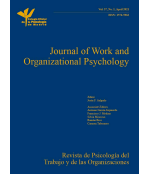




# Journal of Work and Organizational Psychology

<https://journals.copmadrid.org/jwop>



## Effects of Candidate Gender and Qualification on Hiring Recommendations in Asynchronous Video Interview Tools

Edurne Martínez-Moreno, Edurne Elgorriaga, Lorena Gil de Montes, and Olaia Larruskain-Mandiola

University of the Basque Country UPV/EHU, Spain

### ARTICLE INFO

#### Article history:

Received 1 October 2024

Accepted 29 November 2024

#### Keywords:

Asynchronous job interview tool

Rater bias

Competency

Gender

Hiring recommendation

### ABSTRACT

The main objective of this study is to examine the relative influence of candidate competencies vs. rater biases on hiring recommendations made using asynchronous video interview (AVI) tools, while considering a candidate's gender and qualifications. A  $2 \times 2$  within-subject design was employed with 151 HR professionals in Spain to explore the effects of candidate gender (female vs. male) and qualifications (highly qualified vs. semi-qualified) on hiring recommendations. Binary logistic regression and qualitative analyses revealed that although competencies play a strong role biases were the dominant factor influencing hiring recommendations for all candidates. For women, competence was a key predictor. Sociability predicted hiring recommendation of semi-qualified candidates, particularly men, for whom morality also played an important role. First impressions favoured highly qualified women, while nonverbal communication favoured highly qualified men. Consistent with role congruity theory, communal competencies were more valued in women, while agentic competencies were crucial for men.

### Los efectos del género y de la cualificación de la persona candidata en las recomendaciones de contratación en las entrevistas de trabajo asincrónicas

### RESUMEN

El objetivo principal de este estudio es examinar la influencia relativa de las competencias de la persona candidata frente a los sesgos de las y los evaluadores en las recomendaciones para el pase a la siguiente fase del proceso de selección realizadas mediante entrevistas de trabajo asincrónicas (AVI), considerando el género y el grado de cualificación de la persona candidata. Se empleó un diseño  $2 \times 2$  intrasujeto con 151 profesionales de recursos humanos en España para explorar los efectos del género de la persona candidata (mujer vs. hombre) y el grado de cualificación (muy cualificado vs. semicualificado) en las recomendaciones de contratación. Los análisis de regresión logística binaria y los análisis cualitativos revelaron que, aunque las competencias juegan un papel importante, los sesgos fueron el factor dominante que influyó en las recomendaciones de contratación para todas las personas candidatas. En las mujeres, la competencia fue un predictor clave. La sociabilidad predijo la recomendación de contratación de las y los candidatos semicualificados, especialmente en los hombres, para los cuales la moralidad también jugó un papel importante. Las primeras impresiones favorecieron a las mujeres muy cualificadas, mientras que la comunicación no verbal favoreció a los hombres muy cualificados. De acuerdo con la teoría de la congruencia de roles, las competencias comunales fueron más valoradas en las mujeres, mientras que las competencias agénticas fueron cruciales para los hombres.

#### Palabras clave:

Entrevista de trabajo asincrónica

Sesgos del evaluador

Competencias

Género

Recomendación de contratación

The COVID-19 pandemic significantly changed the world, decreasing physical contact among people and promoting the use of information and communication technologies (ICTs) for social interaction and work. Personnel selection processes are not an exception to these changes, with organizations embedding new ICT tools, such as asynchronous video interview (AVI) tools, to assess candidates' suitability for vacant positions. AVI tools are a type of online job interview conducted through web-based video platforms,

where candidates record their responses to a set of questions, which are later assessed by recruitment professionals or artificial intelligent (Dunlop et al., 2022; Liff et al., 2024; Lukacik et al., 2022). In a recent targeted review of digital selection procedures, Woods et al. (2020) concluded that the potential negative effects and biases of using digital selection procedures are still largely unknown, and it remains unclear how these methods may accentuate or mitigate arbitrariness in personnel selection processes. To address this issue, this study has

Cite this article as: Martínez-Moreno, E., Elgorriaga, E., Gil de Montes, L., & Larruskain-Mandiola, O. (2024). Effects of candidate gender and qualification on hiring recommendations in asynchronous video interview tools. *Journal of Work and Organizational Psychology*, 40(3), 164-176. <https://doi.org/10.5093/jwop2024a14>

Funding: This research was supported by a grant from Basque Government Research Groups ('Culture, Cognition, and Emotion' Consolidated Group; IT1598-22)-Correspondence: [edurne.martinez@ehu.es](mailto:edurne.martinez@ehu.es) (E. Martínez Moreno).

ISSN: 1576-5962/© 2024 Colegio Oficial de la Psicología de Madrid. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

two main purposes. First, it aims to examine whether the evaluated competencies of candidates outweigh the biases of raters in hiring recommendations for a Human Resources (HR) technician position using AVI tools within a Spanish context. Second, the study seeks to identify the most salient competencies and rater biases, taking the gender and qualifications of candidates into account.

First, our study will analyze rater behavior and candidate assessments across different genders and qualifications in controlled quasi-experimental settings to address the limited existing literature on the fairness of using AVI tools. In Spain, a study conducted by [González et al. \(2019\)](#), in which four equivalent resumes of fictitious male and female applicants were submitted to 1,372 job postings, found that women were 30% less likely than men to receive a response to their job applications. However, the study also pointed that highly qualified women without children did not experience this form of discrimination during the initial stage of the selection process, as their likelihood of receiving a response was similar to their male counterparts'. This study aims to delve into the factors that influence the selection of women in the later stages of the hiring process.

Most previous studies have focused on candidate reactions to the use of AVI tools (e.g., [Niemitz et al., 2024](#); [Roulin, Pham, et al., 2023](#); [Suen & Hung, 2023](#); [Suen & Hung, 2023](#)), and the few that have examined their effectiveness have reported inconsistent results (e.g., [Gorman et al., 2018, 2016](#); [Langer et al., 2017](#); [Roulin, Lukacik, et al., 2023](#); [Suen et al., 2019](#); [Torres & Gregory, 2018](#)). Second, this study seeks to determine whether the competencies required for men and women for the position of HR technician are the same, or whether, on the contrary, as established by the role congruity theory ([Eagly & Karau, 2002](#); [Heilman et al., 2024](#)), they will be evaluated differently by the raters. Third, we will explore how rater biases operate and which biases are most relevant, considering the AVI tools format and methods used to study rater biases in Spain.

While many studies have reported information on the fairness of face-to-face (FTF) job interviews (e.g., [Alonso & Moscoso, 2017](#); [Alonso et al., 2017](#)), little is known about how technology has transformed selection processes. This study aims to provide new insights into the use of AVI tools in Spain, an area that has not been previously explored. Finally, studies in this field have traditionally used quantitative and deductive methods to estimate the suitability of candidates. This study contributes to a better understanding of the decision-making processes in personnel selection by inductively analyzing the text of evaluative judgments, thus making it possible to identify the main thematic lines on which raters base their decision to recommend or not to recommend candidates.

### Candidate Competencies vs. Rater Biases in AVI Tools

One of the primary concerns for HR professionals and scholars in the selection process is evaluating candidates effectively and fairly. To achieve this, various tools have been developed, with AVI being one such tool used for preselecting candidates. An AVI tool typically comprises four or five questions that candidates cannot preview, although they have 30 seconds to prepare their answers. There are two types of question formats (video and text), but the question format does not influence interview performance ([Niemitz et al., 2024](#)). Candidates then have 2-3 minutes to respond and record their answers using a webcam through various proprietary software platforms. Once the recording starts, it cannot be interrupted or re-recorded ([Dunlop et al., 2022](#)). These interviews are perceived as more useful and easier to use than other preselection tools ([Basch et al., 2022](#)) and offer several advantages: they are fast, economical, and timesaving for employees ([Dunlop et al., 2022](#)).

Despite the advantages of AVI tools, job raters may incur two types of errors: false positives (recommending an unqualified candidate) and false negatives (not recommending a qualified candidate). False

positives are considered worse than false negatives because they can seriously harm both the organization and the person who has been selected. According to the dual-process theory, raters can process candidates' information in two distinct ways ([Derous et al., 2016](#)). On the one hand, some information in job interviews can be processed easily through heuristically driven automatic impression formation, known as Type 1. On the other hand, raters can process candidates' information laboriously and consciously, with their decisions based on controlled judgments (Type 2). Raters aim to conduct their work through Type 2 processes, employing observation sheets, structured interviews, multiple raters, and recordings of interviews to review before making a decision. The use of AVI tools integrates various functionalities, enabling interviewers to conduct personnel selection processes in a deliberate and thoughtful manner. Another issue with AVI tools is that their design may increase the use of stigma-laden heuristics that bias assessments. In an AVI context, raters have the discretion to decide when to stop watching an interview, potentially denying candidates the opportunity to correct a negative first impression ([Lukacik et al., 2022](#)). Furthermore, candidates may not carefully select the setting and timing for recording their answers, inadvertently including background elements that trigger rater biases ([Roulin, Lukacik, et al., 2023](#)). Additionally, the manner in which candidates interact with the camera during their responses can negatively influence raters' perceptions of motivation, interest in the position, or professionalism ([Lukacik et al., 2022](#)).

