UNIVERSITY OF PADOVA

Department of General Psychology

Bachelor's Degree Course in Psychological Science

Final dissertation

Technology-mediated distortions: a review on the biases and misperceptions in employment interviews via computer, telephone and AI

Supervisor:

Prof. Anna Spagnolli

Candidate: Enrico Guarino

Student ID number: 1197677

Academic Year 2021/22

Table of Contents

1. Abstract

Biases in job interviews threaten the objective evaluation of applicants. Similar and different biases exist also in mediated job interviews, where the communication between applicant and interviewer passes through a technological software or hardware. This review synthetises the literature investigating biases in job interviews conducted through telephone, videoconference, asynchronous videos or avatars. Moreover, this review reports perceptions applicants and interviewers had of such modalities. Overall, applicants received lower ratings in mediated interviews compared to face-to-face ones. In fact, lack of nonverbal cues, bad audio/video quality, lags and non-neutral interview locations hinder interviewers in performing objective assessments of applicants. Moreover, the appearance of avatars is another source of bias, as the characteristics of avatars merge with or override those of applicants. Regarding perceptions, interviewers and applicants expressed mainly negative perceptions. In particular, applicants were particularly concerned about privacy and fairness, with the latter being lower for mediated interviews. Furthermore, avatars accentuate biases of face-to-face interviews and can appear “creepy” to applicants. Finally, technological mediation presents other downsides, i.e., increased difficulty in the interviewer-applicant interaction and a rigid and impersonal process. Despite these biases, negative perceptions and downsides, technological mediation brings about simpler and more accessible interviews for applicants and recruiters, along with the chance of a greater level of interview standardisation. To solve the issues of mediated interviews, researchers suggest to do less interviews, pair them with other forms of assessment, standardise interviews more, better inform applicants and make avatars able to transmit more characteristics of their operators, such as nonverbal cues.

2. Introduction

There is no living without working. In fact, finding and securing a job is what allows individuals to provide for themselves and their families, as often work is the only form of income until retirement. Therefore, a fair and objective employment process is of paramount importance in the contemporary world, especially when considering that many individuals change several jobs in a lifetime. However, unpleasant consequences arise if recruiters choose future employees not based on their merits or objective characteristics, but on the recruiters'[1] biases and subjective perceptions. Even more unpleasant is the fact that such penalising distortions may happen even if the recruiter tries to behave in an objective manner. That is, the fault does not lie on the person evaluating the applicant[2], but on the technologies used in the recruitment process, such as those used in job interviews. The adoption of technology in such processes and contexts is not something unheard of, as organisations have been using them for decades, with telephones being one of the first technological media employed (Brenner, Ortner and Fay, 2016, in Langer, König and Fitili, 2018).

In the last years, Artificial Intelligence (AI) started to see implementation in the context of personnel selection as well and became, consequently, the centre of attention for academic research (see Langer et al., 2018; Langer and Landers, 2021; Köchling, Riazy, Wehner and Simbeck, 2021; Wesche and Sonderegger, 2021). Despite the algorithms devised to aid recruiting promised increased reliability and efficiency, many among managers, potential applicants and the public believe AI should not substitute humans in certain tasks, such as job interviews

---

[1] In this review, the terms *rater, interviewer, evaluator* and *recruiter* will be used as synonyms and all will refer to the person working for the organisation that is setting the interview, as well as to who will evaluate the applicant. Interviewers can be professional figures, who rate applicant as part of their job, or participants of a study who are required to think and/or behave as such.

[2] The terms *applicant* and *interviewee* will be used as synonyms. These terms refer to those individuals who accept to participate in a job interview with the hope of becoming an employee of the organisation setting that interview. Applicants may consist of individuals who are looking for a job in real life or participants asked to think and/or behave as such.

(Langer and Landers, 2021). Similarly, Wesche and Sonderegger (2021) found that participants[3] were deterred by the prospect of undergoing an automated AI interview, i.e., one where the AI took the final decision about keeping or discarding the applicant, without the intervention of a human. Subjective perceptions, however, are not enough to claim that AI-enhanced interviews (being them automated or augmented, i.e., aiding humans in judgment) bear problems, such as biases[4] towards applicants. Unfortunately, numerous biases and distortions affecting individuals plague AI, as recent literature is reporting (Lukacik, Bourdage and Roulin, *in press*; Allam, 2021; Köchling et al., 2021; Langer and Landers, 2021; Wesche and Sonderegger, 2021; Srinivasan and Chander, 2021). For instance, AI can be affected by *data-creation bias*, which is an umbrella term that refers to all those biases that occur during the creation of a dataset (i.e., the information that the algorithm uses to learn). An example of this is *sampling bias*, where some instances (e.g., males) are more represented than others (e.g., females). Data-creation bias also includes *negative set bias*, which results from a lack of negative instances of a given feature in the dataset (e.g., a black swan in a set of white swans) and *label bias*, which is related to the inconsistencies in the labelling of data and individual differences/biases inherent in human labellers (Srinivasan and Chander, 2021). Labelling often is a necessary action, as certain AI cannot learn without aid, so designers cannot always avoid this last bias by skipping the labelling procedure altogether (Zorzi, Testolin and Stoianov, 2013). Another category of biases affecting AI is related to the formulation of the problem that the AI must resolve. For instance, a *confounding bias* happens when the AI fails to learn the correct relations between the characteristics of the elements of a dataset, either by not considering all the info in the data or by missing relations

---

[3] The term *participant* refers to the individuals taking part in the studies reported by the examined articles, but in a more general term. As a result, *participant* can be used to refer to those who acted as applicants or as interviewers or, instead, to individuals who did not fit these roles (e.g., they were external judges).

[4] When directed towards individuals, a *bias* is a tendency or predisposition for or against something, sometimes akin to prejudice. If concerning some kind of methodology or process, bias instead refers to a systematic error occurring in sampling and collecting/analysing data. (APA Dictionary, 2022a).

between those characteristics and target outputs (i.e., responses or decisions; Srinivasan and Chander, 2021).

The presence of these and many other biases (see Srinivasan and Chander, 2021, and Köchling et al., 2021, for a more detailed account) clearly make AI still unsuitable for a complete implementation in the context of personnel recruitment. However, this is not the end of the story. After acknowledging the problems AI face, three questions follow naturally: What about other kind of technologies used in job interviews? Which are these technologies? Are they affected by, or at least related to, biases as well? Actually, organisations[5] employ mediums[6] such as telephone, videoconference equipment and video recordings to screen their applicants and employers have been doing that for several years (Brenner et al., 2016, in Langer et al., 2018). Mediated interviews have evolved in form and scope as the technological tools used to perform them evolved over time. More specifically, mediation came first with telephone interviews, where the recruiter and the applicant exchange questions and answers via telephone, restricting, in this way, the information received by both to only the voice of their conversational partner. Subsequently, organisations started to employ videoconference interviews, where applicant and recruiter can both hear and see each other by using camera technologies. Finally, the last major change introduced regards Asynchronous Video Interviews (also referred as "digital interviews"), which present a temporal-displacement factor, i.e., interviewees record themselves while answering the questions provided by the organisation and only after that can interviewers rate them. Actually, recruiters and hiring managers can potentially watch these recordings at any point in time after the interview has been stored. (Brenner et al., 2016 in Langer et al., 2018; Torres, and Mejia, 2017). Consequently, given that the employment of AI in the recruitment process is still

---

[5] The terms *organisation* and *firm* will be used to refer to those enterprises that set-up interviews to hire new employees and for which interviewers work.

[6] Not all technological tools are mediums by default. Essentially, it is possible to define a *medium* as "any means or agency through which messages are transmitted or information is diffused" (APA dictionary, 2022b).

moving its first steps and that literature on the topic is still scarce, this review will focus on the biases and misperceptions associated to the use of telephone, videoconference and asynchronous video interviews in conducting job interview. Moreover, this review will also consider an additional medium that researchers have recently been exploring in their studies, i.e., digital and physical avatars. Therefore, in the next sections, these interview modalities will be thoroughly presented along with the reported biases and misperceptions associated to their use. For convenience, acronyms will substitute the full names of the following modalities: telephone (*TEL*), videoconference (*VC*), asynchronous video interviews (*AVI*) and face-to-face (*FTF*). Despite not entailing technological mediation, this last interview modality will appear in the following sections as researchers used it to make comparisons with the other, mediated modalities.

The next sub-section presents a short but detailed overview on the different interview modalities that originate from the adoption of specific technological mediums (i.e., TEL, VC, AVI and avatars).

2.1 Overview of the interview tools used

The focus of this inquiry, as mentioned before, are technology-mediated interviews. In general, we can define an interview as a synchronous interaction between two parties in which evaluation is central (Blacksmith, Willford, and Behrend, 2016). More specifically, a digital interview (i.e. one employing technology and that is not FTF) is "any procedure that makes use of digital communication technology (i.e. computer-, internet- or mobile-based) for the purposes of assisting organizations during recruitment and selection" (Woods, Ahmed, Nikolaou, Costa, and Anderson, 2019, p. 65, in Wesche and Sonderegger, 2021). Such interviews have the advantages of saving the time and resources used to screen applicants while, at the same time, allowing organisations to access a larger applicant pool and handle these applicants in a more efficient way. Essentially, organisations achieve these benefits by substituting FTF interviews and analogical application submission with web-based applications and on-line interviews (Wesche and Sonderegger, 2021).

