A quasi-experimental investigation of differences between face-to-face and videoconference interviews in an actual selection process

Langer, Markus; Demetriou, Andrew; Arvanitidis, Alexandros; Vanderveken, Stephane; Hiemstra, Annemarie M.F.

DOI
10.1111/apps.12558

Publication date
2025

Document Version
Final published version

Published in
Applied Psychology

Citation (APA)
Langer, M., Demetriou, A., Arvanitidis, A., Vanderveken, S., & Hiemstra, A. M. F. (2025). A quasi-experimental investigation of differences between face-to-face and videoconference interviews in an actual selection process. *Applied Psychology*, *74*(1), Article e12558. https://doi.org/10.1111/apps.12558

Received: 13 September 2023 | Accepted: 10 June 2024

DOI: 10.1111/apps.12558

ORIGINAL ARTICLE

APPLIED PSYCHOLOGY IAAP

A quasi-experimental investigation of differences between face-to-face and videoconference interviews in an actual selection process

Markus Langer[1] | **Andrew Demetriou**[2] | **Alexandros Arvanitidis**[3] | **Stephane Vanderveken**[3] | **Annemarie M. F. Hiemstra**[4]

[1]Department of Psychology, University of Freiburg, Freiburg im Breisgau, Germany

[2]Delft University of Technology, Delft, the Netherlands

[3]European Personnel Selection Office, Brussels, Belgium

[4]Erasmus University Rotterdam, Rotterdam, the Netherlands

Correspondence
Markus Langer, Department of Psychology, University of Freiburg, Arbeits- und Organisationspsychologie, Engelbergerstraße 41, 79085 Freiburg im Breisgau, Germany.
Email: markus.langer@uni-freiburg.de; markus.langer@psychologie.uni-freiburg.de

Funding information
There was no funding for this study.

Abstract

Videoconference interviews are now integral to many selection processes. Theoretical arguments and empirical findings suggest that videoconference interviews may lead to different interview performance ratings in comparison to Face-to-Face (FTF) interviews. This has led to the question of the comparability of the psychometric properties of videoconferences and FTF interviews. However, evidence from actual selection processes stems from the beginning of the century, and recent findings predominantly stem from simulated interview contexts. We present insights from an actual selection process within a large European organization where we had the unique opportunity for a quasi-experimental investigation of differences between videoconference and FTF interviews. Initially, the organization conducted FTF interviews, and after the onset of the COVID-19 pandemic, the interviews were conducted via videoconference. We examine mean differences in applicant performance ratings and evidence for response format-related validity differences. There

were only small, non-significant mean differences and no evidence for response format related validity differences. We discuss possible causes for discrepancies in our findings compared to previous research. Furthermore, we conclude that downstream consequences of differences between FTF and videoconference interviews may be lower than previously expected. We end with a call for research on the interaction between technology-design and selection-tool-design features.

KEYWORDS

job interviews, technology-mediated selection, validity, videoconference

INTRODUCTION

Videoconference interviews are now integral to many selection processes (Basch et al., 2021; Melchers et al., 2021). Whether they are used as screening tools or in the final selection stages, they have become viable alternatives for Face-to-Face (FTF) interviews. This is particularly true in situations where applicants have no option to appear on site of the organization due to living abroad, being limited in their mobility, or in the case of an ongoing pandemic.

Besides empirically testing practical implications such as time-efficiency or cost savings, research comparing videoconference and FTF interview formats (Basch et al., 2021; Blacksmith et al., 2016; Chapman & Rowe, 2001) is based on the fundamental question of the possible differences in psychometric properties between videoconference and FTF interviews (Blacksmith et al., 2016). However, there is little evidence on the psychometric properties of videoconference interviews. On the one hand, this may be due to the expectation that the well-studied psychometric properties of interviews are not altered by conducting an interview via videoconference. On the other hand, there is evidence that casts doubt on the idea that the psychometric properties of interviews are unaffected by the use of videoconference technology. In particular, applicants seem to react more negatively to videoconference interviews, than to FTF interviews (Blacksmith et al., 2016; Melchers et al., 2021). In addition, applicants sometimes receive better (Chapman & Rowe, 2001) and sometimes worse performance ratings in videoconference than in FTF interviews (Basch et al., 2021; Blacksmith et al., 2016). Such differences may also indicate different psychometric properties of these interview formats (Blacksmith et al., 2016).

However, such findings rarely come from a direct comparison of FTF and videoconference interviews in an actual selection process (for an exception, see Chapman & Rowe, 2001). Access to organizations that allow experimental manipulations in their selection processes is rare. Arguably, organizations' legal, procedural, and distributive justice considerations, as well as possible effects on applicant reactions and performance, make it unlikely that organizations will allow such experimental manipulations.

We present data from an actual selection process in a large European organization that provides a unique opportunity to compare FTF and videoconference interviews in a

quasi-experimental design. Specifically, the selection process began prior to the onset of the COVID-19 pandemic in Europe, with applicants traveling to the site of the organization. Travel restrictions in response to the pandemic made it impossible to continue with the FTF interviews. As a result, the selection process was modified, and the remaining applicants participated in videoconference interviews. We examine mean differences in applicant performance ratings and possible evidence of response format related validity differences by investigating correlations between three interviews that were part of the selection process.

BACKGROUND AND HYPOTHESES DEVELOPMENT

Basic differences between FTF and videoconference interviews

Potosky's framework of media attributes (Potosky, 2008) can be used to understand basic differences between FTF and videoconference interviews (Langer et al., 2017) along four dimensions: social bandwidth, interactivity, transparency, and surveillance. The expectation that these basic differences exist and that they are strong enough to affect interviewees and interviewers alike underlies many of the theoretical arguments for why FTF and videoconference interviews may have different psychometric properties.

Social bandwidth relates closely to the assumptions of media richness theory (Daft & Lengel, 1986) and describes the extent to which relevant communication information (e.g., verbal, nonverbal) is exchangeable: A communication medium that enables a high social bandwidth provides communicators with many opportunities to exchange such information. In contrast to FTF interviews, videoconference interviews may provide less social bandwidth. For instance, visual and audio information may be available in comparatively lower resolution. Furthermore, there may be fewer opportunities for interviewers and interviewees to direct nonverbal behavior at each other, and technology-mediation may undermine natural backchanneling behavior (e.g., nodding; Frauendorfer et al., 2014).