In this regard, several studies have confirmed that first impression significantly influences hiring recommendations (e.g., [Buijsrogge et al., 2020](#); [Martín-Raugh et al., 2023](#); [Swider et al., 2016](#)). First impression is easily influenced by a candidate's attractiveness and non-verbal cues, as these cues are readily accessible and can be processed effortlessly. Attractive candidates exhibit more effective non-verbal cues and consequently receive higher scores than less attractive ones due to their greater sense of power ([Tu et al., 2022](#)). Research has demonstrated that both attractiveness and non-verbal communication significantly influence rater and interviewer scores ([Koutsoumpis et al., 2024](#); [Martín-Raugh et al., 2023](#); [Nault et al., 2020](#)). The literature on lookism, or discrimination based on physical appearance, has established that attractiveness influences perceptions of competency, leading to favoritism for attractive candidates ([Hoffman, 2024](#); [Nault et al., 2020](#); [Niu, 2024](#); [Pireddu et al., 2022](#)). The competence stereotype pertains to qualities such as capability and assertiveness ([Fiske, 2018](#); [Fiske et al., 2002](#)). Sociability and morality stereotypes are considered central dimensions for forming a positive view of a person ([Leach et al., 2007](#)). Sociability relates to qualities of friendliness, while morality pertains to sincerity and trustworthiness. Additionally, the literature highlights a strong association between interview performance, professional appearance, and non-verbal communication ([Martín-Raugh et al., 2023](#)).

Few studies have investigated whether a rater's competencies or biases have a greater influence on the scores awarded and, consequently, on hiring recommendations in the context of AVI. A recent study ([Scott & Roulin, 2024](#)), in which 276 senior raters examined the effect of candidates' background settings (home, office, and blurred settings) and response quality on overall rating scores, found that first impressions and candidates' response quality predict overall rating scores. In a hospital-based study comprising 517 observations, findings indicated that scores in communication and conflict resolution competencies significantly influenced overall candidate evaluations, surpassing factors such as aesthetics or procedural order during the use of AVI tools ([Torres & Gregory, 2018](#)).

It is established that certain competencies, defined as underlying characteristics of an individual that causally contribute to superior performance in a job or situation ([Spencer & Spencer, 2008](#)), are desirable for employers depending on the specific job position ([Liff et al., 2024](#)). A field study conducted in Spain with 37 active HR professionals identified six key competencies essential for the

development of HR technicians: teamwork, customer orientation, planning, communication, flexibility, and collaboration (Pereda Marín et al., 2003). The present study focuses on assessing three competencies: teamwork, planning oriented to conflict resolution, and flexibility. To evaluate these competencies, the study used written questions centered on past behaviors, supplemented with observation sheets and the Behaviorally Anchored Rating Scale (BARS). This methodology ensures a structured and standardized approach to assessing each candidate's abilities, facilitating a fair and comprehensive assessment process. By focusing on these competencies, the study aims to identify the most salient attributes that contribute to successful performance in an HR technician role, thereby informing hiring recommendations in the context of AVI.

Building on the findings of the previously mentioned research, which indicate that competency scores hold greater significance than other factors in candidate recommendations, we anticipate that the recommendations for qualified candidates are expected to be driven by the scores they achieve in the competencies relevant to the job position. In contrast, false positives—such as raters endorsing unqualified candidates—will primarily result from rater biases.

*Hypothesis 1:* Candidate competencies (teamwork, planning oriented to conflict resolution and flexibility) will primarily influence hiring recommendations for qualified candidates more than factors related to raters' biases (first impression, attractiveness, non-verbal communication, and stereotypes). Conversely, factors associated with raters' biases will have a stronger influence on hiring recommendations for semi-qualified candidates.

### Candidate Competencies and Gender in Hiring Recommendations in AVI Tools

In the workplace, the evaluation and recruitment of women has been negatively influenced by gender stereotypes. Gender stereotypes can be categorized into two main dimensions: agency and communality (Heilman et al., 2024). Agency, which is typically associated more with men, includes attributes related to task orientation and goal achievement. In contrast, communality, more commonly associated with women, encompasses traits such as kindness and concern for others.

This study focuses on examining teamwork and flexibility competencies, considered as communal competencies. Teamwork encompasses problem-solving within a group, commitment to common goals over individual interests, and sharing resources and information. Flexibility involves adapting behavior to varying situations, maintaining effectiveness across different environments, tasks, responsibilities, and interactions with diverse people. In contrast, planning oriented to conflict resolution is viewed as an agentic competency, focusing on objective-oriented planning such as prioritization, establishing action plans to achieve objectives, and implementing appropriate control and follow-up measures. Agency has been subdivided into competence and assertiveness, while communality into morality and warmth (Abele et al., 2016).

In the context of AVIs, a quasi-experimental study conducted in Germany (Kroll & Ziegler, 2016) analyzed fairness in personnel selection for waiter positions among candidates of different genders and ethnicities. The study found no significant differences in teamwork and communication competency scores between men and women candidates. However, candidates with a Turkish background received higher scores in both compared to native German candidates for this particular job position. This study indicates that there are no differences in how women and men are assessed regarding said competencies. However, it does not establish that these competencies will have the same impact on the candidates' hiring decisions. Therefore, and based on role congruity theory (Eagly & Karau, 2002; Heilman et al., 2024), we expect that communal competencies,

such as teamwork and flexibility, will be prioritized in women for recommending advancement to the next phase of the selection process, while for men the emphasis will be on agentic competencies, like planning oriented to conflict resolution.

*Hypothesis 2:* Competency scores in (a) teamwork and (b) flexibility will have a greater influence on hiring recommendations for women candidates, whereas (c) planning oriented to conflict resolution will have a greater influence on hiring recommendations for men candidates.

### Rater Biases and Candidate Gender in Hiring Recommendations in AVI Tools

Culturally, beauty or attractiveness has historically been a more valued trait in women than in men. Given that physically attractive individuals are often perceived as having more positive traits, such as sociability, honesty, intelligence, and life success (Dion et al., 1972), this perception should provide an advantage for attractive women in selection processes. However, conclusions regarding whether attractiveness has a differential impact on men and women remain inconclusive. A meta-analysis by Hosoda et al. (2003) found no significant gender differences, a result further corroborated by more recent studies (Pireddu et al., 2022). Some studies suggest that attractiveness may benefit men (Johnson et al., 2010; Sheppard & Johnson, 2019), while others argue that it plays a more significant role in the evaluation of women (Mao et al., 2024; Turkmenoglu, 2020).

In Spain, a recent study (Cuadrado et al., 2024) revealed that participants exhibited more favourable attitudes toward women perceived as more attractive compared to those considered less attractive. This study also found that women were considered more competent when applying for male-typed jobs than for neutral-typed jobs. Furthermore, Watkins and Johnston (2000) found that physical attractiveness did not influence evaluations when a job application was exceptional, but it did influence evaluations when the applicant was ordinary. In this regard, Hosoda et al. (2003) found that attractiveness may become decisive when decision-making is difficult or multiple candidates are evaluated consecutively.

Another variable explored to understand biases in job interviews is non-verbal communication (Frauendorfer & Mast, 2015). A recent meta-analysis (Martín-Raugh et al., 2023), which synthesizes findings from 63 studies conducted over the past 70 years, reveals a positive correlation between non-verbal cues and candidate evaluations. The analysis highlights professional appearance, eye contact, and head movements as particularly influential factors in candidate assessments. Additionally, gender differences were noted, with facial expressions and professional appearance having a greater impact on the evaluation of women candidates.

The few studies that have examined the influence of aesthetics in an AVI context are promising. A field study in Taiwan (Suen et al., 2019) found that first impression and physical attractiveness significantly affected structured interview ratings (including AVI tools, synchronous video interviews, and artificial intelligence), but these effects were diminished in the case of AVI tools. They also found no significant differences between women and men regarding both first impression and physical attractiveness. Koutsoumpis et al. (2024), who also identified that attractiveness bias was present in AVI contexts, found similar results: participants considered attractive, regardless of their gender, received higher scores for their interview performance.

Despite the potential positive outcomes of AVI tools, we anticipate that rater bias, influenced by first impressions, attractiveness, and non-verbal communication, will influence hiring decisions, particularly for women candidates. This expectation is based on the fact that raters must evaluate multiple candidates, and as noted by Hosoda et al. (2003), attractiveness often plays a decisive role in hiring

recommendations. Supporting this, recent findings by Cuadrado et al. (2024) in Spain show that attractive women tend to gain a competitive advantage over other candidates in selection processes. Furthermore, Martín-Raugh et al. (2023) emphasizes the close relationship between attractiveness and non-verbal communication. Therefore, we hypothesize that:

*Hypothesis 3:* a) First impression, b) attractiveness (b), and c) non-verbal communication will have a greater influence on hiring recommendations for women candidates compared to men candidates.

In the workplace, women's evaluation and hiring have been adversely affected by gender stereotypes related to competence and sociability. Traditionally, men have been depicted as more competent and brilliant (more aggressive, competitive, dominant, assertive), whereas women have been perceived as more sociable (more caring, empathetic, sensitive, passive) (e.g., Correll, 2017; Duehr & Bono, 2006; Fiske, 2018; Hentschel et al., 2019). There is currently a slight shift in these gender stereotypes. Recent studies have indicated that women candidates are perceived as more competent than men in some cases, and occasionally more trustworthy and socially stable (Leach et al., 2017), although this is not always consistent (Roulin, Lukacik, et al., 2023). Other studies suggest that women who deviate from these stereotypes may be viewed negatively (e.g., Corrington & Hebl, 2018; Rudman & Phelan, 2008): whereas men are expected to be competitive, women are often expected to balance presenting their competence with maintaining warmth and likability. Along these lines, two studies conducted in Italy revealed that competence is the most critical stereotype influencing the hiring recommendation for men candidates, whereas for women candidates all three stereotypes (competence, sociability, and morality) gained importance (Moscatelli et al., 2020), particularly influencing the overall impression through factors like facial morality, facial competence, and attractiveness (Menegatti et al., 2021). Similarly, an audit study by Quadlin (2018), which involved 261 human resource professionals and examined whether men and women receive equal returns on academic performance in hiring, revealed that competence and commitment are valued in hiring decisions only for men. In the case of women, other characteristics, such as being sociable and outgoing, are more appreciated, leading to a penalty for women with good grades. However, a study from the UK (Pireddu et al., 2022) highlighted that attractiveness and competence were equally relevant for hiring decisions for both men and women, while morality and sociability were more critical in assessing men than women for leadership positions. Given the contradictory findings and the lack of testing on the impact of stereotypes on recommendation decisions using AVI tools, we hypothesize that:

*Hypothesis 4:* a) Sociability and b) morality stereotypes will have a greater influence on hiring recommendations for women candidates, whereas c) the competence stereotype will have a greater influence on hiring recommendations for men candidates.