As mentioned above, the last major change in mediated interviews (aside from AI use) was the introduction of AVIs. Given their displaced nature, Toldi (2011) highlights how, during an AVI, "The candidate is only able to give information due to the asynchronous nature of a non-live video interview. They are unable to ask the interviewer questions and receive other information from the employer." (p. 20). In contrast to FTF and VC interviews, AVIs promise to offer more flexibility (no need for scheduling), standardization (no influence of the interviewer on the interviewee) and analytical possibilities (automatic, algorithmic evaluation of the interviews) (Langer, König, and Fitili, 2018). Examples of third-party software companies that allow for AVIs include HireVue, InterviewStream, Vidrecruiter and Interview Rocket (Torres and Gregory, 2018). In their article, Lukacik, Bourdage and Roulin (*in press*) did a remarkable job in conceptualising a model of AVI design and the choices applicants can make before and while completing such interview. The most important aspect highlighted by this model is that applicants and organisations' decisions before the interview can influence an applicant during an AVI, along with the AVI outcomes. For instance, organizations' pre-interview decisions regarding design features (e.g., letting applicants to

re-record answers) could then influence applicants' behaviours in the interview (e.g., use of impression management, interview anxiety), which can ultimately affect post-interview outcomes (e.g., interview performance). More specifically, it is possible to distinguish the characteristics of AVIs that can influence applicants in *Structure and formatting*, *Media features*, *Response formatting features* and *Evaluations features* (Lukacik et al., *in press*). The first characteristic listed above is related to how long questions stays on the screen of the applicant and what prompts such questions to disappear (e.g., a pre-set timer, the applicant clicking a button). *Media features* refer to the presence or absence of video, music or photos introducing the procedure; the presentation of questions to the applicant, which can be written or consist in a recording of the interviewer; and the audio/video quality of the AVI. *Response formatting features* entail all those aspects of an AVI that bound the responses of applicants. In fact, AVIs may or may not allow applicants some extra time to prepare their answers, the chance of stopping the interview and resume it later, or an unlimited time to complete the interview. Moreover, organisations may decide to provide opportunities for self-feedback by allowing applicants to review their responses before submitting them (e.g., re-watching and/or re-recording answers, having a picture-in-picture preview while speaking). Finally, *Evaluation features* refer to the fact that applicants may receive an evaluation score from a human or an AI. As for organisations, applicants' pre-interview decisions can affect interview outcomes as well, even those relevant for organisations (such as evaluator bias and adverse impact; Lukacik et al., *in press*). Part of these decisions affect the aesthetics of the environment or of the applicant, as an individual can choose the location of the interview, the visible background, the lightning of videos and their physical appearance. Other decisions, instead, affect the technological side of the interview, as applicants can act to change the speed of their internet connection, along with the clarity and stability of their image on the screen.

Avatar-mediated interviews aim, among other things, at increasing the fairness of interviews and the *symmetricity* between applicants and interviewers (Nørskov, Damholdt, Ulhøi, Jensen, Ess and Seibt, 2020). Improving on the Fair Proxy Communication (FPC) framework, these researchers argued that providing

anonymity through a proxy only for the interviewer is still unfair in the end. More specifically, a job interview in which the applicant cannot see the interviewer, but the latter can see the former, will result in the applicant's inability to form perceptual biases towards the interviewer, which is a positive feature. However, given that the interviewer can still see the applicant and their non-verbal behaviour, the interviewer can nevertheless form perceptual biases towards the applicant and in general holds more "power" in his/her hands (hence the *asymmetricity*; Nørskov et al., 2020). Consequently, these researchers claimed that both individuals involved in an interview should be allowed to use a proxy avatar, as this re-establishes symmetricity and, most importantly, allows applicants to successfully engage in *impression management*[7]. Impression management strategies need to be finely attuned to the conversational partner's behaviour to work. Therefore, only interviewers can potentially perform such strategies in the asymmetric situation described above, as applicants are not allowed to see the interviewers. (Nørskov et al., 2020). **Figure 1** presents a table describing the different combinations of applicant and interviewer's visual anonymity.

[7] Impression management consists in deliberate behavioural strategies that individuals use to influence the perceptions others have of them, especially by making others attribute positive traits to them (APA Dictionary, 2022c).

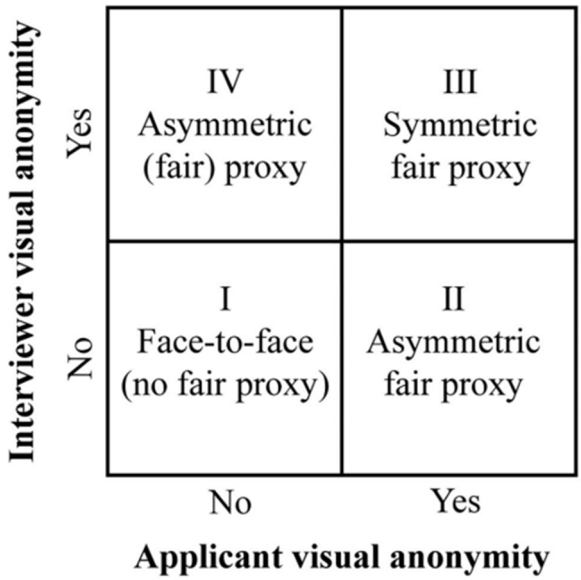

Figure 1. The combination of applicant's and interviewer's visual anonymity leads to four different types of job interview setups, with or without a "fair" proxy (Nørskov et al., 2020).

These last bits of information end the introductory part of this review. The next section, Section 3, will explain the methodology that has been adopted to conduct the research in the online databases and that led to the articles used as references. Such explanation will entail both the criteria used to select articles and those used to discard them, with the criteria for exclusion receiving a particular focus, as they have been applied in a Popperian-like fashion. Section 4 will illustrate, instead, the insights and evidence emerging from the examined articles and it will do so by presenting six research questions, which functioned as a guide to conduct the research in the literature. These questions are:

1) *Is the presence of bias in mediated job interviews supported?*
2) *Are there statistically significant distortions arising from the use of technology?*
3) *What are the psychological and/or physical phenomena reported as exclusive to mediated job interviews?*
4) *Are mediated job interviews suggested as preferable in the study under consideration? If yes/no, why?*
5) *What are the reported pros and cons of mediated job interviews?*

6) *Do authors propose solutions to solve the issues of mediated job interviews?*

In addition, Section 5 will explain what the main limitations of this review are. Section 6 will provide a brief conclusion on the topics discussed, along with some reflections on how future research could address them. Section 7 presents the bibliographical information of the sources used to write this review. Finally, Section 8 contains the Appendix, in which the reader can find additional material related to the topics discussed.

3. Research methodology

Articles on biases and misperceptions in technology-mediated job interviews were searched primarily in the Scopus database, with the subsequent aid of Google Scholar. The first step of the research procedure consisted in dividing the main topic into two sections, namely *Section 2. Introduction* and *Section 4. Biases in mediated job interviews*. Initially, an additional chapter about biases in non-mediated (i.e., FTF) job interviews had been considered, but it was later discarded. The following criteria were used to decide whether or not to keep an article and move it to the next search step: A) Being in English; B) Being related to the topic of biases or distortions in technology-mediated job interviews and C) Being published in a journal or being "in press" (thus, thesis dissertations were excluded). Given the interdisciplinary nature of the topic addressed, publications from fields different from Psychology, such as Computer Science and Economics, have been included in the research process as well.

Thirteen keywords were selected, combined and then inserted in the Scopus searching tool. Looking for articles related to biases in AI-mediated job interviews (Section 2), the primary keywords *artificial intelligence*, *algorithm* and *computer program* were used. These keywords were then combined with secondary keywords *job interview*, *job application*, *personnel selection* and *talent screening*, and both primary and secondary keywords were finally paired with tertiary keywords *bias*, *ethic* and *self-presentation*[8]. The complete list of combinations (according to keyword category) is the following: 1) primary-secondary-tertiary, 2) primary-secondary, 3) primary-tertiary, and 4) secondary-tertiary. Of all the results obtained, only the relevant ones from the combination primary-secondary-tertiary were considered for Section 2. Looking for articles related to biases in mediated job interviews without AI use (Section 4), the primary keywords were *remote*, and *video*. These were then combined with secondary keywords *job interview*, *job application*, *personnel selection* and *talent screening*. Finally, primary

[8] These and the subsequent pairings of keywords were always treated as one search term and logically separated by "AND".

and secondary keywords were paired with tertiary keywords *bias* and *self-presentation*. For Section 4, the results for all the combinations were considered, with some exceptions presented later. **Figure 2** provides a flow diagram summarising all the steps taken during the research process and the number of sources excluded.

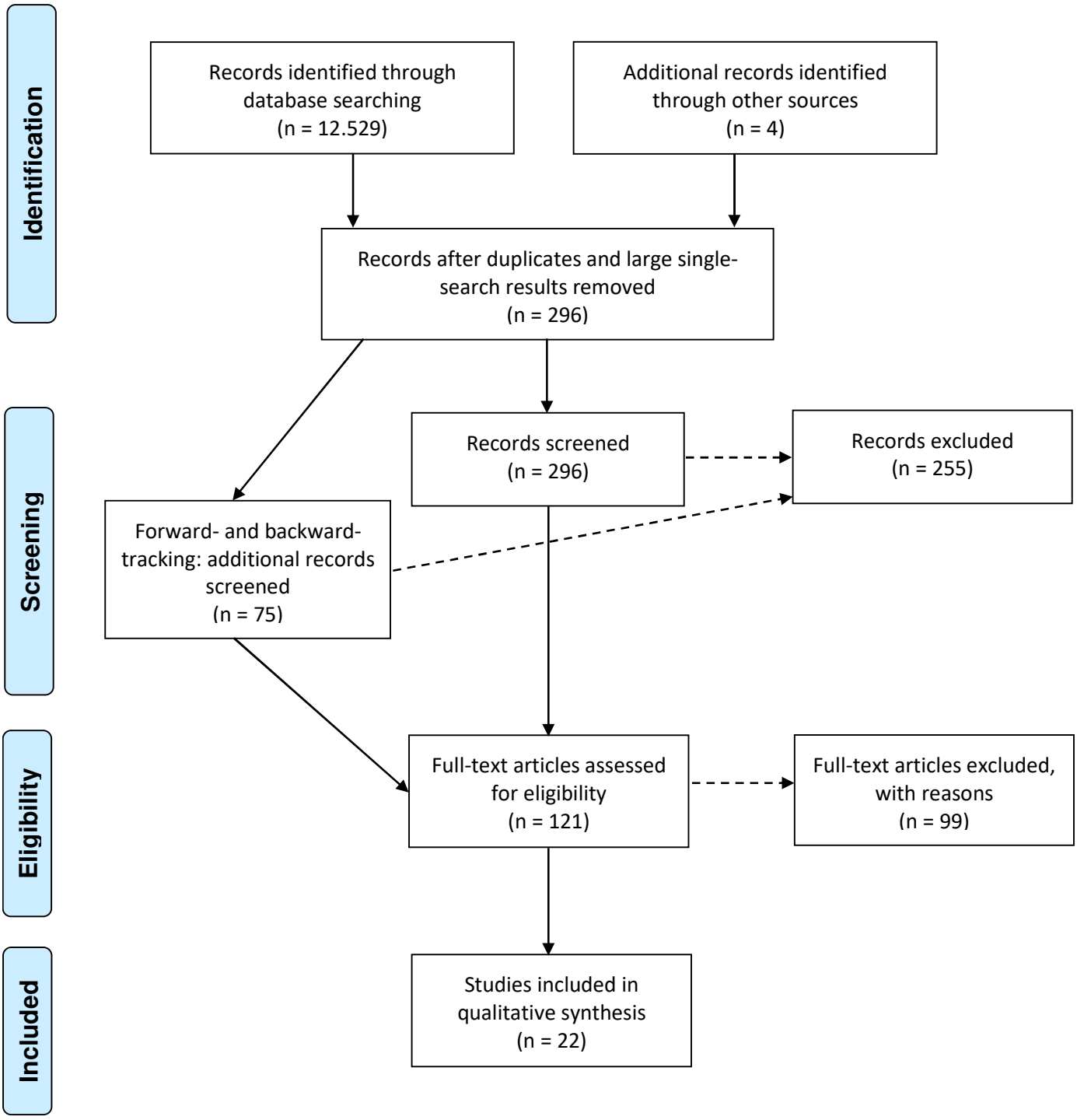

Figure 2. *Customised PRISMA flow diagram of search steps* (see Moher, Liberati, Tetzlaff and Altman, The PRISMA Group, 2009, for the original).