The interactivity of a medium describes the extent to which it is possible to interact during a conversation (Potosky, 2008). Videoconference interviews impede interactivity between interviewer and interviewee because technology mediation can lead to latency in communication (Basch et al., 2021). Furthermore, interactivity can sometimes be completely interrupted by technical problems (e.g., internet connection issues; Fiechter et al., 2018; White & Behrend, 2021).

High transparency means that there are no obstacles during communication and that the interaction partners are not aware that they are using a medium to communicate (Potosky, 2008). Transparency may be lower in videoconference interviews because interviewees are aware that they are communicating with the interviewer through a microphone and camera over the internet (Langer et al., 2017). Similar to interactivity, transparency may be particularly low when there are long latencies and technical problems.

Surveillance means that it might be possible for a third party to interrupt or monitor the conversation (Langer et al., 2017; Potosky, 2008). Closely associated with the aspect of surveillance are applicants' information privacy concerns, which refer to applicants' concerns about their privacy in general and personal data in particular (Basch et al., 2021). Overall, it seems more likely for videoconferences than for FTF interviews that people will have concerns about surveillance, given that sensitive data is transmitted over the internet.

Differences in applicant performance ratings between FTF and videoconference interviews

Based on the differences between FTF and videoconference interviews, research has proposed several reasons for possible mean differences in applicant performance ratings between these interview formats. There are two main lines of reasoning, one arguing for possible negative effects of videoconference interviews on applicant performance ratings (see, e.g., Basch et al., 2021), another arguing for possible positive effects (see, e.g., Chapman & Rowe, 2001). Both see possible causes on the interviewee's and the interviewer's side.

Videoconference interviews will negatively affect applicant performance ratings: On the interviewee's side, there may be less opportunity for impression management (Basch & Melchers, 2020; Blacksmith et al., 2016). Whereas the aforementioned differences (e.g., regarding social bandwidth and interactivity) between FTF and videoconference may have limited effects on verbal impression management, these differences might mitigate nonverbal impression management (Blacksmith et al., 2016). In videoconference interviews, it is more difficult to read the interviewer's facial expressions and other nonverbal behavior, and consequently, it is less possible to react to the interviewer's behavior (Basch et al., 2021; Blacksmith et al., 2016). Technology-mediated communication may also hinder eye contact and smiling (Basch et al., 2021), which can negatively affect rapport building between interviewer and interviewee (Barrick et al., 2009; Blacksmith et al., 2016). Similar reasons also feed into the argument that interviewees may experience greater anxiety in videoconference interviews (Blacksmith et al., 2016), which can lower their performance (Powell et al., 2018; see Melchers et al., 2021, who found no significant difference between FTF and videoconference interviews for applicant anxiety). Furthermore, there is initial evidence that videoconference interviews can elicit stronger privacy concerns in applicants (Basch et al., 2021). Consistent with research showing that privacy concerns can reduce test-taking motivation in personnel selection processes (Bauer et al., 2006), it is possible that videoconference interviews may worsen applicant performance.

On the interviewer's side, videoconferences could make it more difficult to follow what applicants are saying, particularly if there are technical problems (Sears et al., 2013). Furthermore, there is less opportunity for rapport building, less social presence, and interactions with interviewees may feel more taxing (e.g., due to less conversational fluency), which can undermine applicant's likability and thus lower performance ratings (Basch et al., 2021; Sears et al., 2013). Other arguments are that technology-mediation can affect applicant performance ratings because nonverbal signals of intelligence may be more difficult to interpret (Blacksmith et al., 2016; DeGroot & Gooty, 2009). Finally, technical problems may be attributed to the applicant, thus reducing their performance ratings (although the evidence is mixed, with some studies showing that technical problems may be attributed to the applicant, Fiechter et al., 2018, and others showing that technical problems may be attributed to the situation; White & Behrend, 2021).

Videoconference interviews will positively affect applicant performance ratings: On the interviewee's side, the comparatively lower perceived social presence in videoconference interviews could reduce applicants' interview anxiety (Chapman & Rowe, 2001; Short et al., 1976). The argument is that the lower social presence in videoconference interviews may reduce the perceived evaluative nature and thus the perceived pressure of the interview situation. This could attenuate applicants' interview anxiety, which could then improve their interview performance (Powell et al., 2018).

On the interviewer's side, Chapman and Rowe (2001) argue that a combination of attribution processes, naïve theories, and motivated cognition may contribute to better performance ratings in videoconference interviews. Because videoconference interviews offer less social bandwidth, interactivity, and transparency, interviewers are likely to be aware that their conversation is technology-mediated (Chapman & Webster, 2001). This may make the specific situation in which the interview is conducted more salient than in the FTF format. Consistent with attribution theory (Kelley & Michela, 1980), this strong salience of the situation makes it more likely that interviewers attribute possible low performance of applicants to the situation rather than to dispositional characteristics of the applicant (e.g., their suitability for the job; Chapman & Rowe, 2001; Chapman & Webster, 2001). Such attributions might include that the person has technical problems or is unaccustomed to interviews and evaluative situations via videoconference tools. This line of reasoning is supported by Chapman and Webster (2001), who showed that performance attributions are related to perceptions of media richness in that interview performance is more internally attributed in situations of high perceived media richness.

Such attribution processes may be further fueled by interviewers' naïve theories (Wegener & Petty, 1995). Interviewers may have the naïve theory that the videoconference situation may affect their interview ratings or that people in videoconference interviews are at a disadvantage compared to people in FTF interviews. Given that interviewers are often trained to try to evaluate applicants as accurately as possible, they also have an accuracy goal that may motivate them to try to account for effects that may bias their evaluation of applicants (Chapman & Rowe, 2001). This could lead interviewers to (over)compensate for expected negative effects of the videoconference situation by giving more lenient performance ratings. This reasoning is particularly important in our context, where interviewers experienced a switch from the FTF format of the interview to the videoconference format, which could make the videoconference situation more salient, and attributions to the videoconference situation as well as naïve theories about the influence of this situation more likely.

