j.hrmr.2020.100789

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika*, *57*(3), 519–530. https://doi.org/10.1093/BIOMET/57.3.519

Mardia, K. V. (1980). 9 Tests of unvariate and multivariate normality. *Handbook of Statistics*, *1*, 279–320. https://doi.org/10.1016/S0169-7161(80)01011-5

McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant perspectives during selection: A review addressing "So what?,""What's new?," and "Where to next?" *Journal of Management*, *43*(6), 1693–1725. https://doi.org/10.1177/0149206316681846

McCarthy, J. M., & Goffin, R. D. (2003). Is the test attitude survey psychometrically sound? *Educational and Psychological Measurement*, *63*(3), 446–464. https://doi.org/10.1177/0013164402251038

McCarthy, J. M., Hrabluik, C., & Jelley, R. B. (2009). Progression through the ranks: Assessing employee reactions to high-stakes employment testing. *Personnel Psychology*, *62*(4), 793–832. https://doi.org/10.1111/J.1744-6570.2009.01158.X

McHugh, M. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://hrcak.srce.hr/89395

Mejia, C., & Torres, E. N. (2018a). Implementation and normalization process of asynchronous video interviewing practices in the hospitality industry. *International Journal of Contemporary Hospitality Management*, *30*(2), 685–701. https://doi.org/10.1108/IJCHM-07-2016-0402/FULL/PDF

Mejia, C., & Torres, E. N. (2018b). Implementation and normalization process of asynchronous video interviewing practices in the hospitality industry. *International Journal of Contemporary Hospitality Management*, *30*(2), 685–701. https://doi.org/10.1108/IJCHM-07-2016-0402

Mîndrilă, D. (2010). Maximum likelihood (ML) and diagonally weighted least squares (DWLS) estimation procedures: A comparison of estimation bias with ordinal and multivariate non-normal data. *International Journal of Digital Society*, *1*(1), 60–66.

Molton, S. (2021, March 10). *Thoughts on asynchronous interviews | LinkedIn*. LinkedIn. https://www.linkedin.com/pulse/thoughts-asynchronous-interviews-stanley-molton/

Muralidhar, S., Kleinlogel, E. P., Mayor, E., Bangerter, A., Mast, M. S., & Gatica-Perez, D. (2020). Understanding applicants' reactions to asynchronous video interviews through self-reports and nonverbal cues. *Proceedings of the 2020 International Conference on Multimodal Interaction*, 566–574. https://doi.org/10.1145/3382507.3418869

Murphy, S. A. (2020). *Let me tell you about a time: exploring how storytelling and memory affect behavioural storytelling and memory affect behavioural interview performance ratings interview performance ratings* [University of Waterloo]. https://scholars.wlu.ca/etd/2294

Nye, C. D., & Drasgow, F. (2011). Assessing goodness of fit: Simple rules of thumb simply do not work. *Organizational Research Methods*, *14*(3), 548–570. https://doi.org/10.1177/1094428110368562

Reynolds, A. A. (2022, December 22). *2022 review shows hiring was strong. Here's our predictions for 2023*. HireVue. https://www.hirevue.com/blog/hiring/where-weve-been-and-where-were-heading-2022-review-2023-predictions

Robie, C., Christiansen, N. D., Bourdage, J. S., Powell, D. M., & Roulin, N. (2020). Nonlinearity in the relationship between impression management tactics and interview performance. *International Journal of Selection and Assessment*, *28*(4), 522–530. https://doi.org/10.1111/IJSA.12307

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*, 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. *Https://Doi.Org/10.1080/1359432X.2022.2156862*, *32*(3), 333–345. https://doi.org/10.1080/1359432X.2022.2156862

Schmidt, W. V., & Conaway, R. N. (1999). Results-oriented interviewing: principles, practices, and procedures. *Technical Communication*, *46*(3), 393–393.

Schoemann, A. M., Boulton, A. J., & Short, S. D. (2017). Determining power and sample size for simple and complex mediation models. *Social Psychological and Personality Science*, *8*(4), 379–386. https://doi.org/10.1177/1948550617715068

Tilston, O., Krings, F., Roulin, N., Bourdage, J. S., & Fetzer, M. (2024). Reactions to Asynchronous Video Interviews: The Role of Design Decisions and Applicant Age and Gender. *Human Resource Management*.

Tross. (1999). *The effect of interviewee coaching on the structured experience-based interview process and outcomes*. Georgia Institute of Technology.

Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2020). Personnel selection in the digital age: a review of validity and applicant reactions, and future research challenges. *European Journal of Work and Organizational Psychology*, *29*(1), 64–77. https://doi.org/10.1080/1359432X.2019.1681401

Zibarras, L. D., Patterson, F., Holmes, J., Flaxman, C., & Kubacki, A. (2018). An exploration of applicant perceptions of asynchronous video MMIs in medical selection. *MedEdPublish*, *7*(4), 1–16. https://doi.org/10.15694/mep.2018.0000285.1

Table 1. *Study Interview Questions and Probes*

Interview Question	Probes	
	Aspect of Story	
	Task/Action	Result
Can you describe a time when you were asked to provide ideas for solving work issues in which you didn't have much experience?	Can you please tell us more about what you considered in your decisions?	Can you please tell us more about the results of your actions?
Describe a situation where you delegated effectively.	Can you please tell us more about why you chose your course of action(s)?	Can you please tell us more about the outcome of your actions?
Describe a time when you successfully persuaded someone (or several others) to see or do things your way.	Can you please tell us more about your role in the situation, including the tasks/actions you performed?	Can you please tell us more about the impact of your actions?