During the research process (which took place between September and October 2021), the Scopus database yielded 12.529 results. The results of three searches were excluded: *TITLE-ABS-KEY (video AND bias)*, 4040 results; *TITLE-ABS-KEY (remote AND bias)*, 7877 results; and *TITLE-ABS-KEY ("personnel selection" AND bias)*, 315 results. The reason for this exclusion lies in the sheer number of articles (compared to the results of other searches) found for each single result, which was considered as a proxy for the presence of several articles not related to the topic of interest for this review. After the exclusion of the three aforementioned search results and following a check for duplicates, the number of articles diminished to 292. In the meanwhile, my thesis supervisor provided three additional articles, making the total 295. The next step consisted in reading just the titles and abstracts of the selected articles and this led to the exclusion of 200 of them. In order to provide a systematic framework for the rejection process and to reduce arbitrariness, several exclusion criteria were created through a bottom-up process; i.e., whenever an article needed to be rejected for a reason not already represented by an existing criterion, that article received a new exclusion criterion. Thus, ten criteria[9] were created in this way. **Table 1** presents these criteria, along with the number of articles that satisfied each exclusion criterion. Some articles received more than one of them. Moreover, in compliance with rule C (i.e., an article must be already published in a journal or be "in press"), PhD publications were discarded due to a possible lack of peer review.

[9] The criterion *Unrelated topic* is an umbrella category given when conditions to apply the following criteria were satisfied at the same time: *Does not directly assess mediated job interviews*, *Lack of technological mediation* and *Does not directly assess biases*.
Niche topic means that the object of the article is or may be related to job interviews and biases influencing their outcomes, but focuses on very specific positions and issues at the expense of its generalisability (e.g.: residency for job specialisation in the medical field, career advancements for university professors).

Exclusion criteria	N. of articles excluded
Does not directly assess mediated job interviews	84
Lack of technological mediation	5
Does not directly assess biases	85
Unrelated topic	79
Niche topic	3
No AI use	2
Article not available	1
Non-English language only	3
PhD dissertation	2
Little/lack of support for biases	5

Table 2. Overview of the number of exclusion criteria and the number of articles to which they were applied. Yellow boxes contain criteria used only during the first selection step (title and abstract reading), while the blue box shows the one used only during full-text reading. A green box means that the criterion was applied at least once in both steps.

Subsequently, the remaining ninety-two papers were divided according to the topic addressed: nine were identified as related to *Section 2: Introduction*, eight as related to *Section 4: Biases in mediated job interviews*, and seventy-

eight as related to the dropped section on biases in FTF job interviews[10]. Afterwards, the next step consisted in the full-text reading of the articles belonging to Section 2 and Section 4 and in applying exclusion criteria to those deemed not relevant for this review, following the same procedure adopted before. Finally, articles initially placed in the wrong section were moved to the correct one. At the end of this process, nine articles successfully passed the selection, of which five belonged to Section 2 (a sufficient amount for this Section) and four belonged to Section 4 (an insufficient amount). Given the lack of articles for Section 4, the process of *forward and backward tracking* (a.k.a, *snowballing*) was adopted, which consisted in finding additional articles through the references of those already available for Sections 2 and 4, as well as checking articles that cited those of Section 4. This last operation was performed on Google Scholar. However, not all the articles found in this way were kept. In fact, the same selection process adopted above (i.e., the one using exclusion criteria) was used to select the relevant articles among those found on Scholar. More specifically, seventy-five new articles were found in this way, of which twelve were kept after full-text reading. Subsequently, one additional article about AI functioning was added to the bibliography, as it proved necessary in writing part of Section 2. As a result, the total number of articles used as sources is 22[11]. **Table 5** in the Appendix gives an overview of the articles that received a full-text reading and that then became part of the bibliography.

Finally, it is worth mentioning that a handful of citations present in the selected articles provided some corollary material for this review. As such, certain items in the bibliography section are preceded by a disclaimer *(*=articles not consulted directly*) indicating that a direct consultation of that source did not happen.

---

[10] As mentioned before, this sub-topic and its associated articles were dropped due to time constraints, as reading entirely those articles would have taken a considerable amount of extra time. This decision was taken in agreement with my supervisor.

[11] Resulting from the sum of the nine articles found on Scopus, plus the twelve found through forward and backward tracking, plus the one on AI functioning.

4. Biases in mediated job interviews

This section will try to address the issue of biases and, more in general, distortions in mediated job interviews. The aim of the following paragraphs is not to see which distortions of traditional FTF interviews transfer to mediated interviews and, therefore, just a brief coverage of these will be provided and only if needed. Instead, the focus will be on all those distortions that arise due to the presence of technological mediation, i.e., distortions that would not be present otherwise. To do so, six questions guided the examination of the literature on the topic and they received an answer based on the evidence present in such literature. As illustrated in Section 2, these questions are:

1) *Is the presence of bias in mediated job interviews supported?*
2) *Are there statistically significant distortions arising from the use of technology?*
3) *What are the psychological and/or physical phenomena reported as exclusive to mediated job interviews?*
4) *Are mediated job interviews suggested as preferable in the study under consideration? If yes/no, why?*
5) *What are the reported pros and cons of mediated job interviews?*
6) *Do authors propose solutions to solve the issues of mediated job interviews?*

The reader should keep in mind that each examined article provided, at best, three answers to the above questions. Moreover, it was not possible to address each individual issue with information from *all* the articles, as some of them, for instance, just outlined problems of technological mediation without proposing solutions. This happened because those six questions were created to better navigate through the information available and find the relevant pieces of evidence. Therefore, there was little-to-no prior knowledge on the topic of job interviews that could have helped in the creation of such questions. As a result, the literature provided overlapping answers to some questions and this gave rise to two major issues, i.e., a single piece of evidence often addressed Questions 3 and 5 or Questions 1 and 2 together, while Question 4 was relevant only for one article. To solve these issues, Question 5 was addressed by only considering the

more methodological and procedural aspects related to mediation, while the discussion on the subjective and psychological ones was reserved to Question 3. Moreover, Question 1 was merged with Question 2 and the same happened for Question 4 with Question 5, so the merged questions will be treated together in their dedicated sub-sections. Finally, **Table 3** (independent variables) and **Table 4** (dependent variables) present a clustering[12] of the variables encountered in the examined articles.

[12] i.e., a grouping of those variables under a name or concept that express a basic characteristic they have in common.

Table 3. Clustered independent variables

Cluster	Description	Factors
Characteristics of the inter-viewer	Independent variables related to physical characteristics of interviewers or actions performed by them.	Gender: *male vs. female* (Antonio Gómez Jáuregui, D., 2021; Chapman, D. S., 2001) Movement: *mirror vs. random* (Antonio Gómez Jáuregui, D., 2021)
Characteristics of the applicant	Independent variables related to physical and/or psychological characteristics of the participant in real life or in a virtual environment.	Applicant's gender: *male vs. female* (Behrend, T., 2012; Köchling, A., 2021; Silvester, J., 2003) Ethnicity: *Asian vs. Caucasian vs. African-American* (Köchling, A., 2021) Attractiveness: *above median attractiveness score vs. below median attractiveness score* (Straus, S. G., 2001) Perceived avatar attractiveness (Behrend, T., 2012) Educational background: *computer science students vs. non-computer science students* (Langer, M., 2018) Personality trait: *extraverted vs. introverted* (Nørskov, S., 2020)

Cluster	Description	Factors
Medium features	Independent variables related to the technological devices used during real-life or simulated job interviews. Manipulation of these variables consisted in using different devices or in distorting the recordings made through them.	Interview medium: *FTF vs. technology mediation* (Blacksmith, N., 2016); *Audio-only vs. audiovisual vs. teleoperated* (solo task - Bremner, P. A., 2017) *Audio-only vs. teleoperated* (dyadic task - Bremner, P. A., 2017) *FTF vs videoconference* (Chapman, D. S., 2001); *Digital vs. videoconference* (Langer, M., 2017); *Robot-mediated vs. FTF* (Nørskov, S., 2020); *FTF vs. telephone* (Silvester, J., 2003) *FTF vs. telephone vs. videoconference* (interviewers - Straus, S. G., 2001) *Videoconference vs. FTF* (independent raters - Straus, S. G., 2001); Audio and video fluency: *fluent vs. disfluent* (Fiechter, J. L., 2018)
Automation	Independent variable referring to the output given by an AI (neural network) after training on a dataset.	Algorithmic predictions: *BU-NKU algorithm vs. ROCHCI algorithm vs. raw data* (Köchling, A., 2021)

Cluster	Description	Factors
Task features	Independent variables related to the structure, methodology, type and/or characteristics of the interview or task.	INTERVIEWS Interview structure: *high vs. semi-structured vs. low* (Chapman, D. S., 2001) Trial order: *Resume first vs. AVI first* (Torres, E. N., 2018) *First applicant vs. second applicant* (Straus, S. G., 2001) NON-INTERVIEWS Number of persons conducting the task: *solo vs. dyadic, i.e., with a confederate* (Bremner, P. A., 2017) Type of task executed: *Talk about a hobby vs. tell a story using TAT vs. mime a cooking situation* (solo - Bremner, P. A., 2017) *Describe a clip vs. cooperation vs. competition* (dyadic - Bremner, P. A., 2017) Job gender type: *masculine vs. feminine* (Behrend, T., 2012)

Cluster	Description	Factors
Information provided to the participant	Independent variables related to the availability or type of information provided to participants.	Level of information about AI functioning in a personnel selection context: *high vs. low* (Langer, M., 2018) Degree of automation described in the job advertisement: *fully-automated vs. semi-automated vs. non-automated* (Wesche, J. S., 2021) Level of information on automation use in the screening stage: *non-automated vs. automated vs. no information* (Wesche, J. S., 2021) Information on employee benefits: *information vs. no information* (Wesche, J. S., 2021)

Table 4. Clustered dependent variables.