Few studies have empirically examined mean differences in applicant performance ratings between videoconference and FTF interviews. The earliest study that we are aware of reported that applicants received moderately better interview ratings in videoconference interviews $d = 0.45$ (Chapman & Rowe, 2001). The meta-analysis by Blacksmith et al. (2016) reported a moderate negative effect of $d = -0.46$ for applicant performance ratings in videoconference interviews compared to FTF interviews. However, this estimate was based on only three studies with a total of 109 participants, only one field study (the one by Chapman & Rowe, 2001), and its 80% credibility interval included zero. One study that was not included in the meta-analysis was Sears et al. (2013), who reported that interviewees in videoconference condition received moderately lower interview scores ($d = -0.52$) compared to the FTF condition. More recently, Basch et al. (2021) also found a moderate negative effect on performance ratings for videoconference compared to FTF interviews ($d = -0.43$), and Melchers et al. (2021) reported an even stronger effect of $d = -0.62$. These three studies used simulated job interview settings. Further research based on 50 applicants and their interviews in actual selection settings found differences in personality ratings between FTF and videoconference interviews (Michelotti et al., 2021). Conscientiousness, agreeableness, extraversion, and openness to experience were perceived as higher, and neuroticism was perceived as lower in FTF interviews. This could indicate differences between FTF and videoconference interviews that may contribute to higher applicant performance ratings in the former.

Overall, there are arguments for both positive and negative effects of videoconference interviewing on applicant performance ratings. Empirical evidence seems to suggest that applicants in videoconference interviews are more likely to receive lower performance ratings, but there is a lack of evidence that stems from field settings. With this in mind, we decided to ask.

Research Question 1: Are there differences in applicant performance ratings between the videoconference and the FTF interview format?

Response format related validity differences

Blacksmith et al. (2016) concluded that the mean differences in applicant performance ratings between FTF and videoconference interviews may also indicate that there will be response format-related implications for the interviews' validity. For instance, the reduced social bandwidth may mitigate the potentially biasing effects of visual cues (e.g., applicants' physical attractiveness; Sears et al., 2013; Tsai et al., 2012). At the same time, nonverbal and verbal cues that may contain valid information about applicant personality (DeGroot & Kluemper, 2007) may be less available (Blacksmith et al., 2016).

Lower social bandwidth, combined with lower interactivity, could also affect the validity of videoconference interviews as well as what videoconference interviews measure. For instance, less rapport building may reduce sympathy effects on interview ratings (Sears et al., 2013). It could also be more difficult to judge applicants' social skills, which could be particularly important for certain jobs (e.g., salespeople; Blacksmith et al., 2016). In addition, less transparency may lead to more strenuous interactions (White & Behrend, 2021). Again, this may directly reduce the effects of sympathy on interview ratings, and/or it could make it more difficult to validly judge applicants' characteristics.

The aforementioned arguments for possible positive or possible negative effects on applicant performance ratings (on the side of the interviewee and on the side of the interviewer) could also contribute to possible response format related validity differences (Lievens & Sackett, 2017). For instance, if videoconference interviews influence interview anxiety, compared to FTF interviews, this could lead to applicants showing less (or more) behavior relevant to judge their actual characteristics and actual suitability. If applicant likability plays less of a role in interviewers' evaluations in videoconference interview settings, this could reduce unwanted variance in evaluations of applicants' actual suitability. In addition, if attribution processes play a role, the more lenient evaluation of applicants could increase this unwanted variance in applicant performance ratings.

To date, there is limited evidence on the validity of videoconference interviews and even less evidence on response format related validity differences. However, if differences in the response formats affect interviewee behavior and/or interviewer performance ratings, this could affect what the interviews measure. Regardless of the exact nature of these effects on interviewees and interviewers, we could consider these effects to be *method effects* of the interview response format (Lievens & Sackett, 2017). If there are such method effects that influence what the interview measures (i.e., its validity), we might expect that conducting two different interviews in the same format (i.e., both via videoconference) might lead to stronger correlations between those interviews than conducting those two different interviews in different formats (i.e., one in a videoconference format the other in a FTF format).

With the current data, we had the opportunity to investigate correlations between different interviews that were part of the same selection process. There was a *situational interview*, which

was always conducted in a videoconference format. There was a *behavioral interview* and a *field-related interview*, both of which were conducted in person before the onset of the pandemic and in a videoconference format after the onset of the pandemic. If there are effects of the response format that affect what the interview measures, we expect to find higher correlations between the videoconference interview that all applicants had to complete (situational interview) and the videoconference version of the behavioral interview as well as the field-related interview.

> **Hypothesis 1.** There will be a higher correlation between the situational interview and the videoconference format of the behavioral interview and field-related interview than for the FTF format of those interviews.

METHOD

Sample

Analyses were conducted with R 4.2.2 (R Core Team, 2022).[1] We did not preregister the hypotheses or analysis plan, and we did not determine the required sample size in advance because we report analyses on data that was already collected at the time the team of researchers was contacted. Before receiving the data, the authors signed a memorandum of understanding with the organization and applied for ethical approval with the IRB of the main author's institution at the time of analysis.

The sample consisted of $N = 385$ applicants (51% female), of which $n = 289$ conducted the FTF and $n = 96$ the videoconference versions of the behavioral interview and the field-related interview. For data privacy purposes, the organization only provided information about applicant age in three groups: younger (up to 35; $n = 122$), middle (36–40; $n = 122$), and older age (above 41; $n = 141$). Applicants were recruited from all over Europe. Applicants were from Italy (17.7%), France (11.9%), Greece (11.2%), Spain (10.4%), Germany (6.8%), Belgium (6.2%), Romania (4.9%), Poland (4.7%), Hungary (3.1%), Bulgaria (2.9%), Portugal (2.9%), and other European countries each with less than 10 applicants per country. Applicants were allowed to choose their preferred language for the selection tools between English, German, French, and Italian (one of those languages for all selection tools). Most applicants chose English (89.4%).