Regarding the thematic profiles that emerge from the participants' evaluations, we adopt a fundamentally exploratory approach due to the absence of previous work in this area. Generally, we expect to find compatibility between the hypotheses presented in the quantitative part of the study.

## Method

### Sample

The study involved 186 HR professionals located in Spain. Thirty-five participants were excluded from the analysis due to errors in candidate assessments, or because they skipped or failed to respond to assessments for at least three candidates. After data cleaning,

the final sample consisted of 151 HR professionals. Among these participants, 63.3% were women ( $n = 95$ ) with a mean age of 32.41 years ( $SD = 10.20$ ).

In terms of professional roles, 53.5% identified as HR technicians, primarily employed in companies (38.7%,  $n = 58$ ) or HR consulting firms (20.0%,  $n = 30$ ) with more than 250 employees (35.7%,  $n = 46$ ). Regarding educational background and HR experience, a majority held a master's degree (55.7%,  $n = 83$ ) or a bachelor's degree/diploma (38.9%,  $n = 58$ ), and most had less than 5 years of work experience (59.7%,  $n = 89$ ). While 63.8% ( $n = 115$ ) reported frequently or very frequently using competency-based job interviews, only 3.9% ( $n = 5$ ) reported using AVI tools with the same frequency (See Table 1).

**Table 1.** Demographic Profile of Sample

Variable & Descriptive	Frequency	%
Gender		
Female	95	63.3
Male	55	36.7
Education level		
Bachelor's, degree or diploma	58	38.9
Master	83	55.7
Doctoral degree	7	4.7
Other	1	0.7
HR work experience		
Less than 1 year	51	34.2
Between 1 year and 5 years	38	25.5
Between 5 and 10 years	20	13.4
Between 10 and 20 years	25	16.8
More than 20 years	15	10.1
Work experience in organization types		
Company	58	38.7
HR Consulting	30	20.0
TEA (Temporary Employment Agency)	17	11.3
Unemployed	22	14.7
Other	23	15.3
Job position		
HR Director	4	3.1
HR Manager	18	14.0
Chief of staff	3	2.3
Recruitment or HR technician	69	53.5
Other	35	27.1
Number of employees current organization		
Less than 10 employees	19	14.7
Between 10 and 50 employees	31	24.0
Between 50 and 250 employees	33	25.6
More than 250 employees	46	35.7

*Note.*  $N = 151$ . Participants were on average 32.41 years old ( $SD = 10.20$ ), with a maximum age of 70 and a minimum age of 21 years old. Number of participants with missing data in the analysed sample: gender: 1, age: 1, education level: 2, HR work experience: 2, work experience in organization types: 1, job position: 22 and number of employees current organization: 22.

### Design of the Study

The study employed a quasi-experimental design utilizing discrimination-testing methodology. This design was quasi-experimental because participants were randomly assigned to one of four experimental conditions, while considering their work experience to ensure representation of both senior and junior professionals across all conditions. The discrimination-testing methodology involved presenting two subjects with similar profiles and skills for evaluation. The subjects differed in a specific characteristic, such as gender, allowing for the detection of discrimination based on that characteristic.

In sum, this study was a 2 (gender) x 2 (candidate qualification) within-subjects design. Each participant evaluated 4 candidates varying by gender and qualification. The primary dependent variable was the decision to recommend or not recommend the candidate for advancement to the next phase of the selection process at organization. In addition to the final recommendation decision, each candidate was evaluated on measures of rater bias and candidate competencies. Bias measures included responses to competence, morality, and sociability stereotypes, as well as first impression, attractiveness and non-verbal communication of candidates. Competency measures included teamwork, planning oriented to conflict resolution, and flexibility scores.

### Quasi-experiment Preparation

The principal researcher interviewed 9 HR professionals, 4 with proven professional experience in the sector and 5 with fewer than 6 months of experience. The interview consisted of three questions related to teamwork (“Tell me about the last time you worked in a team, your role, your experience”), planning oriented to conflict resolution (“Describe a tense situation you had with a client”), and flexibility competencies (“Describe a situation you had to face suddenly”). The responses from the 9 professionals were disaggregated, resulting in a list of 27 responses—9 for each competency—with 4 from experienced professionals and 4 from less experienced professionals. This step aimed to dissociate the three responses from each interviewed person to ensure heterogeneity and equivalence of the material. The 27 responses were assigned to evaluate two sector experts, independent and external to the study. Each professional received an observation rubric with BARS to evaluate each response on teamwork, planning oriented to conflict resolution and flexibility competencies. The 4 responses with the highest average scores and the 4 responses with the lowest average scores were selected.

Based on the 8 selected responses, 8 interview models were created (4 interview models for qualified candidates and 4 interview models for semi-qualified). All models, whether for qualified or semi-qualified candidates, had the same introduction and farewell questions. This ensured control over years of previous experience, current job quality, and interview conclusion format. 5 men and 5 women were recruited to act as candidates in AVI tools and to record video interpretations for the quasi-experiment. To ensure internal and external validity, the same actor/actress portrayed both the qualified and semi-qualified candidates. This approach maintained neutral and similar contextual and personal characteristics across recordings to mitigate any influence of the recording context on evaluations. Each AVI lasted approximately 10 minutes. In total, 20 AVIs were recorded (video interpretations), manipulating candidate gender and qualification.

A pilot study was conducted with 9 PhD students from the University of the Basque Country to assess video interpretations, observation sheets, and questionnaires. As a result, 8 AVIs (2 women and 2 men candidates, each participating in two interviews) were excluded, because interviewees were perceived as fake.

### Procedure

All participants identified themselves as HR professionals and were contacted either through LinkedIn or email (UPV/EHU alumni now working in HR). They were invited to take part in a selection process using AVI tools. Upon confirming their participation, each participant received an email containing a link to a dedicated webpage created for the study.

The website featured 4 AVI tools: one each for a qualified man candidate, qualified woman candidate, semi-qualified man candidate, and semi-qualified woman candidate. Study participants

were instructed to evaluate these candidates as they would in a real selection process for an HR technician role. They were tasked with assessing the suitability of each interviewee for the position.

Prior to evaluating the candidates, participants were provided with an introduction to the organization where the HR technician would potentially be employed. They observed the competencies and skills of each candidate using observation sheets and evaluated these using BARS. Subsequently, participants had to make a decision whether to recommend each candidate to proceed to the next stage of the selection process. Participants in the study were provided with 4 different links. While for some participants, some actors were qualified, for others they appeared as semi-qualified, and vice versa. To control the possible strange variables, the order of the presentation of the candidates was carried out by means of randomization.

In all cases, we provided information about the research objectives and permission to use the data was requested (participants signed the informed consent), and anonymity and confidentiality were ensured. This study complied with all the ethical requirements in accordance with the University of the Basque Country, UPV/EHU (M10/2019/157), as well as national and international (APA) ethical guidelines.

### Measures

#### Dependent Variables

**Hiring Recommendation.** Own elaboration mono-item scale that designates recommendation frequency regarding each hiring candidate to move on to the next phase of the selection process. Participants answered to the single item “Would you recommend this candidate to move on to the next phase of the selection process?”, where 1 = yes, 0 = no. Subsequently, in order to further explore the reasons for moving the candidate to the next phase, an open question was used to ask them to reason their decision: “Explain why you chose to recommend or not recommend.”

#### Independent Variables

**Competencies.** Participants assessed the candidate in the dimensions of teamwork, planning oriented to conflict resolution and flexibility using BARS for the assessment of candidates (see Table 2). BARS scales range from 0, no evidence of the competency, to 4, high evidence of the competency; but it was recodified in a 5 point Likert scale for further analysis. Inter-rater agreement coefficients were .73 for highly qualified woman, .59 for somewhat qualified woman, .67 for highly qualified man, and .63 for somewhat qualified man.

**First Impression.** In order to evaluate the initial impression of each candidate, after the introductory question to each applicant, participants answered a 4-item scale adapted from the Swider et al. (2016) scale. An example item is: “This candidate appears to be very qualified.” The inter-rater agreement coefficients were as follows: .85 for highly qualified women, .87 for somewhat qualified women, .88 for highly qualified men, and .85 for somewhat qualified men.

**Attractiveness.** To analyze the perception of the participants regarding the physical appearance of the candidates, a 4-item scale was used adapted from the Boor et al. (1983) scale. An example of an item is: “This candidate is physically attractive.” The inter-rater agreement coefficients were: .85 for highly qualified woman, .81 for somewhat qualified woman, .85 for highly qualified man, and .81 for somewhat qualified man.