Clusters	Description	Dimensions
Participants' behavioural measures	Dependent variables related to postural and body behaviours of participants during the task.	Contraction index, quantity of motion, medio-lateral displacement of the centre of pressure x, anterior-posterior displacement of the centre of pressure y (Jáuregui, D. A. G., 2021)

Participants' self-reports/questionnaires	Dependent variables primarily related to participants' feelings about and perceptions of the task, interviewer or situation described. Some dependent variables also assess participants' psychological traits and states.	**QUANTITATIVE** PRE-TASK Big 5 inventory, threat challenge appraisals, perception of virtual interviewer, perceived humanness scores, perceived dominance and warmth (Antonio Gómez Jáuregui, D., 2021) POST-TASK Perceived difficulty and stress; self-assessment manikin (Jáuregui, D. A. G., 2021). Creepiness of situation, privacy concerns, perceived behavioural control, two-way communication, interpersonal treatment, chance to perform, global fairness, organisational attractiveness (Langer, M., 2017). Job relatedness, information known, consistency, open treatment, interpersonal treatment, opportunity to perform, transparency, fairness, creepiness, privacy concerns, overall organisational attractiveness (Langer, M., 2018). Procedural fairness, interactional fairness, behavioural intentions (Nørskov, S., 2020). Communication understanding, conversation fluency, applicant comfort, self-consciousness, interviewer likability (Straus, S. G., 2001). Intentions to apply, organisational attractiveness, expected fairness of the selection procedure, writing of the job advertisement, difference in elo-

		quence between job advertisements, perceived organisational attractiveness and prestige, expected justice of the selection process (Wesche, J. S., 2021) NOT SPECIFIED: Selection procedural justice, perceived fairness, litigation intentions, expectancy of a favourable outcome, perceived difficulty, satisfaction with performance, ratings of the interviewer, intentions to pursue/accept, organisational attractiveness, conversation fluency, self-consciousness, comfort, concerns about technology (Blacksmith, N., 2016) Conscientiousness, neuroticism, job interview score (Köchling, A., 2021) **QUALITATIVE** POST-TASK Description of how participants felt about the selection process (Wesche, J. S., 2021)

Interviewers' impressions and ratings of applicants	Dependent variables related to interviewers' evaluations of applicants and opinions on videoconference technology.	**QUANTITATIVE** PRE-INTERVIEW: Overall impression of the applicant based on written info, appropriateness of applicant's educational background for the position, evaluation of applicants previous experience, educational achievement (Chapman, D. S., 2001) POST-INTERVIEW: Perceived interview performance, screening recommendation (Behrend, T., 2012) Overall impression after interview, communication skills, intelligence, creativity, assertiveness, preparation for the interview, friendliness, leadership ability, appearance, decisiveness, support for their arguments, interpersonal skills, motivation, initiative, likeability, confidence (Chapman, D. S., 2001) Hireability, likeability (Fiechter, J. L., 2018) Applicant rating (Langer, M., 2017) Overall rating of the applicant (Silvester, J., 2003) Applicant's general abilities, likeability, physical attractiveness, communication understanding, conversation fluency (Straus, S. G., 2001) Leadership ability, customer focus, problem-solving skills, ability to deliver results, ability to work in a team, communication skills (Torres, E. N., 2018)

		NOT SPECIFIED: Overall ratings; specific ratings of skills, abilities, competencies, talkativeness, conversation fluency, expressiveness and likeability (Blacksmith, N., 2016) **QUALITATIVE** POST-INTERVIEW Feelings about using videoconference interviews, how videoconference technology may have changed interviews, satisfaction with the use of videoconference technology (Chapman, D. S., 2001) Advancement to the next recruitment stage: *yes vs. no* (Torres, E. N., 2018)
Raters agreement	Dependent variables related to the agreement of a rater with him/herself, other raters and participants' self-ratings.	Intra-judge agreement: *quality of personality judgments based on correlations between different questionnaire items. These items measure the same personality trait and allow to check the internal consistency of each judge* (Bremner, P. A., 2017) Inter-judge agreement: *consensus among judges* (Bremner, P. A., 2017) Self-other agreement: *similarity between personality judgments made by self and others* (Bremner, P. A., 2017) Personality shifts: *extent to which participants shifted from one personality to another, in judges perception, between conditions* (Bremner, P. A., 2017)

Algorithmic ratings and fairness measures	Dependent variables that measure the level of bias present in an AI algorithm.	Rating scores of applicants on conscientiousness, neuroticism and job interview (Köchling, A., 2021) Group-comparison-based fairness measures applied to job interview scores: *DI*, *1/EqOppoFN* (Köchling, A., 2021) Accuracy-based fairness measures applied to job interview score: *BCR*, *NMI*, *1-MSE* (Köchling, A., 2021)

4.1 Biases and distortions in mediated interviews

This sub-section is concerned with the first two research questions: *1) Is the presence of bias in mediated job interviews supported?* and *2) Are there statistically significant distortions due to the use of technology?.* Several studies provided one or more insights into the issues of biases and distortions arising from technology mediation in job interviews. In the following lines, pieces of evidence and personal comments will be presented by roughly following the order of appearance of mediating technologies, i.e., TEL, VC, AVIs and virtual or physical avatars.

4.1.1 Telephone (TEL) interviews

Two studies directly addressed the differences between recruiters' evaluations of applicants in FTF and telephone-mediated interviews, finding statistically significant distortions. In fact, researchers found that applicants communicating via telephone received better ratings than applicants did in FTF and VC interviews (Straus, Miles and Levesque, 2001; Silvester and Anderson, 2003). The telephone medium, in fact, had the effect of suppressing less favourable visual characteristics (i.e., those not considered "beautiful") by restricting the media bandwidth, which reflects the number of different cues or signals that a medium transmits. Therefore, less communication cues were available in TEL condition compared to VC and FTF interviews, where visual information is also available to the recruiter (Straus et al., 2001). Moreover, as a telephone interview yielded a better evaluation even for physical attractive applicants, these researchers speculated that the observed effects might be related to two phenomena, i.e., applicants sounding more enthusiastic to compensate the lack of nonverbal cues or a lower cognitive workload imposed by the telephone medium.

Another significant difference between FTF and TEL interviews is that raters in a field study gave higher interview scores to those applicants in the TEL condition who attributed positive outcomes to more personal causes (e.g., "*They chose me because I had been team captain three years in a row*"; Silvester and Anderson, 2003). However, the same did not happen for FTF interviews. The

interesting details emerging from this study are that participants in the TEL condition actually tended to make *more* personal attributions compared to FTF participants and that recruiters tended to ask more closed questions in TEL interviews, although this last effect did not reach full statistical significance. Along with these distortions, which create differences between modalities, the researchers evidenced an "equaliser" effect in the TEL condition. In FTF interviews, recruiters asked a higher amount of open-ended questions to males, whom also produced more discourse, while no gender differences emerged in TEL interviews (Silvester and Anderson, 2003). From a practical perspective, if companies are able to control for the positive bias towards TEL applicants (e.g., by interviewing all applicants through telephone), this last effect suggests a positive role of TEL in curbing gender biases in the context of job interviews. Unfortunately, a more recent study found a distortion in TEL interviews opposite to the one just discussed, i.e., applicants received lower ratings in this condition compared to participants who did a FTF interview (Torres and Gregory, 2018). This discrepancy may be the result of an implicit difference between the methodologies of the previous studies and the one of Torres and Gregory or, more interestingly, it may be due a different perception recruiters had of the telephone medium. In fact, it is probable that a viable reason for this discrepancy is that technology-mediated interviews were rarer compared to the present days, as VC interviews and AVIs gained much more popularity since the early 2000s. Therefore, recruiters might have been more moderate and benevolent in older studies, because they and/or the applicants were not so accustomed to the modality. Although speculative in nature, this remark could be of interest for future studies.

4.1.2 Videoconference (VC) interviews

In general, the literature evidences two major biases attributable to VC interviews (and to AVIs as well): *Background bias* and *Cognitive Load bias* (Allam, J. S., 2021; Lukacik et al., *in press*). *Background bias* arises because applicants of job interviews, while being evaluated trough VC or AVIs, are located in their homes. Therefore, the presence of a non-neutral background can reveal aspects of the candidate's private life that would not be accessible to interviewers

in traditional FTF interviews, such as the candidate's living situation. This kind of info may prompt, as a result, the interviewer to make undesired assumptions about the applicant and bias the interviewer in producing the final score (Allam, J. S., 2021). The reverse is also true, although the interviewer can actively limit such bias by choosing a neutral background or location, such as their workplace. (Lukacik et al., *in press*). *Cognitive Load bias*, instead, arises due to the specific layout of videoconference platforms, as interviewers need to focus unnaturally on camera to demonstrate interest in the conversation, extrapolate more meaning from the conversation because of a lack of nonverbal cues (as there is a decreased bandwidth) and there is a small window providing immediate and continuous feedback of their performance. This increased use of mental resources can cause interviewers to be more reliant on mental shortcuts and implicit biases in rating candidates (Allam, J. S., 2021). Interestingly, this bias can also affect participants while they are recording themselves for an AVI and hinder their ability to manage their impression, especially when deception is involved (Lukacik et al., *in press*).

In a study[13] comparing VC and FTF modalities, as well as manipulating their level of structure and the gender of interviewers, researchers observed that interviewers rated applicants undergoing a VC interview higher than they did for applicants in the FTF condition (Chapman and Rowe, 2001). Although authors found other variables involved in this discrepancy, the interview medium played a substantial role by accounting for 15% of the variance in ratings. To explain these results, the researchers proposed that the VC medium helped applicants to reduce anxiety during the interview, as the interviewer was not physically present. In alternative, interviewers may have made external attributions and applied naïve theories of bias corrections in favour of the VC applicants, such as being more benevolent, as the VC technology was relatively new at the time the study took place (Chapman and Rowe, 2001).