We decided to include data from all applicants applying for an "Administrator" position within an overall hiring process in the field of "Economy". Specifically, applicants applied to an "Administrator" position in the field of "Economy" and then had to choose one out of five different specializations that they thought is the best fit for their expertise. Of the applicants, 40.7% applied for the specialization Competition Law, 28.0% for Financial Rules applicable to the EU, 23.6% for Financial Law, 4.4% for EU budget, and 3.3% for Economic and Monetary Union Law. We decided to include all this data because a) the specializations were very similar and all belonged to the overall field of economy, b) because the behavioral interview and the situational interviews were exactly the same for all positions, and for the field-related interview the structure of the interview was the same with questions being slightly different for the different specializations but still within the field of economy, and c) because the set of interviewers was the same across specializations, and they were instructed and trained in the same way. This means that not only was the experience from the side of the applicants similar across the different specializations, but also on the side of the interviewers, we had similar conditions across the specializations.

Procedure

The application process started in June 2019. The first selection stage included an online process where applicants had to respond to questions regarding their CV that focused on the required experience for the respective position. Applicants who passed this stage constitute the sample included in this paper. Those applicants participated in a case study, a situational interview, a behavioral interview, and a field-related interview. The case study was conducted before the onset of the pandemic and was always conducted on the site of the organization at different test centers across Europe. We report correlations with this selection tool, as this may be of interest to readers. The situational interview was introduced after the onset of the pandemic and was always conducted online. The behavioral interview and field-related interview were first conducted as FTF (from January 2020 to March 2020), and after March 2020, the onset of the COVID-19 pandemic, they were conducted via videoconference. Comparisons between the FTF and videoconference formats were thus possible for the behavioral and the field-related interview.

Measures

A set of raters evaluated applicant performance for all selection tools. The raters were experts in the respective field for which the organization was hiring. For every applicant, there were different sets of two raters who conducted the interview and then discussed applicant responses and agreed on applicant ratings. This made it impossible to calculate inter-rater reliabilities. All raters received training for their task. This involved face-to-face training sessions before the pandemic and online sessions after the onset of the pandemic. For the current selection processes specifically, the training included two and a half days of training on general selection methodology, followed by specific training for each of the different interviews. All raters received an additional half-day of training for each of the interviews, where they practiced interviews with mock applicants.[2] For the situational and the behavioral interview, the interview questions and rating scales were provided by the organization. For the field-related interview, the raters received a training to learn the mechanism for developing the field-related questions, and rating scales, and then developed those materials for the field-related interview. Given that the field-related interview was fully structured, the interviewers developed those materials once and then used them consistently throughout the selection process.

For every selection tool, applicants were rated on different competencies on a scale from 1 (low performance) to 10 (high performance).[3] For the current analyses, we averaged over the ratings for the competencies for the individual selection tools to calculate a single score for each selection tool (i.e., one for the case study, situational interview, behavioral interview, and the field-related interview).

The case study included a computer-based test based on a scenario where applicants were faced with various problems that they were asked to solve or to which they had to react (e.g., prepare a presentation, draft email responses) based on the provided materials. The case study was rated on the competencies "analysis and problem solving", "communication", "delivering quality and results", and "prioritizing and organizing".

The situational interview was semi-structured (given that the interviewers were allowed to ask follow-up questions) and aimed at assessing general competencies by asking the applicant

how they would react in a hypothetical situation. Prior to the interview, the applicants were invited to read an online written assignment and background information to prepare for the situational interview. The situation described in the assignment, as well as additional situations, formed the basis for questions during the interview. During the interview, the applicants were allowed to consult the written assignment and the background information. Applicants received a single main question per competency (i.e. six main questions) and possible follow-up questions. For the semi-structured interview forms (situational and behavioral), the interviewers received training and instructions on the follow-up questioning approach to ensure that the follow-up questioning was as standardized as possible. The questions started with the description of the situational context, followed by the specific question. As an example, there was a question that started with a situational description that a colleague was appointed to participate in a high-level meeting but, due to an unexpected cause, is not available. The question for the applicants was then: "Your task is to replace your colleague and take over their duties, how would you best prepare yourself to replace them in this situation?" The situational interview was used to rate the competencies "analysis and problem solving", "learning and development", "prioritizing and organizing", "resilience", "working with others", and "leadership".

The behavioral interview was semi-structured and aimed at assessing general competencies by asking the applicant how they had handled a specific situation in the past. Applicants received a single main question per competency (i.e. five main questions[4]) and possible follow-up questions. Applicants were asked to provide the interviewers with concrete examples of situations from the past where they had shown these competencies. A sample question was, "Can you give us an example of a situation where it was hard to reach a consensus with your colleagues within a working group you participated?" The behavioral interview was used to rate the competencies "communication", "delivering quality and results", "learning and development", "resilience", "working with others", and "leadership".

The field-related interview followed a fully structured interview approach and was designed to evaluate specific skills and knowledge required for the position. In contrast to the behavioral interview and situational interview, which both focused on general competencies, the field-related interview focused on specifics for the position (e.g., specific technical expertise). Thus, the field-related interview questions differed slightly among the different specializations that applicants applied for, but all questions were related to the field of "Economy". Applicants received a score between 1 and 100 for the field-related interview (we recoded this rating to reflects a range between 1 and 10). A sample question was, "From your experience in procurement activities, could you give us an example of a recent project where you were involved in the tendering process?"

Available applicant demographics were gender, age group, nationality, their first language, and the language they chose to for the interview.

RESULTS

Table 1 shows correlations, means, and standard deviations for all variables. Tables 2 and 3 provide correlations, means, and standard deviations for all variables for the subsample of applicants who participated in the FTF format of the behavioral interview and field-related interview (Table 2) and the subsample of applicants who participated in the videoconference format (Table 3).

TABLE 1 Means, standard deviations, and correlations for the overall sample.

Variable	*M*	*SD*	1	2	3	4	5
1. Case study performance	5.11	1.01					
2. Situational interview performance	6.26	0.88	.25** [.15, .34]				
3. Behavioral interview performance	6.36	0.67	.19** [.09, .28]	.26** [.16, .35]			
4. Field-related interview performance	6.69	1.08	.12* [.02, .22]	.29** [.20, .38]	.22** [.12, .31]		
5. Gender	0.49	0.50	.04 [−.06, .14]	−.01 [−.11, .09]	.02 [−.08, .12]	.06 [−.04, .16]	
6. Face-to-face vs. videoconference	0.25	0.43	−.08 [−.17, .02]	.00 [−.10, .10]	−.04 [−.14, .06]	−.02 [−.12, .08]	.11* [.01, .21]

Note: Gender is coded as 0 = female, 1 = male. Face-to-face vs. videoconference is coded as 0 = face-to-face interview, 1 = videoconference interview.
$N = 385$.
*$p < .05$, and **$p < .01$.