**Non-verbal Communication.** The ECO-CNV scale by Roso-Bas et al. (2017) which evaluates body expression, facial expression, gaze at the camera, naturalness of speech and fluency was adapted. An example of item is: “Appropriate gestures illustrating speech.” The

**Table 2.** Observation Sheet with Behaviorally Anchored Rating Scale (BARS)

Evidence level	Competencies
	<b>TEAMWORK:</b> Perform tasks and solve problems as part of a group, establishing commitments to achieve a common goal above individual interests and sharing resources and information.
No evidence	There is no evidence of the candidate in this competence. It cannot be assessed.
Level 1	Does not ask team members to contribute and provide ideas/suggestions to the team.
Level 2	Asks team members to contribute and provide ideas/suggestions to the team, but does not maintain a team approach to problems.
Level 3	Asks team members to contribute and provide ideas/suggestions to the team, but either does not maintain a team approach towards problems or does not share information or help colleagues.
Level 4	Asks team members to contribute and provide ideas/suggestions to the team, helps them if they need it and also keeps them informed at all times, fostering a team approach towards problems.
	<b>PLANNING ORIENTED TO CONFLICT RESOLUTION:</b> Define priorities, establishing the necessary action plans to achieve the set objectives and the appropriate control and monitoring measures.
No evidence	There is no evidence of the candidate in this competence. It cannot be assessed.
Level 1	Does not identify the necessary priorities for the established objectives.
Level 2	Identifies the priorities to achieve the set objectives, but does not establish the appropriate action plans.
Level 3	Identifies the priorities for achieving the set objectives and establishes the appropriate action plans, but does not establish control and monitoring measures.
Level 4	Identifies the priorities to achieve the set objectives, and establishes the appropriate action plans, as well as control and monitoring measures.
	<b>FLEXIBILITY:</b> Modifies behavior, adapting it to situations of change or ambiguity and maintaining effectiveness in different settings, with different tasks, responsibilities and people.
No evidence	There is no evidence of the candidate in this competence. It cannot be assessed.
Level 1	Not capable of modifying their behavior to adapt it to situations of change or ambiguity.
Level 2	Capable of modifying their behavior to adapt it to situations of change or ambiguity, but is neither effective nor efficient in such situations.
Level 3	Capable to modify behaviour to adapt to changing or ambiguous situations in an efficient and effective manner, but unable to take on or perform different roles as required by the task or the moment.
Level 4	Capable to modify their behavior to adapt it to situations of change or ambiguity in an effective and efficient manner, and is also able to assume or perform different functions as required by the task or the moment and remain effective.

inter-rater agreement coefficients were .74 for women categorized as highly qualified, .67 for somewhat qualified woman, .80 for man classified as highly qualified, and .76 for evaluated as somewhat qualified man.

**Stereotypes.** To assess stereotypes about men and women, a 9-item scale was elaborated from the work of Fiske et al. (2018), Fiske et al. (2002), and Leach et al. (2007; 2017). Responses were answered on a 5-point Likert-type response scale (1 = *not at all*, 5 = *a lot*). Three items measure morality (e.g., “They are honest”), another three items measure sociability (e.g., “They are kind”) and three more items measure competence (e.g., “They are intelligent”). The inter-rater agreement coefficients for the morality stereotype were .84, .71, .90, and .71 for highly qualified woman, somewhat qualified woman, highly qualified man, and somewhat qualified man, respectively. Similarly, for the sociability stereotype, the coefficients were .76, .76, .73, and .79 for highly qualified woman, somewhat qualified woman, highly qualified man, and somewhat qualified man, respectively. Lastly, for the competence stereotype, the coefficients were .72, .81, .78, and .79 for highly qualified woman, somewhat qualified woman, highly qualified man, and somewhat qualified man, respectively.

The response scale for all instruments were a 5-point Likert scale where 1 = *strongly disagree* and 5 = *strongly agree*.

### Manipulation Check Variable

**Candidate Qualification.** To check the manipulation of perceived qualification of the candidates, we asked participants to indicate the degree to which they consider the candidate is qualified for the vacant position. Responses were answered on a 5-point Likert-type response scale (1 = *unqualified*, 5 = *fully qualify*).

### Analysis

Quantitative data analysis was performed using IBM SPSS Statistics 27. First, preliminary analysis were performed: descriptive analysis and Cronbach alpha. The general assumption of multicollinearity

between the variables studied was calculated using Spearman's rho correlations. No major violations were found, as there is no correlation greater than .70 between the recommendation to proceed to the next stage of the selection process and other variables.

To determine the experimental manipulation, repeated measure ANOVAs, Cochran Q and McNemar tests were calculated. We conducted repeated measure ANOVAs to examine if there were statistically differences in qualification perceptions between four conditions. Cochran Q tests were performed to test raters are more likely to select qualified candidates than semi-qualified candidates. In addition, McNemar tests were performed to determine the differences between candidates (Table 3). As there are 4 profiles (p1, p2, p3, and p4) the possible comparisons are 6: p1-p2, p1-p3, p1-p4, p2-p3, p2-p4, and p3-p4.

Binary logistic regression models were run to test the hypotheses of the study. Four separate binary logistic regression models with forward Wald to control for confounding were run to test hypothesis 1. The remaining hypotheses of the study were tested using binary logistic regression calculations with the enter method. The relationship between the predictors and the dependent variable was estimated using the odd ratio (OR) statistic. Values greater than 1 indicate that an increase in the predictor variable is associated with an increased likelihood of recommendation, while values less than 1 are associated with a decreased likelihood of recommendation.

To delve deeper into the content of the value judgments of the AVI tools made by the participants, all open-ended responses to the question “Explain why you chose to recommend or not recommend” were used as a corpus for lexical analysis using the Iramuteq software (version 0.7 alpha 2). By using an automated form of analysis, the study reduces reliability and validity problems typically associated with text analysis (Klein & Licata, 2003). The analysis is based on a contingency table resulting from associating portions of text (Elementary Context Units, ECUs) with each word. From this contingency table, a matrix of squared distances is generated, such that ECUs are considered closer if they share common words (Reinert, 1986). The program proceeds with a

**Table 3.** Findings for Nonparametric McNemar Test to Identify Which Candidate Profiles Differ between Them

	Recommendation Frequency		Profile Comparison	Differential Value of Recommendation between Candidates (McNemar test)	p
	%	N			
Profile 1 (p1) Qualified woman	91.2%	125	p1-p2 (Qualified woman-Qualified man)	0.47	.700
Profile 2 (p2) Qualified man	89.0%	113	p1-p3 (Qualified woman-Semi-qualified woman)	51.57	.001
Profile 3 (p3) Semi-qualified woman	48.5%	64	p1-p4 (Qualified woman-Semi-qualified man)	101.04	.001
Profile 4 (p4) Semi-qualified man	16.9%	24	p2-p3 (Qualified man-Semi-qualified woman)	41.67	.001
			p2-p4 (Qualified man-Semi-qualified man)	92.04	.001
			p3-p4 (Semi-qualified woman-Semi-qualified man)	30.31	.001

**Table 4.** Spearman's Correlation Matrix for Qualified Woman

	Mean	SD	Qualified Woman											
			1	2	3	4	5	6	7	8	9	10		
1. Hiring recommendation	--	--	--											
2. Teamwork	3.927	1.381	.372**	--										
3. Planning oriented to conflict resolution	3.745	1.207	.328**	.461**	--									
4. Flexibility	4.080	1.244	.442**	.496**	.486**	--								
5. First impression	4.084	0.650	.416**	.322**	.326**	.456**	--							
6. Attractiveness	3.306	0.687	.114	.233**	.093	.164	.154	--						
7. Non verbal communication	4.295	0.644	.323**	.296**	.142	.367**	.530**	.229**	--					
8. Competence	4.213	0.591	.417**	.505**	.499**	.629**	.519**	.162*	.334**	--				
9. Sociability	4.130	0.505	.275**	.397**	.336**	.412**	.239*	.436**	.365**	.455**	--			
10. Morality	4.056	0.600	.238**	.403**	.408**	.431**	.357**	.147	.368**	.561**	.648*	--		

\* $p < .05$ , \*\* $p < .01$ .

**Table 5.** Spearman's Correlation Matrix for Qualified Man

	Mean	SD	Qualified Man											
			1	2	3	4	5	6	7	8	9	10		
1. Hiring recommendation	--	--	--											
2. Teamwork	4.142	1.221	.444**	--										
3. Planning oriented to conflict resolution	3.714	1.259	.374**	.473**	--									
4. Flexibility	4.075	1.128	.317**	.456**	.435**	--								
5. First impression	3.778	0.737	.109	.135	.270**	.221*	--							
6. Attractiveness	3.297	0.675	-.163	.106	.173*	.160	.344**	--						
7. Non verbal communication	4.015	0.732	.257**	.212*	.209*	.327**	.516**	.359**	--					
8. Competence	4.042	0.634	.281**	.511**	.344**	.484**	.256**	.270**	.320**	--				
9. Sociability	4.124	0.536	.163	.263**	.218*	.286**	.181*	.224**	.336**	.489**	--			
10. Morality	3.998	0.665	.270**	.394**	.273**	.317**	.096	.153	.308**	.512**	.610*	--		

\* $p < .05$ , \*\* $p < .01$ .

top-down hierarchical cluster analysis, resulting in sets of ECUs that best differentiate the vocabulary and assist in interpretation. Following previously established procedures (Idoaga Mondragon et al., 2023), the most significant vocabulary in each class was selected using the following criteria: an expected value of the word greater than 5, a chi-square association statistical test in each class, and the word appearing in the class in a percentage greater than 50%.

## Results

### Checking Manipulation

Raters were asked about the perceived degree of qualification of the candidates. Both qualified candidates (woman,  $X = 4.167$ ,  $SD = 0.760$ ; man,  $X = 4.015$ ,  $SD = 0.828$ ) were perceived as highly qualified, while the semi-qualified candidates (woman,  $X = 3.264$ ,  $SD = 0.879$ ; man,  $X = 2.633$ ,  $SD = 0.879$ ) were perceived as low or somewhat qualified. Our results showed that there were significant differences between the candidates regarding hiring recommendation,  $Q(3) = 191.354$ ,  $p < .001$ . Both highly qualified candidates (woman, 91.2%,  $n = 125$ ; man, 89.0%,  $n = 113$ ) were recommended

to a greater extent than semi-qualified candidates (woman, 48.2%,  $n = 64$ ; man, 16.9%,  $n = 24$ ). Therefore, experimental manipulation worked well.