[13] Only two of the examined articles specifically addressed potential differences between VC and FTF interviews, so information on this aspect is probably incomplete.

Finally, a meta-analysis by Blacksmith et al. (2016) provides further insights on VC, although within the bigger framework of mediated interviews. They found that interviewers' ratings and applicants' perceptions were lower in mediated (i.e., TEL and VC) compared to FTF interviews. Two variables named *Study setting* and *Publication date*, however, moderated lower interviewers' ratings. In fact, a larger negative effect was found in real interviews compared to laboratory ones and in new studies compared to the old ones. This brings about serious concerns about the external validity (and the generalisability) of some studies, especially the ones on TEL interviews, which are often dated and rare nowadays. As already hinted, applicants preferred FTF interviews to mediated ones, with the least appreciated being VC interviews. According to the researchers, this is attributable to the restriction on applicants' ability to impression manage (i.e., influence the interviewer's perceptions) imposed by technological mediation (Blacksmith et al., 2016). However, Blacksmith and colleagues found a second-order sampling error for this second effect, so these results should be approached with caution. Similarly, AVIs exacerbate this impediment on applicants' ability to manage impressions, as other researchers found lower general intentions to use impression management in AVI than FTF or VC interviews (Lukacik et al., *in press*). It is worth of note that Blacksmith et al.'s (2016) meta-analysis also included the studies of Chapman and Rowe (2001), Silvester et al. (2003) and Straus et al. (2001). This meta-analysis, therefore, allowed the researchers to insert those studies within a higher framework, useful for making comparisons across different investigations and test the generalisability of single-study results. In fact, this proved to be essential in discovering an overall negative influence of technological mediation on ratings of applicants, despite the positive effects suggested by the three initial studies.

4.1.3 Asynchronous Video Interviews (AVIs)

Interviews conducted through recorded videos present biases and distortions as well. The following results are innovative with respect to Blacksmith et al.'s (2016) meta-analysis because they were published later. The first critical aspect that affects interviewees' ratings in AVIs is the fluency of the recording.

Fluency exerts a great influence on recruiters' scorings of an applicant's variable called *hireability*, i.e., the probability of the interviewing organisation to employ a given applicant (Fiechter, Fealing, Gerrard and Kornell, 2018). Moreover, the same researchers were not able to rule out a possible significant decrease in *likeability* of applicants due to disfluent videos. The reasons behind these effects are worth of further inquiry, as, in one of the two experiments, researchers *explicitly* instructed the participants who posed as interviewers to discount eventual video or audio lags. This, however, did not happen. Therefore, researchers speculated that disfluency affected interviewers at an implicit level or interviewers disagreed with the instruction received, as they deemed poor AV quality to be a sign of an unprepared candidate. A simpler explanation is that the cognitive load derived from also having to press buttons may have hindered interviewers' attention and, therefore, their efforts to discount AV disfluency (Fiechter et al., 2018). If true, this last phenomenon would be of great worry for professional recruiters and HRM departments, as a lower level of attention could easily counteract explicit instructions and training, with an increased risk for discrimination.

Although a fluent and well-recorded AVI is a first step towards a fair evaluation, it is not enough. In fact, the order in which a recruiter sees the AVI (i.e., before or after other material informative of the candidate) is another source of distortions (Torres and Gregory, 2018) and the same goes for acquiring information about candidates with different modalities (e.g., résumé vs. AVI; Langer, König and Krause, 2017; Torres and Gregory, 2018). More specifically, if a recruiter sees the AVI of a candidate before their résumé, that recruiter will give a higher score to that candidate compared to a situation in which résumé inspection came first (Torres and Gregory, 2018). A possible explanation for this is that résumés create higher expectations that, if not satisfied during the subsequent AVI inspection, will lead to harsher judgments of the candidate. When recruiters inspection AVIs first, instead, they create lower expectations of the candidate or simply lack background knowledge that would influence the scores negatively. On a side note, AVIs received higher overall ratings compared to résumés in this study, although the effect size was small (Torres and Gregory, 2018). In addition, Langer et al. (2017) found a difference in ratings between VC and AVIs, with

participants receiving better ratings in the latter. Interestingly, qualitative data about interviewers' reactions reveal how they perceived applicants as responding in a more detailed manner in AVIs, while they felt applicants gave shorter answers in VC interviews (Langer et al, 2017).

4.1.4 Virtual and physical avatars

Research in various disciplines has also investigated distortions and reactions in interviews where the interviewer and/or the applicant were seeing a proxy of the other and, therefore, not the actual body or image of their conversational partner. Virtual avatars can significantly influence recruiters' assessment of a candidate's performance, as the appearance of avatars seems to have an effect on ratings even when aesthetics is not relevant to the evaluated performance (Behrend, Thompson and Sharek, 2012). A similar bias is exacerbated when applicants have the freedom to create an avatar that does not match the appearance of the candidate, as in Behrend et al. (2012) study. This finding, therefore, hints at the possibility of a halo-like effect[14] being transferred to virtual avatars too, despite their obvious simplification of human appearance. In another investigation, interviewee of a within-subject study judged a human-virtual agent interview to be as threatening as the human-human interview before the interaction and as difficult but with lower perceived stress after the interaction. Nevertheless, the absolute ratings on the stress scale for both interviews were close (Jáuregui, Giraud, Isableu and Martin, 2021).

Other studies, instead, explored participants' fairness perceptions and judge agreement of mediated interviews using a robot proxy (Bremner, Celiktutan and Gunes, 2017; Nørskov et al., 2020). In an interview where a humanoid robot capable of motion-capture technology, (such as the NAO) mediates the interaction, external observers had more difficulties in expressing a judgment about the participant's personality. In fact, robot mediation hindered *Inter-judge agreement*, i.e., the degree to which different observers expressed the same judgment about

---

[14] A *halo effect* is a rating bias in which evaluation of an individual in one dimension influences further judgments on other dimensions (APA Dictionary, 2022d).

an applicant, and *Self-other agreement*, i.e., the degree of concordance between the applicant self-perception and an observer's judgment (Bremner et al., 2017). However, the most remarkable finding is that observers attributed personality stereotypes, based on appearance, to the robot when participants were speaking through it. The result is that observers actually incorporated cues of the robot's appearance into the personality judgments of interviewees, in some cases even overriding the cues conveyed through speech by the interviewees themselves (Bremner et al., 2017). Robot proxies, other than distorting personality assessment, are also responsible of distorted perceptions made by applicants. In fact, compared to FTF interviews, the robot-mediated condition led applicants to express lower fairness perceptions and intentions to accept the job (Nørskov et al, 2020) and organisations should not underestimate such effects. Indeed, based on this last piece of evidence, firms adopting these kind of interviews face the added risk of losing high-quality applicants.

4.2 Psychological phenomena of mediated interviews

This sub-section deals with the research question *3) What are the psychological and/ or physical phenomena reported as exclusive to mediated job interviews?*. The focus of this sub-section will mainly be on interviewers and applicants' perceptions, expressed either by real-life interviewers and applicants or by participants assuming one of these roles during a study[15].

Chapman and Rowe (2001), in their study comparing FTF to VC interviews, provide interesting qualitative data on interviewers' perceptions and impressions of the two interview modalities. More than 67% of interviewers reported that they preferred FTF interviews and that VC was of no particular help. Some

[15] The term *perceptions* refers to how people immediately feel about and understand an interview modality, along with how they feel and understand their relationship with the interviewer. The same holds true for interviewers, who instead develop perceptions of applicants (this definition was adapted from Langer and Landers, 2021, whom by contrast addressed perceptions of AI-assisted job interviews).

also reported difficulties in assessing applicants' *Appearance*, *Confidence*, *Assertiveness* and nonverbal behaviour, as well as attribution of error. Although it seems straightforward that this is mainly due to the technology available at the time (more than 25% reported image and audio issues), nonverbal behaviour remains difficult to assess even nowadays, as interviewers cannot see an applicant's body in its entirety. In spite of this, some interviewers (16%) declared to have had an easier time in taking more notes, checking their watches or looking at résumés in the VC condition. Moreover, these interviewers said that it was also easier to concentrate on the speech of the applicant and that they felt less compelled to appear positive with the interviewee in the VC condition compared to the FTF one. A minority of the interviewers involved in the study reported the dimensions of *Communication skills*, *Friendliness* and *Support for Arguments* as easier to assess in the VC compared to the FTF modality (Chapman and Rowe, 2001). Probably, the increased attention devoted to applicants' speech played a role in this perceived easiness of assessment. Similarly, speech played a role also in Silvester and Anderson study (2003) as applicants in the TEL condition produced significantly more personal attributions, i.e., they attributed the cause of an event as something unique to them as individuals and this did not happen in the FTF condition.

Aside from the previous investigation, most of the studies in the articles focused instead on applicant's reactions, which are predominantly negative about the use of VC, AVI and avatars in job interviews. In fact, a much recent article reports lower perceptions of fairness and likelihood to accept a job offer in VC and TEL interviews compared to FTF ones (Torres and Mejia, 2017). In addition, applicants felt lower perceptions of procedural justice in interviews where communication happened through a computer and a more impersonal interaction in VC compared to FTF (Torres and Mejia, 2017; Langer et al., 2017). AVIs received the majority of attention in the literature and more than one study reported more negative reactions towards AVIs than towards other forms of mediated interviews. Applicants reported, related to AVIs, a lack of real-time feedback and the inability of portraying oneself realistically, as well as more feelings of creepiness and concerns about their privacy (Torres and Mejia, 2017; Langer et al., 2017).

Langer et al. (2017) propose that creepiness is enhanced in AVIs because participants are somewhat forced to monitor their performance while recording the interview, or because they simply are not sufficiently accustomed with the format. Privacy concerns, on the other hand, arise since participants feel that a third party, not involved in the recruitment procedure, may watch their recording (Langer et al., 2017; see the Surveillance facet in Potosky, 2008, for a more detailed explanation on such concerns). In addition, applicants perceive AVIs as more impersonal than VC and FTF interviews, as well as "cold" and inappropriate for proposing a final job offer, with the potential and unwanted outcome of well-fitting applicants deciding to abandon the selection procedure (Langer et al., 2017). Despite a general negative perception of the AVI format, participants of the examined studies highlighted some benefits and positive aspects of the medium, such as those related to the dimensions of *Job-relatedness*, *Opportunity to perform* and *Fairness* (Lukacik et al., *in press*). Indeed, when researchers emphasised the standardisation aspect of AVIs, participants rated their fairness higher. In addition, participants asked to imagine an AVI design that allowed interviewee 60 seconds of preparation (before giving a time-restricted answer) rated it similarly to VC on *Fairness* and *Opportunity to perform* (Lukacik et al., *in press*). Although this is a point in favour of AVI, participants still perceived VC interviews as more unfair than FTF ones (as previously mentioned), so FTF appears to be the overall preferred modality according to the evidence available.