TABLE 2 Means, standard deviations, and correlations for the subsample of applicants who took part in the face-to-face format of the behavioral interview and the field-related interview.

Variable	*M*	*SD*	1	2	3	4
1. Case study performance	5.15	1.01				
2. Situational interview performance	6.26	0.90	.25** [.14, .35]			
3. Behavioral interview performance	6.38	0.66	.14* [.03, .25]	.23** [.12, .34]		
4. Field-related interview performance	6.71	1.10	.10 [−.01, .21]	.30** [.19, .40]	.19** [.08, .30]	
5. Gender	0.45	0.50	.01 [−.11, .12]	−.01 [−.12, .11]	.02 [−.09, .14]	.03 [−.08, .15]

Note: Gender is coded as 0 = female, 1 = male.
$N = 289$.
*$p < .05$, and **$p < .01$.

Research question and hypothesis

Research Question 1 asked whether there are differences in applicant performance ratings between the videoconference and the FTF interview format? Given the unequal sample sizes, we conducted Levene's tests to check for possible significant differences in the variances of the interview performance ratings of the two samples. For both the behavioral interview, $F(1, 383) = 2.33$, $p = 0.13$, and the field-related interview, $F(1, 383) = 0.51$, $p = 0.48$, the variances of the samples were comparable. We thus decided to calculate independent sample *t*-tests with pooled

TABLE 3 Means, standard deviations, and correlations for the subsample of applicants who took part in the videoconference format of the CBI and FRI.

Variable	*M*	*SD*	1	2	3	4
1. Case study performance	4.98	0.99				
2. Situational interview performance	6.27	0.81	.25* [.05, .43]			
3. Behavioral interview performance	6.31	0.73	.31** [.12, .48]	.34** [.16, .51]		
4. Field-related interview performance	6.65	1.02	.19 [−.01, .37]	.28** [.09, .46]	.29** [.09, .46]	
5. Gender	0.58	0.50	.16 [−.04, .35]	−.01 [−.21, .19]	.02 [−.18, .22]	.17 [−.03, .36]

Note: Gender is coded as 0 = female, 1 = male.
$N = 96$.
*$p < .05$, and **$p < .01$.

variances. We also tested the robustness of our findings by additionally calculating several Welch's two-sample *t*-tests, where the results were nearly identical to the results we report in the manuscript. Although the mean values for both the behavioral interview and the field-related interview were numerically higher for the FTF interviews, there was no statistically significant difference between the conditions and small effect sizes (behavioral interview: $t(383) = 0.78$, $p = 0.22$, $d = 0.09$ [−0.14, 0.32]; field-related interview: $t(383) = 0.49$, $p = 0.31$, $d = 0.06$ [−0.17, 0.29]). The effect sizes differed substantially from prior findings, where applicants in the videoconference format received either moderately worse ($\sim d = -0.40$ – -0.60) or moderately better ratings ($d = 0.46$). Therefore, the response to Research Question 1 is that there were no significant differences in applicant performance ratings.

Hypothesis 1 proposes that there will be a higher correlation between the situational interview and the videoconference format of the behavioral interview and field-related interview than for the FTF format of those interviews. As Tables 2 and 3 show, for both FTF and videoconference interviews, the correlations with ratings on the other selection tools were small to moderate, positive, and all significant. Although the numeric values of the correlations between the behavioral and the situational interview (.34) were higher for the videoconference format than for the FTF format (.23), this difference was not statistically significant, $z = 1.01$, $p = 0.16$. The numeric values for the correlations between the field-related interview and the situational interview were slightly lower (.28) for the videoconference format than for the FTF format (.30), which was not in line with the hypothesis and also not statistically significant, $z = -0.18$, $p = 0.57$. All confidence intervals for the relevant correlations overlapped substantially. Thus, there was no support for Hypothesis 1.

Additional analyses

Given the quasi-experimental nature of our research, we examined the comparability of the videoconference and the FTF subsample. Table 4 shows the demographic characteristics as well

TABLE 4 Distribution of demographics in the videoconference and in the face-to-face sample.

	Videoconference	Face-to-face
Gender	Female = 41.7%	Female = 54.7%
	Male = 58.3%	Male = 45.3%
Age	Young = 32.3%	Young = 31.5%
	Middle = 25.0%	Middle = 33.9%
	Old = 42.7%	Old = 34.6%
Interview language (chosen by the applicant)	English = 79.2%	English = 92.7%
	Other = 20.8%	Other = 7.3%
First language	English = 4.8%	English = 17.7%
	Other = 95.2%	Other = 82.3%
Specialization	Competition law = 14.6% (14)	Competition law = 49.5% (143)
	Financial rules = 2.1% (2)	Financial rules = 36.3% (105)
	Financial law = 52.1% (50)	Financial law = 14.2% (41)
	EU budget = 17.7% (17)	EU budget = 0%
	Economic and monetary union law = 13.5% (13)	Economic and monetary union law = 0%

Note: $n_{videoconference} = 96$, $n_{FTF} = 289$.

as the specialization that applicants chose, grouped by the two subsamples. We also calculated linear regressions using the dummy-coded version of the videoconference versus the FTF format and added demographic information as predictors where it made sense – also as an interaction term (e.g., the interaction between the interview format and gender).

For gender, there is a slight imbalance with more female applicants who took the FTF version of the interview. Including gender as a predictor in the regression would not change the findings in such a way that we would have come to different conclusions. It was also not a significant predictor, and there was also no significant interaction between gender and the interview format.

For age, there were relatively balanced subsamples. Including age as a predictor in the regression would not change the findings in such a way that we would have come to different conclusions. Age was a significant predictor in the regression for the behavioral interview: applicants from the middle age group received lower ratings than applicants from the young age group. There was no interaction between age and the interview format. For the field-related interview, age was not a significant predictor.

For the interview language, there was a slight imbalance in that English was chosen as the interview language comparably more often in the FTF format. Including the interview language as a predictor in the regression (English vs. other language) would not change the findings in such a way that we would have come to different conclusions. It was also not a significant predictor, and there was also no significant interaction between interview language and the interview format.