Note. Two probes are provided for each story aspect. Across questions and within a given story aspect, the probes are conceptually the same but reworded to mitigate participant learning effects on outcomes of interest.

Table 2. *Descriptive Information and Correlations for all Variables in Study 1*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. Availability of Probing Condition	0.49	0.50	-								
2. Opportunity to Perform	3.51	0.94	.20**	ω = .93							
3. Honest Self Promotion	3.61	0.84	.04	.36**	ω = .67						
4. Honest Defensive IM	2.10	1.03	.12*	.20**	.15**	ω = .71					
5. Interview Ratings	2.83	0.69	.29**	.11*	.17**	-.07	-				
6. Interview Motivation	4.61	0.59	-.01	.20**	.23**	.05	.06	ω = .91			
7. Traditional Interview Experience	13.06	17.30	-.05	-.00	.03	.03	.01	-.06	-		
8. AVI Experience	1.13	2.20	.06	.09	-.02	.13*	-.08	-.01	.30**	-	
9. Number of Probes Used[a]	5.77	0.80	NA[b]	.12	-.04	.10	.19**	.03	.06	-.09	-
10. Response Length (Sum)	403.83	169.84	.75**	.23**	.09	.12*	.42**	.09	-.02	.05	.77**

Note. Scale reliabilities are presented along the diagonal, in italics. For Availability of Probing Condition, 0 = *AVI without Probing* and 1 = *AVI with Probing*. *M* and *SD* are used to represent mean and standard deviation, respectively. IM = Impression Management. AVI = Asynchronous Video Interview. [a] Correlations are calculated with a subset of the sample filtered for participants in the AVI condition with probes. [b] Due to the subsetted sample with probes, a correlation could not be computed with the Availability of Probing Condition variable.

* indicates $p < .05$. ** indicates $p < .01$.

Table 3. *Summary of SEM Measurement Model Results*

Estimate	Alternative Model 1: One-Factor	Alternative Model 2: Null Model	Hypothesized Model
Fit Indices			
χ^2(df)	1083.34(152)***	1396.55(153)***	272.73(138)***
CFI/TLI	0.73/0.70	0.64/0.60	0.96/0.95
RMSEA	0.12	0.14	0.05
SRMR	0.14	0.13	0.06
χ^2(df) difference relative to the hypothesized model	811(14)***	1124(15)***	
Sample size (*n*)	404 observations, 41 missing patterns		
Direct and Indirect Effects			
H1: Condition → Opportunity to Perform Perceptions		0.21(0.10)***	0.21(0.10)***
H2: Condition → Honest Self-promotion		-0.01(0.13)	-0.01(0.13)
H2: Condition → Honest Defensive IM		0.05(0.09)	0.05(0.09)
H3: Condition → Opportunity to Perform Perceptions → Honest Self-promotion		0.12(0.06)**	0.12(0.06)**
H3: Condition → Opportunity to Perform Perceptions → Honest Defensive IM		0.06(0.03)*	0.06(0.03)*
H4: Condition → Interview Ratings		0.38(0.08)***	0.38(0.08)***
H5: Condition → Honest Self-promotion → Interview Ratings		-0.00(0.03)	-0.00(0.03)
H5: Condition → Honest Defensive IM → Interview Ratings		-0.01(0.02)	-0.01(0.02)
H6: Condition → Opportunity to Perform Perceptions → Interview Motivation		0.06(0.03)*	0.06(0.03)*
H7: Opportunity to Perform → Interview Motivation → Interview Ratings		0.01(0.02)	0.01(0.02)

Note. Values in the table represent standardized structural coefficients for each respective model test. Standard errors are in parentheses. The model difference tests provide evidence of which model best fits the data. All estimates are based on 1,000 boot-strapped samples. ***$p < .001$, ** $p < .01$, * $p < .05$.

Table 4. *Descriptive Information and Correlations for all Variables in Study 2*

Variable	*M*	*SD*	1	2	3	4	5	6	7	8	9
1. AVI with Probing	0.34	0.48	-								
2. AVI Matching Probe Response time	0.33	0.47	-.51**	-							
3. Opportunity to Perform	3.37	1.03	.10	-.10	ω = .94						
4. Honest Self Promotion	3.73	0.96	.01	-.01	.35**	ω = .80					
5. Honest Defensive IM	2.49	1.29	.15*	-.07	.22**	.39**	ω = .87				
6. Interview Performance	2.47	0.86	.20**	-.07	.15*	.14*	-.00	-			
7. Traditional Interview Experience	13.11	17.44	-.01	-.04	.04	.15*	-.04	.14*	-		
8. AVI Experience	2.09	4.31	.07	-.06	.13*	.15*	.18**	-.03	.31**	-	
9. Number of Probes Used[a]	5.78	0.69	NA[b]	NA[b]	.01	.01	-.05	.16	.10	.07	-
10. Response Length (Sum)	358.33	189.55	.60**	-.20**	.08	.08	.02	.49**	.13*	.05	.61**

Note. Scale reliabilities are presented along the diagonal, in italics. For AVI with Probing, 0 = *AVI without Probing* and 1 = *AVI with Probing*. For AVI Matching Probe Response Length, 0 = *AVI without Probing* and 1 = *AVI Matching Probe Response Length*. *M* and *SD* are used to represent mean and standard deviation, respectively. IM = Impression Management. AVI = Asynchronous Video Interview. [a] Correlations are calculated with a subset of the sample filtered for participants in the AVI condition with probes. [b] Due to the subsetted sample with probes, a correlation could not be computed with the AVI with Probing and AVI Matching Probe Response Length condition variables. * indicates $p < .05$. ** indicates $p < .01$.