Finally, it is the turn of virtual and physical avatars. Both versions are sources of issues related to fairness. In fact, virtual avatars that interviewers perceived as attractive had more chances to continue through a selection procedure compared to less attractive ones, while using a robot avatar during an interview led applicants to lower perceptions of fairness and intentions to accept a job compared to FTF interviews (Torres and Mejia, 2017; Nørskov et al, 2020). One peculiar effect digital-avatar mediation has on interviewers is of extending a gender-specific bias to both genders. In fact, while in FTF interviews applicants' attractiveness influences women's evaluations more, attractiveness of digital avatars affects both genders equally (Behrend et al, 2012). Behrend et al. (2012) suggest that this happens because raters use the avatar's appearance as a benchmark

to assess job-relevant characteristics while following a “beautiful = good” inference, similarly to what happens with the halo effect (as anticipated in the previous sub-section). Regarding physical avatars, they can distort the assessment of personality and create meaningful differences in self-perceptions of interviewees and the perceptions others have of them, as testified in the study about observers’ agreement on personality judgments (Bremner et al., 2017). When participants who played the role of applicants were asked about their experience with a physical avatar, some stated they perceived the robot as "a doll" and "bizarre", thus indicating a potential uncanny valley effect[16] (the robot had a minimal human design, that prompted uneasiness in participants because of its ambiguous features; Nørskov et al., 2020). Nørskov and colleagues claim that applicants perceived their robot-mediated interview as more unfair because a technological tool was placed in-between applicants and interviewers. In this sense, the interview design probably contributed to those negative perceptions since the robot mediation did not allow applicants to adjust their self-promotion tactics, as the robot only mimicked the operator's head movement, emulated the operator's lip movements and transmitted the operator's speech. (Nørskov et al., 2020).

4.3 Pros and cons of mediated interviews

The following paragraphs will present evidence tackling research questions *4) Are mediated job interviews suggested as preferable in this study? If yes/no, why?* and *5) What are the reported pros and cons of mediated job interviews?*. The literature reports many advantages for the AVI and robot-avatar modalities, as well several disadvantages for the aforementioned modalities plus VC and TEL interviews[17].

---

[16] The *uncanny valley effect* consists in the unpleasant feeling arising by seeing a robot that appears very similar to a living human (Cambridge Dictionary, 2022).

[17] A possible reason behind the lack of reported benefits of TEL interviews is that organisations adopted this technology early and, reasonably, the hype behind it had probably deflated long ago. Consequently, this review might have missed old reports on the benefits and excitement behind the adoption of the TEL technology. It is possible to advance a similar explanation for VC interviews, because only a handful of

In general, reported advantages of adopting technology-mediated interviews consist in a larger number of applicants available, simplified and faster selection procedures, increased ability of selecting adequate or high-potential applicants and a simplified storage/usage of applicants' information (Torres and Meija, 2017; see Stone et. al, 2013, for the original source). More specifically, many articles report on the advantages of AVIs, which include the possibility for applicants to apply from anywhere and at any time - saving time and resources for both the organisation and the applicant - further automation of the process and increased validity and reliability (Dafoulas, Pateli, Turega, 2002; Torres and Gregory, 2018; Lukacik et al., *in press*). In particular, researchers report higher reliability and validity due to the increase in interview structure (Torres and Gregory, 2018; Lukacik et al., *in press*). In fact, AVIs allow to ask consistent questions across applicants and longer interviews and responses, while they do not allow applicants' questions (that are difficult to standardise). Furthermore, given that organisations can share applicants' recordings quickly, AVIs also support a panel rating system, with the additional advantage that distortions linked to interviewer presence are absent (Lukacik et al., *in press*). On a side note, the increase in interview structure is often welcomed, as ratings in such interviews correlate more with job performance compared to unstructured interviews (Torres and Gregory, 2018). Dafoulas et al. (2002) report in their commentary article that the adoption of asynchronous interviews would bring about a decrease of biased decision-making for interviewers, a decrease of unwanted psychological burdens in applicants, equal opportunities for candidates and an improvement of relationships between organisations and applicants. However, it must be specified that these authors devised asynchronous interviews as essentially online questionnaires with open-ended questions, without the presence of a video or recording of applicants. Although the effects outlined above are for sure positive, no trace of written asynchronous interviews was found in the literature and, therefore, there is no material on this kind of modality. Still, it can be claimed with much more

---

articles from the late 90's and the early 2000s are present here and all of them describe VC as something more or less established in recruitment procedures.

certainty that such desired outcomes are rare in AVIs, as demonstrated by the distortions and negative perceptions that applicants have about them, which are reported in sub-sections 4.1 and 4.2. Finally, with regard to robot avatars, Nørskov et al. (2020) propose that applicants who faced discrimination in the past will perceive them more positively because robot-mediated interviews promise more fairness and objectivity, at least on paper.

The literature reports several downsides of technological mediation in job interviews. They range from simpler ones, such as audio and video lag, issues with image clarity and lack of responsiveness (Chapman and Rowe, 2001; Commissaris, 2021) to more sophisticated ones. The formers, although trivial, may negatively influence the ability of recruiters to connect with applicants. Furthermore, insufficient eye contact, as well as poor camera angles, lightning and backgrounds can achieve a similar negative effect too (Commissaris, 2021). Moreover, there is evidence that audio lags in telephone conversations decrease feelings of “solidarity” and “belonging” and similar perceptions arise as well when the video is present but degraded (Commissaris, 2021). Despite AVIs looking like a suitable medium to circumvent lag issues, they present a fundamental shortcoming, i.e., recruiters can arbitrarily decide to skip applicants’ recordings without actually viewing those (Lukacik et al., *in press*). This issue, absent in FTF, TEL or VC interviews, becomes of paramount importance in AVIs, since discarding an applicant based on the recruiter’s first impressions presents a high risk of bias and discrimination.

Medium drawbacks constitute a burden not only for the applicant, but also for organisations and interviewers, as those who evaluate applicants using AVIs cannot ask them follow-up questions and the videos they evaluate are pre-recorded. Consequently, the standardisation and structuring of the process, although beneficial, may feel rigid and prevent the applicant from increasing their scores (Torres and Gregory, 2018; Dafoulas et al., 2002). Finally, a general poverty of communication and lack of nonverbal clues (as mentioned in Sub-sections 2.1 and 4.2) and resistance of the HR staff in adopting new procedures constitute

additional problems that organisations must bear in mind when adopting technology-mediated interviews (Dafoulas et al., 2002; Allam, J. S.,2021; Lukacik et al., *in press*).

4.4 How to improve mediated interviews

This final sub-section presents suggestions advanced by researchers that address the last research question *6) Do authors propose solutions to solve issues of mediated job interviews?*. In fact, the authors that explored TEL, VC, AVI and robot-mediated interviews often proposed tips or solutions that organisations can adopt to ameliorate the issues of these interview modalities. Some of these solutions are on the extreme-side of the continuum, consisting in simply doing less VC interviews and AVIs and more paper-and-pencil tests (Fiechter et al., 2018) or in pairing AVIs with other assessment measures (Torres and Mejia, 2017). Blacksmith et al. (2016), instead, suggests an approach that helps in tailoring the interview to the specific applicant, i.e., to consider how interview outcomes are influenced by individual differences and perceptions of interviewees. Others propose solutions aimed at modifying the problematic aspects of technology-mediated interviews, or at implementing policies or rules that help to avoid misperceptions and biases. For instance, to avoid the Background bias and distortions linked to audio-video quality in AVIs, organisations could incorporate a higher level of interview structure (e.g., consistent instructions/recommendations to applicants about recording), use anchored rating scales, give clear instructions or training to recruiters, use multiple raters and make applicant assessment based only on audio recordings (Lukacik et al., *in press*). Other solutions applicable to AVIs are to provide applicants with information about the interview procedure (in order to improve their perceptions and reactions) and use videos to present the organisation and the future job to applicants (Langer et al, 2017). This last suggestion, however, may become problematic if the management uses such videos to manage their impression or the image of the company in a deceiving way (Lukacik et al., *in press*). With regard to objective assessment, Dafoulas et

al. (2002) suggest an interesting model of interview structure based on three pillars: the *Interview module*, the *Scoring module* and the *Tracking module*. The *Interview Module* includes operations for providing the content and structure of an electronic interview questionnaire, as well as for conducting online interviews. The *Scoring Module*, then, includes operations for assigning weights to questions and grades to interviewees' answers. This module also contains the tools for calculating the applicant's scores based on his/her answers to the interview questions. Finally, the *Tracking Module* includes operations that allow the review of candidates' interviews, scores and other information about them, such as résumé and references (Dafoulas et al., 2002). **Figure 3** depicts the interview structure including the three modules just explained.

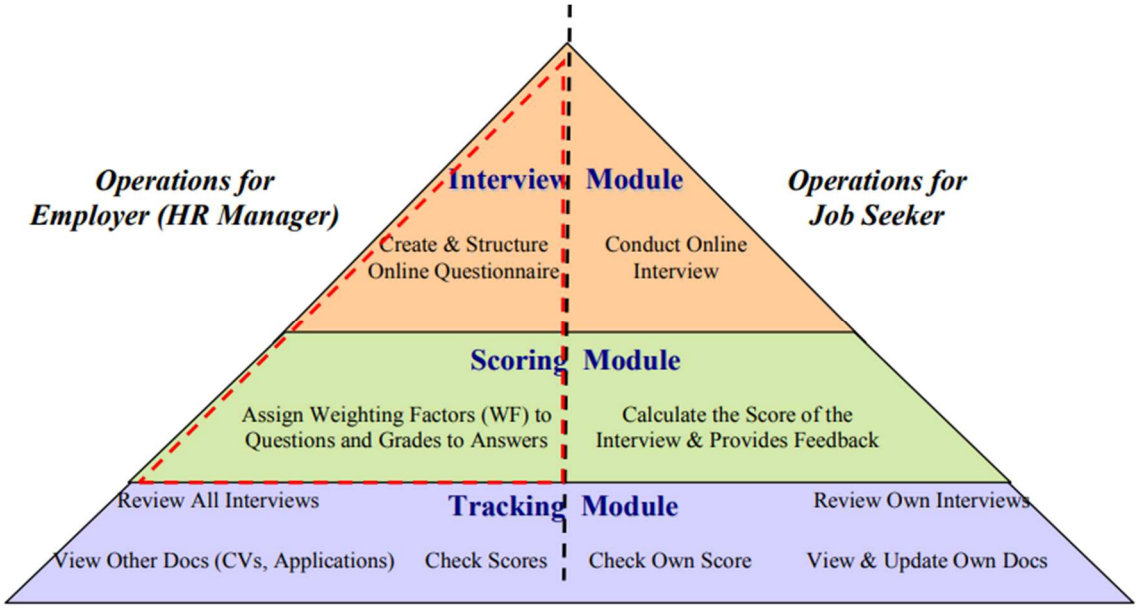

Figure 3. The model for online screening applications, based on the three pillars (Dafoulas et al., 2002).