For applicants' first language, there was a slight imbalance in that there were comparably more applicants with English as their first language in the FTF condition. Including first language as a predictor in the regression (English vs. other language) would not change the

findings in such a way that we would have come to different conclusions. Also, it was not a significant predictor, and there was also no significant interaction between the first language and the interview format.

For specialization, there were imbalances in the sense that for EU Budget and Economic and Monetary Union Law, there were no FTF interviews, whereas for Financial Rules, there were only a small number of videoconference interviews that had been conducted. This is a result of the quasi-experimental nature of the research and of our choice to include samples from different specializations (for our reasoning for why we included applicants from different specializations in the field of Economy in our dataset, see the Sample section in the Methods section).

DISCUSSION

In this paper, we present a quasi-experimental investigation of possible differences between FTF and videoconference interviews based on data from an actual selection process of a large European organization. The main findings of this investigation were that a) there were only small, non-significant mean differences in performance between applicants who participated in the FTF and videoconference interviews, and b) there was no evidence of response format related validity differences.

Theoretical implications

No evidence for mean differences between FTF and videoconference interviews

Contrary to meta-analytic findings (Blacksmith et al., 2016) and in contrast to recent findings (Basch et al., 2021; Melchers et al., 2021), our data provided no evidence of differences in performance ratings between FTF and videoconference interviews. There are several possible reasons why our findings are not consistent with the commonly found moderate differences in candidate performance ratings – either positive or negative – for videoconference interviews compared to FTF interviews.

First, not finding differences between FTF and videoconference interviews may indicate that there are smaller basic differences between these interview formats than expected. We used Potosky's framework of media attributes (Potosky, 2008) to highlight possible differences between these interview formats, as have been highlighted in other research (Basch et al., 2021). If these differences (in terms of social bandwidth, interactivity, transparency, and surveillance) are smaller than expected, this undermines any theoretical reasoning that builds on these basic differences (e.g., with respect to rapport building, nonverbal impression management, attribution processes).

Second, recent studies reporting differences between FTF and videoconference interviews were conducted in simulated contexts. It is possible that there are features of simulated interviews that affect interviewee performance ratings in videoconference interviews. For instance, if participants are less motivated to perform well in simulated contexts, challenging aspects of videoconference interviews (e.g., technical problems, less rapport building, privacy concerns) may more strongly influence interviewee behavior. In contrast, in actual selection contexts, applicants may be more motivated to cope with possible challenges. Regarding Michelotti et al.

(2021) as the only recent study that we are aware of from an actual selection setting, the difference to our study was that their study was based on only 50 applicants and focused entirely on personality ratings. Potentially, the assessment of general personality facets is different from the assessment of overall suitability for a job and job-related competencies.

Third, the field study by Chapman and Rowe (2001) that provided evidence of positive effects of videoconference formats on applicant performance ratings was published in 2001. The main theoretical rationale of their study was a combination of attribution processes, naïve theories, and motivated cognition effects: interviewers may be inclined to attribute negative applicant performance to the videoconference interview situation rather than to dispositional characteristics of the interviewee. Since then, videoconference technologies have become commonplace, even before the pandemic. Consequently, interviewers may not be more lenient in their evaluation of applicants in videoconference interviews because they may believe that interviewees today should be accustomed to videoconference situations. Alternatively, since 2001, the social bandwidth and interactivity of videoconference interviews have improved along with technical advances. This improves what the framework of media attributes (Potosky, 2008) calls the transparency of the interview: interviewers are less likely to know that they are communicating with applicants via a videoconference tool. This may also be related to the fact that interviewers are less likely to be aware that videoconferencing is a situational factor that may affect the interview. As a result, interviewers may be less likely to attribute poor applicant performance to the videoconference situation, which would make it less likely that interviewers assess applicants more favorably.

Fourth, other kinds of attribution processes, however, may have played a stronger role in our context. In particular, it may have been the unique situation of our quasi-experimental study that led to the current findings. When the interviews were changed to the videoconference format, it was not possible to conduct FTF interviews due to the pandemic. Thus, while there may have been effects of the videoconference interview that detrimentally affected applicant performance, everyone involved was aware of the situational constraint that a technology-mediated interview was the only option available. From the interviewers' perspective, the context of the pandemic may have been perceived as an important situational factor that could have contributed to more lenient ratings of applicants. In other words, after the onset of the pandemic, applicants may have received a "pandemic bonus" for their ratings. This could have reduced potential differences between the FTF and videoconference interviews, contradicting the findings in recent years of negative effects of videoconference interviews on applicant ratings (Basch et al., 2021; Melchers et al., 2021). Unfortunately, we did not have information to examine this possible effect.

Finally, from the interviewees' perspective, consistent with the justice literature (Gilliland, 1993; Hausknecht et al., 2004), it is possible to expect lower applicant performance in cases where applicants have negative reactions to the interview process. In cases where there is no other option than videoconferencing, it may alleviate negative reactions and thus also reduce negative effects on performance. Unfortunately, this will need to be explored in future research, as there was no data to test this assumption in the current case.

No evidence for response format related validity differences

We had a unique constellation with high ecological validity that allowed us to evaluate correlations between different interviews conducted via different response formats in an actual high-

stakes selection process. We proposed that if there are effects on either the interviewee or the interviewer that contribute to differences between videoconference and FTF interviews, there may be a common-method effect: correlations between different interviews conducted using the same response format could be stronger than correlations between different interviews conducted using different response formats (Lievens & Sackett, 2017).

The findings were not in line with this proposition, and consequently, there was no evidence of response format related validity differences between the videoconference and the FTF version. Although we found a numerically higher correlation for the videoconference version of the behavioral and the situational interview, this finding was not significant. In addition, the correlation for the videoconference version of the field-related interview and the situational interview was not in line with our hypothesized direction of effects. Consequently, differences between the response formats that may affect applicant behavior (e.g., effects on interview anxiety; effects on impression management) and/or interviewer ratings (e.g., dampening of applicant likability effects, less rapport building, attribution processes with respect to applicant performance) may be weaker than expected based on prior work (Basch et al., 2021; Blacksmith et al., 2016; Chapman & Rowe, 2001).