### Hypothesis Testing

Correlation analyses are presented below to identify the variables associated with the recommendation of highly qualified woman (Table 4), highly qualified man (Table 5), semi-qualified woman (Table 6) and semi-qualified man (Table 7).

Our results of four binary logistic regression indicated that the final step of the model explained 76.2% (Nagelkerke's  $R^2$ ) of the variance for highly qualified women (step 5), 54.6% (Nagelkerke's  $R^2$ ) for highly qualified men (step 6), 56.1% (Nagelkerke's  $R^2$ ) for semi-qualified women (step 7), and 34.9% (Nagelkerke's  $R^2$ ) for semi-qualified men (step 7).

Hypothesis 1 expected that candidate competences will primarily influence hiring recommendations for highly qualified candidates more than factors related to raters' biases. Conversely, for semi-qualified candidates, factors associated with raters' biases will have a stronger influence on hiring recommendations. The variables predicting the hiring recommendations of candidates are

**Table 6.** Spearman's Correlation Matrix for Semi-Qualified Woman

	Mean	SD	Semi-qualified Woman									
			1	2	3	4	5	6	7	8	9	10
1. Hiring recommendation	--	--	--									
2. Teamwork	2.471	1.349	.292**	--								
3. Planning oriented to conflict resolution	2.691	1.380	.265**	.347**	--							
4. Flexibility	2.978	1.462	.525**	.380**	.280**	--						
5. First impression	4.050	0.670	.273**	.148	.206*	.257**	--					
6. Attractiveness	3.097	0.620	.202*	.076	.029	.125	.260**	--				
7. Non verbal communication	3.893	0.647	.297**	.228**	.228**	.305**	.264**	.152	--			
8. Competence	3.462	0.682	.566**	.327**	.315**	-.143	.557**	.245**	.323**	--		
9. Sociability	3.413	0.707	.508**	.309**	.155	-.143	.444**	.200*	.312**	.435**	--	
10. Morality	3.865	0.606	.288**	.175*	.231**	-.130	.310**	.125	.404**	.494**	.462**	--

\* $p < .05$ , \*\* $p < .01$ .

**Table 7.** Spearman's Correlation Matrix for Semi-qualified Man

	Mean	SD	Semi-qualified Man									
			1	2	3	4	5	6	7	8	9	10
1. Hiring recommendation	--	--	--									
2. Teamwork	2.352	1.060	.311**	--								
3. Planning oriented to conflict resolution	2.352	1.178	.256**	.461**	--							
4. Flexibility	2.218	1.203	.263**	.362**	.282**	--						
5. First impression	3.359	0.813	.230**	.371**	.270**	.300**	--					
6. Attractiveness	3.074	0.629	.168*	.153	.090	-.030	.145	--				
7. Non verbal communication	3.471	0.837	.297**	.427**	.293**	.171*	.428**	.054	--			
8. Competence	3.166	0.618	.271**	.200*	.274**	.135	.368**	.278**	.254**	--		
9. Sociability	2.909	0.716	.363**	.422**	.187*	.291**	.436**	.290**	.380**	.367**	--	
10. Morality	3.088	0.748	.322**	.348**	.313**	.208*	.262**	.204*	.457**	.300**	.348*	--

\*\*  $p < .01$ , \*  $p < .05$ .

detailed in Table 8, ranked by effectiveness. For highly qualified women, competence ( $OR = 342.893$ ), first impression ( $OR = 13.429$ ), and teamwork ( $OR = 2.406$ ) primarily account for their hiring recommendations. Among highly qualified men, non-verbal communication ( $OR = 3.711$ ) emerges as the most influential predictor, followed by planning oriented to conflict resolution ( $OR = 2.039$ ), teamwork ( $OR = 2.013$ ), and attractiveness ( $OR = 0.149$ ). For semi-qualified women, competence ( $OR = 6.671$ ), sociability ( $OR = 3.366$ ), and flexibility ( $OR = 1.534$ ) play significant roles. In the case of semi-qualified men, sociability ( $OR = 3.397$ ), morality ( $OR = 2.593$ ), and planning oriented to conflict resolution ( $OR = 1.640$ ) are noteworthy predictors. Therefore, hypothesis 1 is partially supported, as anticipated, given that rater biases strongly influence the hiring recommendations for both semi-qualified and qualified candidates.

Hypothesis 2 expected that competency scores in a) teamwork and b) flexibility will have a greater influence on hiring recommendations for women candidates, whereas c) planning oriented to conflict resolution will have a greater influence on hiring recommendations for men candidates. According to Table 8, men perceived as planning oriented to conflict resolution had a 2.039 times higher probability of being hiring recommended if they were highly qualified and a 1.640 times higher probability if they were semi-qualified. For highly qualified women, teamwork competency emerges as the most predictive ( $OR = 2.406$ ), whereas for semi-qualified women, flexibility is prominent ( $OR = 1.534$ ). Therefore, hypothesis 2 is partially supported by our data.

Regarding rater biases, our data indicated that first impression predicts the hiring recommendations of qualified women ( $OR = 13.429$ ), while non-verbal communication predicts the hiring recommendations of qualified men ( $OR = 3.711$ ). Surprisingly, qualified men who are considered attractive have .149 times lower probability of being recommended. Thus, hypothesis 3 is partially confirmed.

Our results also showed that highly qualified women perceived as competent had 342.892 times higher probability of being recommended for hiring, and 6.671 times higher in the case of semi-qualified women. Another difference between highly qualified and semi-qualified candidates is that sociability is significant only for semi-qualified candidates (women,  $OR = 3.366$ ; men,  $OR = 3.397$ ). For semi-qualified men, being moral increases the likelihood of recommendation by 2.593 times. Consequently, Hypothesis 4 is not supported by our data.

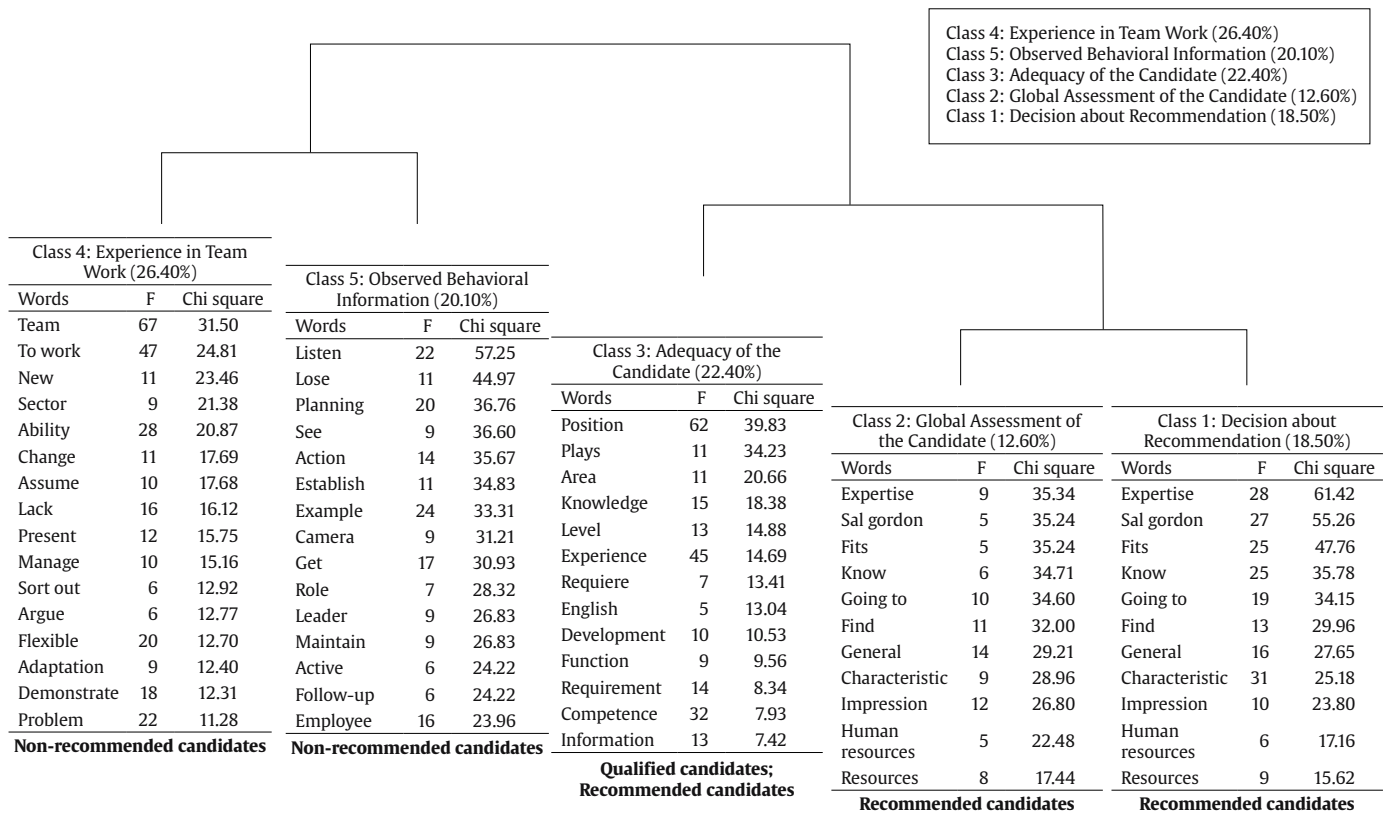
### Qualitative Analyses

The corpus consisted of 25.382 words, with 2.996 unique words. The descending hierarchical cluster analysis divided the corpus into 492 ECUs. The results revealed two main clusters. The first cluster focused on aspects related to the evaluation of a candidate's self-presentation (Classes 4 and 5), while the second cluster was related to the decision-making process regarding a candidate's recommendation to proceed to the second phase (Classes 3, 2, and 1). The first cluster was composed of experience in teamwork (Class 4, 26.4%) and behavioral information observed in the AVI tools (Class 5, 20.1%). Both Classes were associated with the decision not to recommend the candidate for the second phase ( $p < .05$ ). The second cluster contained three Classes: a candidate's suitability for the position (Class 3, 22.4%), the decision on the recommendation (Class 1, 18.6%), and overall assessment of the candidate (Class 2, 12.6%). In this second cluster, a candidate's suitability for the position and the candidate's overall assessment were associated with the decision to recommend the candidate for the second phase ( $p < .04$  and  $p < .03$ , respectively). Furthermore, a candidate's suitability for the position was also associated with the candidate's qualification (Class 3,  $p < .05$ ). (See Figure 1).