Researchers have suggested improvements for virtual and physical avatars as well. Concerning virtual avatars, organisations should act towards a standardisation of interviewees' avatars (even at the cost of lowering the resemblance between the interviewee and their avatar), as well as better inform candidates

about the importance of their digital appearance (with the latter being the preferred solution; Behrend et al, 2012). Finally, physical avatars (i.e., robots) should be able to transmit nonverbal communication cues, with the aim of autonomous cues generation based on the operator personality (i.e., based on the psychological characteristics of who is controlling the robot). Furthermore, the robot appearance should become customisable to resemble the individual applicant, while organisations employing such avatars should train operators before the actual interview (Bremner et al., 2017).

5. Limitations

This review presents several limitations concerning the methodology used to find sources and the nature of these sources. The research was primarily conducted on Scopus, using Google Scholar as an aid when an article of interest was absent in the former database. Although this strategy proved to be efficient in selecting the material, a more thorough inspection of Scholar and other databases could have yielded a richer article pool and, consequently, a more detailed review on the literature on biases and misperceptions in mediated job interviews. Moreover, as mentioned in Section 3, the issue of biases and misperceptions in FTF interviews were dropped and, therefore, have not been tackled in this review. This prevented the possibility of drawing meaningful connections between biases in FTF interviews and other modalities, especially for the likely transfer of biases from FTF interviews to TEL, VC, AVI and avatar-mediated ones. Despite hints on the presence of this phenomenon in Section 4 of this review, an actual and useful discussion is missing. In addition, an impending and close deadline influenced the writing of this review. It is irrefutable that more content, insights and comparisons between interview modalities could have been provided with more time available.

Another type of limitation is related to the sources examined. As Blacksmith et al. (2016) highlighted, some of the studies discussed in this review present problems related to sampling. Others present another relevant methodological issue, as they relied on asking to participants to imagine being interviewed or interviewing an applicant, without actually involving participants in such activities. This kind of procedure is typical of vignette studies, i.e., those that employ a "carefully prepared description of a person, object or situation, representing a systematic combination of properties" (Atzmüller and Steiner, 2010, p. 128, in Nørskov et al., 2020). The problem of vignette studies is that they may reflect more the view of a third party (i.e., people who just observe or hear about technology-mediated interviews) rather than a second party (i.e., people directly affected by technology-mediated interviews). This means that participants asked to

imagine being involved in such interviews might not reflect reactions and behaviours of real applicants (see Langer and Landers, 2021 for a similar situation, but with AI decisions) As a result, the generalisability of vignette studies is clearly controversial.

A further limitation of this review is that certain mediated interview modalities were ignored. For instance, Blacksmith et al. (2016) includes in their meta-analysis studies on *computer-mediation* and *interactive voice response*. However, these and the other ignored modalities are only briefly mentioned in the examined articles and, therefore, there is not enough material to produce a meaningful discussion on them. As a result, those modalities were excluded from this review, but a truly comprehensive review on technological mediation in job interviews should consider them as well.

Finally, as mentioned in Section 3, the research questions that have been used to explore the literature on the topic were actually ill-suited. In fact, there are two main problems with them. The first one consists in the initial lack of knowledge on the topic of interest, which led to devising questions as general as possible in order to catch all the relevant phenomena reported in the literature. Although those questions may have succeeded in this, they could certainly have been more specific and, therefore, more useful in highlighting connections between psychological phenomena and in prompting interesting insights as a result. In practice, this shortcoming became evident when merging together Question 1 with Question 2 and, later, Question 4 with Question 5 proved to be necessary. The second problem consists in the strong theoretical bias of the research process. In fact, given that the focus of the research were biases and misperceptions in mediated interviews, the studies found during the research process dealt predominantly with biases and misperceptions as a result. Actually, only one of these studies suggested mediated interviews as preferable, while others did not explicitly express a stance on the issue. Therefore, this review undeniably ignored several articles providing a positive view on mediated interviews and might have provided an overly-negative perspective about them. Furthermore, it is highly probable that acknowledging these articles and including them in this review would

have solved the problem inherent in question 4, i.e., the near absence of articles suggesting mediated interviews as preferable.

6. Discussion and Conclusion

Research on biases and misperceptions in technology-mediated job interviews has primarily focused on the use of telephones, videoconferences and asynchronous virtual interviews. Moreover, in these recent years, researchers have been publishing studies on virtual and physical avatars, as well as on the use of AI as evaluators of applicants' performance. This review introduces some of the biases researchers found in those AI and thoroughly examines the biases and misperceptions associated to the use of the other mediums (TEL, VC, AVIs and avatars) in job interviews. Overall, technological mediation in job interviews is problematic because it maintains some biases of face-to-face interviews, give rise to new ones and applicants form negative perceptions of it. For instance, biases related to applicants' physical attractiveness affect also digital avatars, low audio/video quality affects the chances of applicants being employed and individuals perceive lower fairness in adopting mediated modalities (Behrend et al. 2012; Fiechter et al., 2018; Torres and Mejia, 2017; Lukacik et al., *in press*; Nørskov et al, 2020). Section 2 of this review introduced six research questions: 1) *Is the presence of bias in mediated job interviews supported?* 2) *Are there statistically significant distortions arising from the use of technology?* 3) *What are the psychological and/or physical phenomena reported as exclusive to mediated job interviews?* 4) *Are mediated job interviews suggested as preferable in the study under consideration? If yes/no, why?* 5) *What are the reported pros and cons of mediated job interviews?* 6) *Do authors propose solutions to solve the issues of mediated job interviews?*

With regard to questions *1) Is the presence of bias in mediated job interviews supported?* and *2) Are there statistically significant distortions arising from the use of technology?*, it is possible to conclude that biases exist, at least potentially, in all the technology-mediated modalities examined (TEL, VC, AVI and avatars) and that numerous reports of them are present in the literature. In general, despite individual studies suggesting that mediated interviews inflate the ratings of applicants, reviews and recent evidence highlight that applicants receive lower ratings in mediated interviews. In fact, lack of nonverbal cues, bad audio/video quality, lags and non-neutral interview locations hinder interviewers in performing

objective assessments of applicants. Moreover, the appearance of avatars is another source of bias, as the characteristics of avatars merge with or override those of applicants.

With regard to question *3) What are the psychological and/ or physical phenomena reported as exclusive to mediated job interviews?*, the main psychological effect that mediation prompted in interviewers and applicants are negative perceptions. Applicants were particularly concerned about privacy and fairness, with the latter varying across interview modalities, but being ultimately lower for mediated interviews compared to FTF ones. Moreover, avatars accentuate biases of FTF interviews and can appear "creepy" to applicants. However, a minority of individuals reported positive perceptions of mediated interviews, such as interviewers emphasising an easier assessment of applicants.

With regard to questions *4) Are mediated job interviews suggested as preferable in the study under consideration? If yes/no, why?* and *5) What are the reported pros and cons of mediated job interviews?*, a group of researchers advocates for written asynchronous questionnaires, while others do not express a firm opinion, not even for other forms of technological mediation in interviews. Nevertheless, technological mediation brings about simpler and more accessible interviews for applicants and recruiters, with AVIs increasing the level of standardisation. In spite of this, technological mediation has its downsides, mainly difficulties in interaction between interviewer and applicant, as well as a rigid and impersonal process.

With regard to question *6) Do authors propose solutions to solve the issues of mediated job interviews?*, researchers essentially suggest to do less interviews, pair them with other forms of assessment, standardise interviews more, better inform applicants and make avatars able to transmit more characteristics of their operators, such as nonverbal cues.

Given that the literature evidences several problems affecting technological mediation in job interviews, organisations and recruiters acknowledging these drawbacks may feel discouraged in employing mediated interviews and, as a result, stick to FTF ones. However, it is reasonable to claim that following the advices reported in Sub-section 4.4 (i.e., how to solve the problems of mediated

interviews) and providing a standardised interview procedure is enough to avoid discrimination or biased evaluations of applicants. In fact, as suggested by one of the examined articles, standardisation of the interview modality is of paramount importance. As a result, if an organisation starts to employ AVIs or another modality to screen a set of applicants, they should screen all those applicants with that modality (Langer et al., 2017). Conversely, letting applicants choose their preferred modality, although a good idea at a first glance, can lead to potentially disastrous consequences, because applicants screened with an FTF interview would probably receive higher scores than those screened with another modality. Moreover, this choice would open an enormous space for legal actions against the organisation who gave applicants such freedom, as it can be considered a form of disparate treatment. One a side note, if standardisation of the interview procedure is a primary objective an organisation is striving for, the adoption of some mediated interview modalities becomes controversial. In fact, if the increased tendency to make personal attributions in the TEL modality actually generalises to real-life contexts, HR managers and recruiters adopting such modality should inevitably adjust their rating scales in order to balance this phenomenon, otherwise they risk inflating the ratings of applicants, which then could become less predictive of actual job performance. A much “rougher” solution would be to drop the TEL modality altogether in favour of another one; indeed, the lower amount of research produced on TEL seems to suggest that this already happened or is happening in organisations, but it is not clear if this is due to the same shortcomings highlighted in this review or for other reasons. According to Blacksmith et al.’s (2016) meta-analysis, it is possible to advance a similar argument also for ratings in VC interviews, as applicants received lower scores in VC compared to FTF interviews and, as a result, HR managers and recruiters should adapt their scales as well when adopting VC. However, it must be mentioned again that this meta-analysis dealt with articles from 2007 and backwards, therefore AVIs are absent and the abovementioned finding does not automatically transfer to them (although it could not be excluded that Blacksmith and colleagues included this modality within the label “computer-mediated” or “interactive voice response”).