Main practical implication

Our findings cast doubt on the previously reported effect of videoconferencing on applicant performance ratings in interviews (Basch et al., 2021; Blacksmith et al., 2016; Chapman & Rowe, 2001; Melchers et al., 2021). Whereas previous research has cautioned against using FTF and videoconference response formats as alternatives in the same selection stage (e.g., Basch et al., 2021), our findings suggest that the choice of a videoconference interview may not be as influential on applicant performance ratings as previously assumed. This may have implications for practice, where practitioners wonder about the effects of interviewing some applicants on site and others via videoconference. However, given mixed study results on applicant performance ratings in FTF versus videoconference interviews (Basch et al., 2021; Chapman & Rowe, 2001), and our finding of no significant differences, it may still be advisable to be cautious when using different response formats at the same selection stage.

Limitations and future research

There are two key limitations to the current findings. First, there are clear limitations associated with quasi-experimental designs. For instance, participants were not randomly assigned to the experimental conditions, so there may be individual differences that are differentially distributed across the two conditions. Additionally, because the quasi-experimental manipulation occurred at a specific point in time, time-related confounding effects (as described earlier in the example of the possible "pandemic bonus") cannot be ruled out. Thus, any comparison with prior research, which has often been conducted in randomized experimental studies, is ultimately tentative.

Second, our findings are based on archival data. This means that we did not have the opportunity to collect data that could have helped to rule out alternative explanations for our findings. For example, pre-interview impressions of applicants can impact final applicant performance ratings (Dougherty & Turban, 1994). It would have been intriguing to examine if the difference between pre-interview impressions of applicants and post-interview ratings varies

between videoconference and FTF interviews. Unfortunately, such data is not available. Nor do we have any insights on applicant reactions to the interview that could have also influenced applicants interview performance (Hausknecht et al., 2004). If applicant reactions were positive due to the pandemic situation, this could have contributed to the unexpected findings that there was no difference between the FTF and videoconference format. However, this should only be understood as speculation about how applicant reactions could have affected our findings, as we have no evidence on this issue.

For future research, there is a need to examine the psychometric properties of different variants of videoconference interviews and other technology-mediated selection approaches. Although our findings may be interpreted optimistically regarding the comparability of FTF and videoconference interview formats, our findings may not generalize to other interview formats or other selection tools. With respect to other interview formats, asynchronous video interviews (AVIs) are more different from FTF interviews, particularly regarding the media attributes of interactivity and transparency, and research has reported that AVIs may also not be comparable to synchronous versions of videoconference interviews (e.g., because applicants can receive better evaluations in AVIs; Langer et al., 2017). The specific technology-design features of AVIs may cause stronger effects on the psychometric properties of interviews and may have stronger effects on applicants with certain characteristics.

With respect to other selection tools, in assessment centers there are also parts that could be conducted face-to-face or via videoconferencing tools, and for which there is even less research on the implications of the response format than for interviews. For example, for group discussions, videoconference formats may be particularly challenging, for example, due to their lower social bandwidth. Lower social bandwidth reduces the ability to signal an intent to respond to another applicant's comment, making turn-taking more complex. To address such issues, videoconferencing technology offers features such as virtual hand raising, but how this affects psychometric properties, applicant reactions, or the fairness of selection tools remains a question for future research. There is reason to believe that technology-design features (e.g., asynchronous vs. synchronous response format; virtual hand raising) and selection-tool-design features (e.g., interview structure; number of people in the group discussion) interact to produce a complex set of effects worthy of future investigation.

CONCLUSION

Research comparing FTF and videoconference interviews is based on the fundamental question of the comparability of the psychometric properties of the two interview formats. Research showing differences in applicant performance ratings between these interview formats has raised concerns about the comparability of their psychometric properties. In our data from an actual selection process, there was no evidence of differences in applicant performance ratings between FTF and videoconference interviews. Furthermore, there was also no evidence of response format related validity differences. Although this allows for (tentative) optimistic conclusions regarding the comparability of the psychometric properties of FTF and videoconference interviews, it is necessary to examine the validity of technology-mediated selection approaches more directly, as there are several theoretical propositions (see, e.g., Basch et al., 2021; Chapman & Rowe, 2001; Potosky, 2008) that provide arguments for effects on the validity, on applicant reactions, or on fairness that may result from the interplay of technology-design and selection-tool-design features.

ACKNOWLEDGEMENTS

Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

Alexandros Arvanitidis and Stephane Vanderveken are employees of the organization that granted us access to the archival data.

DATA AVAILABILITY STATEMENT

Data cannot be made available since this is data from an actual selection process.

ETHICS APPROVAL

The use of the archival data was approved by the ethics commission of the first author's former institution (Saarland University) Ethics approval number 21-18, Date 31st of August 2021.

ENDNOTES

[1] The following packages were used for the analyses: cocor, car, sjPlot, psych, apaTables, and tidyverse (Diedenhofen & Musch, 2015; Fox & Weisberg, 2019; Lüdecke, 2022; Revelle, 2022; Stanley, 2021; Wickham et al., 2019)

[2] The organization wanted us to mention that they have since adjusted their selection process and each of the selection tools. In addition, each rater now provides a separate rating for the applicant's responses, which is then averaged across the two raters.

[3] The competencies were defined by the behaviors that raters were instructed to assess in each selection tool. See the following link for a complete list of behaviors associated with the competencies that were assessed in the current selection process https://epso.europa.eu/sites/default/files/documents//general/anchors/en_overview_anchors_for_publication.xlsx

[4] There was no question related to the Communication competency. Interviewers were required to assess this competency based on the applicant's overall interview performance.