**Table 8.** Binary Logistic Regressions to Explain Recommendation of the Candidates

Variables	B	SE	Wald	p	Exp (B)	%95 CI for Exp (B)	
						Lower limit	Upper limit
<b>Qualified woman</b>							
Competence	5.837	2.493	5.481	.019	342.893	2.587	450.998
First impression	2.597	1.204	4.651	.031	13.429	1.267	142.317
Teamwork	0.878	0.393	4.997	.025	2.406	1.114	5.196
Constant	-27.254	9.014	9.143	.002	0.001		
<b>Qualified man</b>							
Non-verbal communication	1.311	0.473	7.681	.006	3.711	1.468	9.381
Planning oriented to conflict resolution	0.712	0.320	4.960	.026	2.039	1.089	3.817
Teamwork	0.700	0.316	4.897	.027	2.013	1.083	3.740
Attractiveness	-1.903	0.651	8.530	.003	0.149	0.042	.535
Constant	-1.108	2.087	0.282	.596	0.330		
<b>Semi-qualified woman</b>							
Competence	1.898	0.515	13.564	.001	6.671	2.430	18.314
Sociability	1.214	0.421	8.327	.004	3.366	1.476	7.676
Flexibility	0.428	0.184	5.438	.020	1.534	1.071	2.199
Constant	-12.202	2.400	25.848	.001	0.001		
<b>Semi-qualified man</b>							
Sociability	1.223	0.395	9.566	.002	3.397	1.565	7.372
Morality	0.953	0.425	5.032	.025	2.593	1.128	5.963
Planning oriented to conflict resolution	0.495	0.226	4.768	.029	1.640	1.052	2.556
Constant	-9.889	1.965	25.324	.001	0.001		



**Figure 1.** Results of the Descending Cluster Analysis Using the Reinert Method. Explanations Given about the Decision Made in the Selection Process of the Candidates.

An example of an ECU from the Classes in the first cluster were the following: “Three years of experience in the sector. He has training and development, is motivated by personal goals but less so by the company’s goals. He argues that teamwork is important to him, but his arguments reveal a selfish attitude towards colleagues. Lacks the ability to share or listen to others’ ideas. He should have

shown more empathy...” (Class 4, candidate not recommended) (chi square = 167.69); “When explaining the problem with the other team members, she displayed a very inflexible attitude and offered little support to the workers who were covering for their colleagues. It does not seem to me that she managed the problem well, as when she was put in charge of leading the team, she did not mention any

activities carried out.” (Class 5, candidate not recommended) (chi square = 377.61).

Examples of ECUs in the second cluster include the following:

“She possesses the right skills for the position, but her work experience does not cover all the areas of activity in which he would be performing his job. Even so, I believe she has sufficient ability to acquire knowledge in these areas and perform the job satisfactorily. I think she has potential, the ability to learn, and the motivation for the position” (Class 3, candidate recommended) (chi square = 156.24); “Overall, I liked him. He effectively defended his profile by emphasizing his professional trajectory and addressing specific situations to highlight his competencies.” (Class 2, candidate recommended) (Chi Square = 202.95); “I would recommend that the candidate proceed to the next phase of the process because a face-to-face interview would be advisable, in which some other nuances that require spontaneous conversation could be appreciated. Further explore his performance” (Class 1, candidate recommended) (chi square = 310.38).

However, there are also examples in this cluster in this line:

“For him, this position is a step backward; he has already worked as an HR manager and says he is in a project similar to ours, from which he is leaving due to burnout. If the manager of our company is not planning to retire and leave the person we hire as their successor, this guy will also get tired and eventually leave. If it is true that his profile is entirely suitable for the operational part, in terms of motivation, he would start off very strong but would eventually get tired and look for another challenge, or he would continue looking for an HR manager position while working with us. He is already a senior profile, and for this position, we would look for something more junior, like the first candidate.” (Class 2, candidate not recommended) (chi square = 190.99).

The results suggest that the justification for not recommending candidates for the second phase is based on candidate deficiencies based on the evaluation of social and organizational competencies. In contrast, the justifications put forth to reason the recommendation of candidates are mostly based both on the fit between a candidate adequacy to the applied position and on a comprehensive evaluation of their profile.

## Discussion

The aim of this study was to identify the most significant competencies and rater biases in hiring recommendations for a HR technician position using AVI tools in a Spanish context, taking into account the gender and qualifications of the candidates.

Contrary to previous research findings (Scott & Roulin, 2024; Torres & Gregory, 2018), our results revealed that competencies were not the primary influence on hiring recommendations, although they did play an important role. In line with this, qualitative analyses showed that justifications for not advancing candidates to the second phase were primarily based on perceived deficiencies in social and organizational competencies. Conversely, recommendations to advance candidates were largely driven by their alignment with position requirements, combined with a comprehensive evaluation of their overall profile. Consistent with the role congruity theory (Eagly & Karau, 2002; Heilman et al., 2024), this study confirms that raters place greater value on competencies associated with gender roles when hiring recommending candidates. Our results highlight that for women it is crucial to appear competent and possess communal competencies (such as teamwork and flexibility) to be recommended by the raters, while for men morality and agentic competency (planning oriented to conflict resolution) are crucial.

In contrast with previous studies (Menegatti et al., 2021; Moscatelli et al., 2020; Pireddu et al., 2022; Quadlin, 2018), for women appearing competent is more important than it is for men

in order to be recommended. There is a clear distinction between highly qualified women and semi-qualified women concerning the additional variables associated with competence. For highly qualified women, the first impression is particularly important, while for semi-qualified women, being sociable is the key. Literature in this area has demonstrated that first impression does influence overall rating scores (Swider et al., 2016; Torres & Gregory, 2018; Tu et al., 2022), although to a lesser extent in AVI tools (Suen et al., 2019; Torres & Gregory, 2018). Our results provide evidence that the importance of a candidate's first impression in AVI tools varies by gender. It is crucial only for highly qualified women, highlighting the significance of appearing competent, not just being qualified for the position.

In the case of highly qualified men, our results suggest that appearing competent, having effective non-verbal communication, and being attractive may lead to an overvaluation of the candidate by raters. According to previous research (Martín-Raugh, et al., 2023), we found that non-verbal communication explains qualified men' recommendation. Non-verbal presence is associated with eye contact, smiling, and expansive body posture (Tu et al., 2022), but also with professional appearance (Martín-Raugh, et al., 2023). Therefore, highly qualified men whose non-verbal communication is effective are more likely to be recommended to the extent that they are seen as competent.

Our results indicate that while highly qualified men are perceived as attractive, they are less likely to be recommended for hiring. Some authors (Tu et al., 2022) have pointed out attractiveness and non-verbal communication are closely linked, since attractive candidates display more effective non-verbal cues and higher scores than less attractive ones. Highly qualified men may be viewed as overqualified for the position and inspire fear in the raters that they may soon leave the job. Birkelund et al. (2022) suggested that women candidates are perceived as more stable staff than men candidates. Turnover intention is one of the most alarming factors for the organizations since it involves discomfort and lower performance at work. Therefore, one of the basic principles of personnel selection processes is to avoid hiring a candidate who is overqualified for the position. The raters in this study confirmed that, alleging in their recommendation justifications doubts about recommending qualified male, since they could leave the job and look for others of higher rank. In the evaluation of qualified men, the halo effect (Thorndike, 1920) may be at play, where additional characteristics are attributed to a person based on a few observed traits. In this case, it would result in a more positive overall judgment based on characteristics such as competence or attractiveness. This could explain why participants in the study report that qualified men are perceived as overqualified for the position. The halo effect is a judgment error, as evidenced by the fact that the profiles of qualified men and women are similar; however, over qualification is only attributed to men. In this line, a study, comprising four experimental investigations into managers' hiring decisions, revealed that signals of high capability are not necessarily perceived as advantageous in the selection process. This is because highly capable candidates are seen as having more opportunities to leave the job and are perceived as less committed to the organization (Galperin, et al., 2020). Consequently, being perceived as highly competent may work against men in hiring decisions.

Consistent with previous studies (e.g., Corrington & Hebl, 2018; Rudman & Phelan, 2008), our data show that sociability is a highly valued characteristic in interviews, but only for semi-qualified candidates. It is important to note that there is a gender-based difference: semi-qualified women must also appear competent to be recommended for hiring, whereas for semi-qualified men, demonstrating moral qualities is more important. Indeed, young men may be perceived as less sociable if they are perceived as competent and are penalized in the recruitment process, while this does not happen in the young women case (Krings et al., 2023).

## Implications for Practice

The checking manipulations conducted to verify that the use of AVI tools correctly identified qualified and semi-qualified candidates suggest that this tool is effective in recognizing competent candidates. However, despite the fact that the sample in this study consisted of experienced professionals, biases continue to emerge, even with their training and expertise. This leads us to two conclusions regarding the practical implications: first, the use of AVI tools is not suitable for all stages of the selection process, and second, it is essential to continue addressing this issue in the organizational context at different levels.

Regarding the timing of using the AVI tool, it is considered suitable for the early stages as a preliminary screening mechanism for evaluating competencies. However, given that these tools are not free of biases, it is crucial that they be complemented with in-depth job interviews in the later stages of the selection process to ensure a more thorough and accurate assessment of the candidates.