REFERENCES

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

Basch, J. M., & Melchers, K. G. (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., 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*, 601–621. https://doi.org/10.1177/0149206306289829

Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the employment interview: A meta-analysis. *Personnel Assessment and Decisions*, *2*, 2. https://doi.org/10.25035/pad.2016.002

Chapman, D. S., & Rowe, P. M. (2001). The impact of videoconference technology, interview structure, and interviewer gender on interviewer evaluations in the employment interview: A field experiment. *Journal of Occupational and Organizational Psychology*, *74*(3), 279–298. https://doi.org/10.1348/096317901167361

Chapman, D. S., & Webster, J. (2001). Rater correction processes in applicant selection using videoconference technology: The role of attributions. *Journal of Applied Social Psychology*, *31*(12), 2518–2537. https://doi.org/10.1111/j.1559-1816.2001.tb00188.x

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science*, *32*(5), 554–571. https://doi.org/10.1287/mnsc.32.5.554

DeGroot, T., & Gooty, J. (2009). Can nonverbal cues be used to make meaningful personality attributions in employment interviews? *Journal of Business and Psychology*, *24*(2), 179–192. https://doi.org/10.1007/s10869-009-9098-0

DeGroot, T., & Kluemper, D. (2007). Evidence of predictive and incremental validity of personality factors, vocal attractiveness and the situational interview. *International Journal of Selection and Assessment*, *15*(1), 30–39. https://doi.org/10.1111/j.1468-2389.2007.00365.x

Diedenhofen, B., & Musch, J. (2015). Cocor: A comprehensive solution for the statistical comparison of correlations. *PLoS ONE*, *10*(4), e0121945. https://doi.org/10.1371/journal.pone.0121945

Dougherty, T. W., & Turban, D. B. (1994). Confirming first impressions in the employment interview: A field study of interviewer behavior. *Journal of Applied Psychology*, *79*(5), 659–665. https://doi.org/10.1037/0021-9010.79.5.659

Fiechter, J. L., Fealing, C., Gerrard, R., & Kornell, N. (2018). Audiovisual quality impacts assessments of job candidates in video interviews: Evidence for an AV quality bias. *Cognitive Research: Principles and Implications*, *3*(1), 47. https://doi.org/10.1186/s41235-018-0139-y

Fox, J., & Weisberg, S. (2019). *An R companion to applied regression* (Third ed.). Sage. https://socialsciences.mcmaster.ca/jfox/Books/Companion/

Frauendorfer, D., Schmid Mast, M., Nguyen, L., & Gatica-Perez, D. (2014). Nonverbal social sensing in action: Unobtrusive recording and extracting of nonverbal behavior in social interactions illustrated with a research example. *Journal of Nonverbal Behavior*, *38*(2), 231–245. https://doi.org/10.1007/s10919-014-0173-5

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

Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, *57*(2), 639–683. https://doi.org/10.1037/e518612013-214

Kelley, H. H., & Michela, J. L. (1980). Attribution theory and research. *Annual Review of Psychology*, *31*, 457–501. https://doi.org/10.1146/annurev.ps.31.020180.002325

Langer, M., König, C. J., & Krause, K. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings. *International Journal of Selection and Assessment*, *25*(4), 371–382. https://doi.org/10.1111/ijsa.12191

Lievens, F., & Sackett, P. R. (2017). The effects of predictor method factors on selection outcomes: A modular approach to personnel selection procedures. *Journal of Applied Psychology*, *102*(1), 43–66. https://doi.org/10.1037/apl0000160

Lüdecke, D. (2022). sjPlot: Data visualization for statistics in social science (R package version 2.8.12) [computer software]. https://CRAN.R-project.org/package=sjPlot

Melchers, K. G., Petrig, A., Basch, J. M., & Sauer, J. (2021). A comparison of conventional and technology-mediated selection interviews with regard to interviewees' performance, perceptions, strain, and anxiety. *Frontiers in Psychology*, *11*, 603632. https://doi.org/10.3389/fpsyg.2020.603632

Michelotti, M., McColl, R., Puncheva-Michelotti, P., Clarke, R., & McNamara, T. (2021). The effects of medium and sequence on personality trait assessments in face-to-face and videoconference selection interviews: Implications for HR analytics. *Human Resource Management Journal*, *1748-8583*, 12425–12310. https://doi.org/10.1111/1748-8583.12425

Potosky, D. (2008). A conceptual framework for the role of the administration medium in the personnel assessment process. *Academy of Management Review*, *33*(3), 629–648. https://doi.org/10.5465/amr.2008.32465704

Powell, D. M., Stanley, D. J., & Brown, K. N. (2018). Meta-analysis of the relation between interview anxiety and interview performance. *Canadian Journal of Behavioural Science*, *50*(4), 195–207. https://doi.org/10.1037/cbs0000108

R Core Team. (2022). R: A language and environment for statistical computing. https://www.R-project.org/

Revelle, W. (2022). psych: Procedures for psychological, psychometric, and personality research (R package version 2.2.9) [computer software]. https://CRAN.R-project.org/package=psych

Sears, G., Zhang, H., Wiesner, W. D., Hackett, R. W., & Yuan, Y. (2013). A comparative assessment of videoconference and face-to-face employment interviews. *Management Decision*, *51*, 1733–1752. https://doi.org/10.1108/MD-09-2012-0642

Short, J., Williams, E., & Christie, B. (1976). *The social psychology of telecommunications*. Wiley.

Stanley, D. (2021). apaTables: Create American Asychological association (APA) style tables (R package version 2.0.8) [computer software]. https://CRAN.R-project.org/package=apaTables

Tsai, W.-C., Huang, T.-C., & Yu, H.-H. (2012). Investigating the unique predictability and boundary conditions of applicant physical attractiveness and non-verbal behaviours on interviewer evaluations in job interviews. *Journal of Occupational and Organizational Psychology*, *85*(1), 60–79. https://doi.org/10.1348/2044-8325.002003

Wegener, D. T., & Petty, R. E. (1995). Flexible correction processes in social judgment: The role of naive theories in corrections for perceived bias. *Journal of Personality and Social Psychology*, *68*(1), 36–51. https://doi.org/10.1037//0022-3514.68.1.36

White, J., & Behrend, T. S. (2021). Rater attributions of malfunctions in videoconference interviews. *Work, Aging and Retirement*, *9*(2), 190–203. https://doi.org/10.1093/workar/waab025

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., … Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. https://doi.org/10.21105/joss.01686

How to cite this article: Langer, M., Demetriou, A., Arvanitidis, A., Vanderveken, S., & Hiemstra, A. M. F. (2025). A quasi-experimental investigation of differences between face-to-face and videoconference interviews in an actual selection process. *Applied Psychology*, *74*(1), e12558. https://doi.org/10.1111/apps.12558