In terms of actions that can be implemented to prevent the emergence of stereotypes, at individual level, this study demonstrates a link between the activation of gender stereotypes and candidate recommendations. Understanding this association could benefit raters by enabling them to improve their selection processes and reduce bias in hiring decisions based on stereotypical thinking. Therefore, training raters in the use of AVI tools and in recognizing the biases that may emerge during the selection process is essential for promoting fairness. Providing objective information about different groups and highlighting the emergence of bias are effective training strategies to reduce discrimination and promote equal opportunities. Recent studies on ethnic discrimination (Derous et al., 2020) have confirmed that prejudice can be reduced through targeted training. However, the same study emphasized the need for tools to sustain the long-term impact of such training. While training may initially reduce hiring discrimination, the effects tend to diminish over time, allowing discrimination to resurface. Future research will focus on developing new training methods that ensure equal access to employment opportunities in a more durable way.

Likewise, the training method and guide could be used as part of the formal university curriculum for training future professionals. The 2030 Agenda emphasizes the need to build more inclusive societies (Goal 11) and promote equality (Goals 5 and 10) for all. This can only be achieved through quality education (Goal 4) that addresses these issues and facilitates the creation of decent jobs (Goal 8) accessible to everyone, without any form of discrimination.

At organizational level, it would be interesting to develop a guide to ensure non-discriminatory e-selection processes that allow consultants and companies to have guidelines for action to make effective selection processes and ensuring equal opportunities.

## Limitations and Future Research

Although this study offers insight into AVI tools functioning in the preselection process, it is not without limitations. First, the small sample does not allow the results to be generalized nor does it allow for more complex analyses that allow establishing models about decision-making in personnel selection. We call for researchers to carry out structural equation models that allow us to delve deeper into the results found in this study.

Another limitation of this study pertains to its design, as it exclusively focused on the selection of younger profiles. Consequently, additional research is needed to explore the outcomes for more mature or senior profiles. Future studies should aim to replicate this research while considering other forms of discrimination that may arise based on candidate profile, such as ethnicity, sexual orientation, or disability. Additionally, it would be valuable to investigate other types of jobs, including those related to senior officials, to provide a

more comprehensive understanding of the selection process. It would also be interesting for future research to analyze whether the biases identified in this study persist when using artificial intelligence to evaluate candidates through AVI tools.

Likewise, we call for researchers to investigate the biases that may influence the rating of candidates. Our study suggests that while raters focus on the competence scores assigned to candidates when making recommendations, these scores can also be affected by rater biases. Understanding these biases is crucial for ensuring a fair and equitable selection process.

## Conclusion

AVI tools have been recognized as valuable instruments that help raters focus on assessing competencies while minimizing biases (Kroll & Ziegler, 2016; Scott & Roulin, 2024; Suen et al., 2019; Torres & Gregory, 2018). However, contrary to findings from studies in other fields and geographical contexts, our results indicate that while raters are able to differentiate between qualified and semi-qualified candidates, their decisions remain influenced by biases. This highlights the importance of using AVIs as preliminary and complementary tools, alongside other methods such as in-depth interviews, while also emphasizing the need to raise awareness and provide training for evaluators on the potential biases that may arise during the selection process.

## Conflict of Interest

The authors of this article declare no conflict of interest.

## References

- Abele, A. E., Hauke, N., Peters, K., Louvet, E., Szymkow, A., & Duan, Y. (2016). Facets of the fundamental content dimensions: Agency with competence and assertiveness—Communion with warmth and morality. *Frontiers in Psychology, 7*, Article 1810. <https://doi.org/10.3389/fpsyg.2016.01810>
- Alonso, P., & Moscoso, S. (2017). Structured behavioral and conventional interviews: Differences and biases in rater ratings. *Revista de Psicología del Trabajo y de las Organizaciones, 33*(3), 183-192. <https://doi.org/10.1016/j.rpto.2017.07.003>
- Alonso, P., Moscoso, S., & Salgado, J. F. (2017). Structured behavioral interview as a legal guarantee for ensuring equal employment opportunities for women: A meta-analysis. *The European Journal of Psychology Applied to Legal Context, 9*(1), 15-23. <https://doi.org/10.1016/j.ejpal.2016.03.002>
- Basch, J. M., Melchers, K. G., & Büttner, J. C. (2022). Preselection in the digital age: A comparison of perceptions of asynchronous video interviews with online tests and online application documents in a simulation context. *International Journal of Selection and Assessment, 30*(4), 639-652. <https://doi.org/10.1111/ijsa.12403>
- Birkelund, G. E., Lancee, B., Larsen, E. N., Polavieja, J., Radl, J., & Yemane, R. (2022). Gender discrimination in hiring: Evidence from a cross-national harmonized field experiment. *European Sociological Review, 38*(3), 337-354. <https://doi.org/10.1093/esr/jcab043>
- Boor, M., Wartman, S. A., & Reuben, D. B. (1983). Relationship of physical appearance and professional demeanor to interview evaluations and rankings of medical residency applicants. *The Journal of Psychology, 113*(1), 61-65. <https://doi.org/10.1080/00223980.1983.9923557>
- Buijsrogge, A., Duyck, W., & Derous, E. (2020). Initial impression formation during the job interview: Anchors that drive biased decision-making against stigmatized applicants. *European Journal of Work and Organizational Psychology, 30*(2), 305-318. <https://doi.org/10.1080/1359432X.2020.1833980>
- Correll, S. J. (2017). Reducing gender biases in modern workplaces: A small wins approach to organizational change. *Gender and Society, 31*(6), 725-750. <https://doi.org/10.1177/0891243217738518>
- Corrington, A., & Hebl, M. (2018). America clearly is not ready for a female president: Why? *Equality, Diversity and Inclusion, 37*(1), 31-43. <https://doi.org/10.1108/EDI-08-2017-0175>
- Cuadrado, I., López-Rodríguez, L., Estevan-Reina, L., Constantín, A. A., & Robles, A. (2024). Effects of attractiveness and occupation type on attitudes toward working women and men. *Basic and Applied Social Psychology, 46*(2), 83-96. <https://doi.org/10.1080/01973533.2024.2302464>

- Deros, E., Buijsrogge, A., Roulin, N., & Duyck, W. (2016). Why your stigma isn't hired: A dual-process framework of interview bias. *Human Resource Management Review*, 26(2), 90-111. <https://doi.org/10.1016/j.hrmr.2015.09.006>
- Deros, E., Nguyen, H. H. D., & Ryan, A. M. (2020). Reducing ethnic discrimination in resume-screening: A test of two training interventions. *European Journal of Work and Organizational Psychology*, 30(2), 225-239. <https://doi.org/10.1080/1359432X.2020.1756907>
- Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, 24(3), 285-290. <https://doi.org/10.1037/h0033731>
- Duehr, E. E., & Bono, J. E. (2006). Men, women, and managers: Are stereotypes finally changing? *Personnel Psychology*, 59(4), 815-846. <https://doi.org/10.1111/j.1744-6570.2006.00055.x>
- Dunlop, P. D., Holtrop, D., & Wee, S. (2022). How asynchronous video interviews are used in practice: A study of an Australian-based AVI vendor. *International Journal of Selection and Assessment*, 30(3), 448-455. <https://doi.org/10.1111/ijssa.12372>
- Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological Review*, 109(3), 573-598. <https://doi.org/10.1037/0033-295x.109.3.573>
- Fiske, S. T. (2018). Stereotype content: Warmth and competence endure. *Current Directions in Psychological Science*, 27(2), 67-73. <https://doi.org/10.1177/0963721417738825>
- Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology*, 82(6), 878-902. <https://doi.org/10.1037/0022-3514.82.6.878>
- Fraundorfer, D., & Mast, M.S. (2015). The impact of nonverbal behavior in the job interview. In A. Kostić & D. Chadee (Eds.), *The social psychology of nonverbal communication*. Palgrave Macmillan. [https://doi.org/10.1057/9781137345868\\_11](https://doi.org/10.1057/9781137345868_11)
- Galperin, R. V., Hahl, O., Sterling, A. D., & Guo, J. (2020). Too good to hire? Capability and inferences about commitment in labor markets. *Administrative Science Quarterly*, 65(2), 275-313. <https://doi.org/10.1177/0001839219840022>
- González, M. J., Cortina, C., & Rodriguez, J. (2019). The role of gender stereotypes in hiring: A field experiment. *European Sociological Review*, 35(2), 187-204. <https://doi.org/10.1093/esr/jcy055>
- Gorman, C. A., Robinson, J., & Gamble, J. S. (2018). An investigation into the validity of asynchronous web-based video employment-interview ratings. *Consulting Psychology Journal: Practice and Research*, 70(2), 129-146. <https://doi.org/10.1037/cpb0000102>
- Heilman, M. E., Caleo, S., & Manzi, F. (2024). Women at work: Pathways from gender stereotypes to gender bias and discrimination. *Annual Review of Organizational Psychology and Organizational Behavior*, 11(1), 165-192. <https://doi.org/10.1146/annurev-orgpsych-110721-034105>
- Hentschel, T., Heilman, M. E., & Peus, C. V. (2019). The multiple dimensions of gender stereotypes: A current look at men's and women's characterizations of others and themselves. *Frontiers in Psychology*, 10, Article 11. <https://doi.org/10.3389/fpsyg.2019.00011>
- Hofmann, B. (2024). Aesthetic Injustice. *Journal of Business Ethics*, 189, 217-229. <https://doi.org/10.1007/s10551-023-05401-4>
- Hosoda, M., Stone-Romero, E. F., & Coats, G. (2003). The effects of physical attractiveness on job-related outcomes: A meta-analysis of experimental studies. *Personnel Psychology*, 56(2), 431-462. <https://doi.org/10.1111/j.1744-6570.2003.tb00157.x>
- Idoiaga Mondragon, N., Yarritu, I., Saez de Cámara, E., Beloki, N., & Vozmediado, L. (2023). The challenge of education for sustainability in higher education: Key themes and competences within the University of the Basque Country. *Frontiers in Psychology*, 14, Article 1158636. <https://doi.org/10.3389/fpsyg.2023.1158636>
- Johnson, S. K., Podratz, K. E., Dipboye, R. E., & Gibbons, E. (2010). Physical attractiveness biases in ratings of employment suitability: Tracking down the "Beauty is beastly" effect. *The Journal of Social Psychology*, 150(3), 301-318. <https://doi.org/10.1080/00224540903365414>
- Klein, O., & Licata, L. (2003). When group representations serve social change: The speeches of Patrice Lumumba during the Congolese decolonization. *British Journal of Social Psychology*, 42(4), 571-593. <https://doi.org/10.1348/01446660322595284>
- Koutsoumpis, A., Ghassemi, S., Oostrom, J. K., Holtrop, D., van Breda, W., Zhang, T., & de Vries, R. E. (2024). Beyond traditional interviews: Psychometric analysis of asynchronous video interviews for personality and interview performance evaluation using machine learning. *Computers in Human Behavior*, 154, 1-18. <https://doi.org/10.1016/j.chb.2023.108128>
- Krings, F., Manoharan, S., & Mendes de Oliveira, A. (2023). Backlash or high self-promotion at hiring depends on candidates' gender and age. *International Journal of Selection and Assessment*, 31(3), 362-375.
- Kroll, E., & Ziegler, M. (2016). Discrimination due to Ethnicity and Gender: How susceptible are video-based job interviews? *International Journal of Selection and Assessment*, 24(2), 161-171. <https://doi.org/10.1111/ijssa.12138>
- 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/ijssa.12191>
- Leach, C. W., Carraro, L., Garcia, R. L., & Kang, J. J. (2017). Morality stereotyping as a basis of women's in-group favoritism: An implicit approach. *Group Processes & Intergroup Relations*, 20(2), 153-172. <https://doi.org/10.1177/1368430215603462>
- Leach, C. W., Ellemers, N., & Barreto, M. (2007). Group virtue: The importance of morality (vs. competence and sociability) in the positive evaluation of in-groups. *Journal of Personality and Social Psychology*, 93(2), 234-249. <https://doi.org/10.1037/0022-3514.93.2.234>
- Liff, J., Mondragon, N., Gardner, C., Hartwell, C. J., & Bradshaw, A. (2024). Psychometric properties of automated video interview competency assessments. *Journal of Applied Psychology*, 109(6), 921-948. <https://doi.org/10.1037/apl0001173>
- Lukacik, E. R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, 32(1), Article 100789. <https://doi.org/10.1016/j.hrmr.2020.100789>
- Mao, Y., Hu, W., Xu, C., & Sun, Y. (2024). Beauty and flexible employment in the digital age: The mediating role of social capital. *Journal of the Knowledge Economy*. <https://doi.org/10.1007/s13132-024-01824-0>
- Martin-Raugh, M. P., Kell, H. J., Randall, J. G., Anguiano-Carrasco, C., & Banfi, J. T. (2023). Speaking without words: A meta-analysis of over 70 years of research on the power of nonverbal cues in job interviews. *Journal of Organizational Behavior*, 44(1), 132-156. <https://doi.org/10.1002/job.2670>
- Menegatti, M., Pireddu, S., Crocetti, E., Moscatelli, S., & Rubini, M. (2021). The Geneva de'Benci effect: Competence, morality, and attractiveness inferred from faces predict hiring decisions for women. *Frontiers in Psychology*, 12, Article 658424. <https://doi.org/10.3389/fpsyg.2021.658424>
- Moscatelli, S., Menegatti, M., Ellemers, N., Mariani, M. G., & Rubini, M. (2020). Men should be competent, women should have it all: Multiple criteria in the evaluation of female job candidates. *Sex Roles*, 83(5-6), 269-288. <https://doi.org/10.1007/s11199-019-01111-2>
- Nault, K. A., Pitesa, M., & Thau, S. (2020). The attractiveness advantage at work: A cross-disciplinary integrative review. *Academy of Management Annals*, 14(2), 1103-1139. <https://doi.org/10.5465/annals.2018.0134>
- Niemitz, N., Rietheimer, L., König, C., & Langer, M. (2024). The (null) effects of video questions on applicant reactions in asynchronous video interviews: Evidence from an actual hiring context. *International Journal of Selection and Assessment*, 32(3), 421-428. <https://doi.org/10.1111/ijssa.12470>
- Niu, Y. (2024). The impact of lookism on employment decisions: evidence from China. *Equality, Diversity and Inclusion*, 43(2), 188-210. <https://doi.org/10.1108/EDI-12-2022-0359>
- Pereda Marín, S., Berrocal Berrocal, F., & Sanz Gómez, P. (2003). Los perfiles de exigencias en la ocupación del profesional de recursos humanos. *Psicología desde el Caribe*, 12, 13-38.
- Pireddu, S., Bongiorno, R., Ryan, M. K., Rubini, M., & Menegatti, M. (2022). The deficit bias: Candidate gender differences in the relative importance of facial stereotypic qualities to leadership hiring. *British Journal of Social Psychology*, 61(2), 644-671. <https://doi.org/10.1111/bjso.12501>
- Quadlin, N. (2018). The mark of a woman's record: Gender and academic performance in hiring. *American Sociological Review*, 83(2), 331-360. <https://doi.org/10.1177/0003122418762291>
- Reinert, M. (1986). Un logiciel d'analyse lexicale: ALCESTE. *Les cahiers de l'Analyse des Données*, 11(4), 471-481.
- Roso-Bas, F., Pades-Jiménez, A., & Ferrer-Pérez, V. A. (2017). Competencia comunicativa: validación de una escala para evaluar la comunicación no verbal durante el discurso. *Estudios sobre Educación*, 32, 95-113. <https://doi.org/10.15581/004.32.95-113>
- Roulin, N., Lukacik, E. R., Bourdage, J. S., Clow, L., Bakour, H., & Diaz, P. (2023). Bias in the background? The role of background information in asynchronous video interviews. *Journal of Organizational Behavior*, 44(3), 458-475. <https://doi.org/10.1002/job.2680>
- Roulin, N., Pham, L. K. A., & Bourdage, J. S. (2023). Ready? Camera rolling... action! Examining interviewee training and practice opportunities in asynchronous video interviews. *Journal of Vocational Behavior*, 145, Article 103912. <https://doi.org/10.1016/j.jvb.2023.103912>
- Roulin, N., Wong, O., Langer, M., & Bourdage, J. S. (2023). Is more always better? How preparation time and re-recording opportunities impact fairness, anxiety, impression management, and performance in asynchronous video interviews. *European Journal of Work and Organizational Psychology*, 32(3), 333-345. <https://doi.org/10.1080/1359432X.2022.2156862>
- Rudman, L. A., & Phelan, J. E. (2008). Backlash effects for disconfirming gender stereotypes in organizations. *Research in Organizational Behavior*, 28, 61-79. <https://doi.org/10.1016/j.riob.2008.04.003>
- Scott, C., & Roulin, N. (2024). Does background type and blurring affect performance ratings in video interviews? *Personnel Assessment and Decisions*, 10(1), 12-20. <https://doi.org/10.25035/pad.2024.01.002>
- Sheppard, L. D., & Johnson, S. K. (2019). The femme fatale effect: attractiveness is a liability for business women's perceived truthfulness, trust, and deservingness of termination. *Sex Roles*, 81(11-12), 779-796. <https://doi.org/10.1007/s11199-019-01031-1>
- Spencer, L. M., & Spencer, P. S. M. (2008). *Competence at work models for superior performance*. John Wiley & Sons.

- Suen, H. Y., Chen, M. Y. C., & Lu, S. H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior, 98*, 93-101. <https://doi.org/10.1016/j.chb.2019.04.012>
- Swider, B. W., Barrick, M. R., & Harris, T. B. (2016). Initial impressions: What they are, what they are not, and how they influence structured interview outcomes. *Journal of Applied Psychology, 101*(5), 625-638. <http://dx.doi.org/10.1037/apl0000077>
- Swider, B. W., Harris, T. B., & Gong, Q. (2022). First impression effects in organizational psychology. *Journal of Applied Psychology, 107*(3), 346-369. <https://doi.org/10.1037/apl0000921>
- Thorndike, E. L. (1920). A constant error in psychological ratings. *Journal of Applied Psychology, 4*(1), 25-29. <https://doi.org/10.1037/h0071663>
- Torres, E. N., & Gregory, A. (2018). Hiring manager's evaluations of asynchronous video interviews: The role of candidate competencies, aesthetics, and resume placement. *International Journal of Hospitality Management, 75*, 86-93. <https://doi.org/10.1016/j.ijhm.2018.03.011>
- Tu, M. H., Gilbert, E. K., & Bono, J. E. (2022). Is beauty more than skin deep? Attractiveness, power, and nonverbal presence in evaluations of hirability. *Personnel Psychology, 75*(1), 119-146. <https://doi.org/10.1111/peps.12469>
- Turkmenoglu, M. A. (2020). Exploring appearance-based discrimination in the workplace. In M. A. Turkmenoglu & B. Cicek (Eds.), *Contemporary global issues in human resource management*, (pp. 23-35). Emerald Publishing Limited. <https://doi.org/10.1108/978-1-80043-392-220201004>
- Watkins, L. M., & Johnston, L. (2000). Screening job applicants: The impact of physical attractiveness and application quality. *International Journal of Selection and Assessment, 8*(2), 76-84. <https://doi.org/10.1111/1468-2389.00135>
- 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>