Regarding perceptions, the literature shows that individuals involved in mediated interviews generally develop negative perceptions of such, irrespective of the particular hardware or software used, although applicants and/or interviewers praised some aspects of VC interviews and AVIs. An interesting phenomenon highlighted by some researchers consists in these negative perceptions being higher when organisations start to adopt a new interview modality, while they become lower after that modality consolidates in recruiting practices. More specifically, this seems to have happened for VC interviews, although enough evidence to claim that it happened for sure is lacking. Therefore, it is possible to argue that negative perceptions applicants and interviewers have about mediating technologies will eventually lessen with time, as the public becomes more accustomed to such modalities. This could be of great help to researchers and organisations alike, as negative perceptions can effectively become an obstacle for the implementation of new technologies in the workplace, as for the case of expert systems in the '90s (Langer and Landers, 2021). As a result, documenting (or not) the existence of this phenomenon constitutes an interesting issue for future research. On the other side, it is obvious that applicants and interviewers will not refrain from forming negative perceptions about technological mediation because researchers still have to collect enough evidence on that. Therefore, the problem is still relevant and organisations need to address it as soon as possible. It is perhaps possible to ameliorate this issue by providing more information to participants about how the application process and the technological tools used in it actually work, as Langer et al. (2017) suggest. However, two problems affect this possible solution, i.e., organisations can provide this information in a deceptive way and providing additional information can actually backfire by inducing further negative perceptions in applicants. (Lukacic et al., *in press*; Langer et al., 2020). As reported by Langer et al. (2020) for information on how automated selection procedures work, an *indirect* positive effect on overall organisational attractiveness was counterbalanced by a *direct* negative effect of information on overall organisational attractiveness. That is, participants posing as applicants perceived the information on how the AI worked as both positive and negative. Moreover, such information was not able to decrease feeling of *creepiness* and

privacy concern and, most importantly, participants interpreted information in different ways despite the standardised modality in which they received that information (i.e., the wording and format were the same for all of them). In fact, the group of computer science students was the one for which information had the greatest effect and not the other group as hoped, which was composed by people who had less acquaintance with technological tools and informatics (Langer et al., 2020). This phenomenon is of great concern to companies, especially if observed in other studies too, as information is usually provided to aid those who are not knowledgeable about a procedure so to lower their possible biases and misconceptions. As a result, researchers could (and should) focus on this issue with the aim of finding an effective way to provide information to these people. Alternatively, research could explore other ways to lower unjustified perceptions and biases in applicants, maybe by allowing them to play with a simplified version of the recruiting AI in a game-like setting.

In synthesis, the evidence collected in this review clearly illustrates that interview modalities based on TEL, VC, AVI and avatars have important issues that organisations should keep in mind when adopting them. However, it would be foolish to stop using such modalities for the aforementioned reasons and adopt a conservative stance, because the advantages are manifold and, actually, those modalities are already becoming outdated. In fact, AI is the next frontier in the area of job interviews and there already are companies selling their algorithms to recruiters looking for more economical and rapid selection processes. In this sense, the role of researchers is not to blindly discourage the use of technological mediation in job interviews, as there are no flawless interview modalities, but to help organisations, applicants and stakeholders through their research. In fact, the issues highlighted in Sections 2 and 4 provide useful insights that researchers in psychology, informatics and other fields can use to formulate research questions relevant from both a theoretical and a practical point of view. Indeed, advancements of knowledge in the area of mediated of job interviews is, in my opinion, fundamental in devising a set of best practices that will secure a fair selection procedure for all kinds of applicants. Therefore, although it may sound clichéd, more research is actually needed.

7. Bibliography

Allam, J. S. (2021). Rebuttal from dr allam. *Chest, 160*(4), 1190-1191. doi:10.1016/j.chest.2021.05.022

APA Dictionary of Psychology (2022a, January 30): Definition of bias. https://dictionary.apa.org/bias

APA Dictionary of Psychology (2022b, January 28): Definition of medium. https://dictionary.apa.org/medium

APA Dictionary of Psychology (2022c, January 28): Definition of impression management. https://dictionary.apa.org/impression-management

APA Dictionary of Psychology (2022d, January 28): Definition of halo effect. https://dictionary.apa.org/halo-effect

* Atzmüller, C., and Steiner, P. M. (2010). Experimental vignette studies in survey research. Methodology 6, 128–138. doi: 10.1027/1614-2241/a000014

Behrend, T., Toaddy, S., Thompson, L. F., & Sharek, D. J. (2012). The effects of avatar appearance on interviewer ratings in virtual employment interviews. *Computers in Human Behavior, 28*(6), 2128-2133. doi:10.1016/j.chb.2012.06.017

Blacksmith, N., Willford, J. C., & Behrend, T. S. (2016). Technology in the employment interview: A meta-analysis and future research agenda. *Personnel Assessment and Decisions*, *2*(1), 2.

Bremner, P. A., Celiktutan, O., & Gunes, H. (2017). Personality perception of robot avatar teleoperators in solo and dyadic tasks. *Frontiers Robotics AI, 4*(MAY) doi:10.3389/frobt.2017.00016

* Brenner, F. S., Ortner, T. M., & Fay, D. (2016). Asynchronous video interviewing as a new technology in personnel selection: The applicant's point of view. Frontiers in Psychology, 7(863). https://doi.org/10.3389/fpsyg.2016.00863.

Cambridge Dictionary, (2022, February 07): Definition of uncanny valley. https://dictionary.cambridge.org/dictionary/english/uncanny-valley

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. doi:10.1348/096317901167361

Commissaris, C. (2021). Is my mic working? audiovisual quality biases in the virtual residency interview. *Academic Medicine,* , 615-616. doi:10.1097/ACM.0000000000003979

Dafoulas, G. A., Pateli, A. G., & Turega, M. (2002). Business-to-employee cooperation support through online job interviews. Paper presented at the *Proceedings - International Workshop on Database and Expert Systems Applications, DEXA, , 2002-January* 286-292. doi:10.1109/DEXA.2002.1045912 Retrieved from www.scopus.com

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) doi:10.1186/s41235-018-0139-y

Jáuregui, D. A. G., Giraud, T., Isableu, B., & Martin, J. -. (2021). Design and evaluation of postural interactions between users and a listening virtual agent during a simulated job interview. *Computer Animation and Virtual Worlds,* doi:10.1002/cav.2029

Köchling, A., Riazy, S., Wehner, M. C., & Simbeck, K. (2021). Highly accurate, but still discriminatory: A fairness evaluation of algorithmic video analysis in the recruitment context. *Business and Information Systems Engineering, 63*(1), 39-54. doi:10.1007/s12599-020-00673-w

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. doi:10.1111/ijsa.12191

Langer, M., König, C. J., & Fitili, A. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection. *Computers in Human Behavior*, *81*, 19-30.

Langer, M., & Landers, R. N. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. *Computers in Human Behavior, 123* doi:10.1016/j.chb.2021.106878

Lukacik, E. -., Bourdage, J. S., & Roulin, N. (in press). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review,* doi:10.1016/j.hrmr.2020.100789

Nørskov, S., Damholdt, M. F., Ulhøi, J. P., Jensen, M. B., Ess, C., & Seibt, J. (2020). Applicant fairness perceptions of a robot-mediated job interview: A video vignette-based experimental survey. *Frontiers in Robotics and AI, 7* doi:10.3389/frobt.2020.586263

Silvester, J., & Anderson, N. (2003). Technology and discourse: A comparison of face-to-face and telephone employment interviews. *International Journal of Selection and Assessment, 11*(2-3), 206-214. doi:10.1111/1468-2389.00244

Srinivasan, R., & Chander, A. (2021). Biases in AI systems. *Queue, 19*(2), 45-64. doi:10.1145/3466132.3466134

Straus, S. G., Miles, J. A., & Levesque, L. L. (2001). The effects of videoconference, telephone, and face-to-face media on interviewer and applicant judgments in employment interviews. *Journal of Management, 27*(3), 363-381. doi:10.1016/S0149-2063(01)00096-4

* Toldi, N.L., 2011. Job applicants favor video interviewing in the candidate-selection process. Employ. Relat. Today 38 (3), 19–27.

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. doi:10.1016/j.ijhm.2018.03.011

Torres, E. N., & Mejia, C. (2017). Asynchronous video interviews in the hospitality industry: Considerations for virtual employee selection. *International Journal of Hospitality Management, 61*, 4-13. doi:10.1016/j.ijhm.2016.10.012

Wesche, J. S., & Sonderegger, A. (2021). Repelled at first sight? expectations and intentions of job-seekers reading about AI selection in job advertisements. *Computers in Human Behavior, 125* doi:10.1016/j.chb.2021.106931

* Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2019). Personnel selection in the digital age: A review of validity and applicant reactions, and future research challenges. European Journal of Work & Organizational Psychology, 29(1), 64–77. https://doi.org/10.1080/1359432X.2019.1681401

Zorzi, M., Testolin, A., & Stoianov, I. P. (2013). Modeling language and cognition with deep unsupervised learning: a tutorial overview. *Frontiers in psychology*, *4*, 515.

*=articles not consulted directly

8. Appendix

Table 5. Bibliographic information on the literature examined.

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Allam, J. S. (2021). Rebuttal from dr al-lam.	**Review**	X
Behrend, T. (2012). The effects of avatar appearance on interviewer ratings in vir-tual employment interviews.	**Empirical study** 1 x 2 x 2 experimental between-subjects design: Perceived av-atar attractiveness x Applicant gender (*M vs. F*) x Job gender type (*stereotypically male vs. stereotypically female*)	No live interview, but a text-based one. Avatars were created using Haptek's PeoplePutty software.
Blacksmith, N. (2016). Technology in the employment interview: A meta-analysis and future research agenda.	**Review** Meta-analysis	Telephone; Videoconference; Computer mediation; Interactive voice response.

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Bremner, P. A. (2017). Personality perception of robot avatar teleoperators in solo and dyadic tasks.	**Empirical studies** SOLO TASK STUDY 3 x 3 experimental design: Tasks (*hobby vs. story vs. mime*) x Communication Condition (*Audio Only vs. AudioVisual vs. TeleOperation*) DYADIC TASK STUDY 3 x 2 experimental design: Tasks (*informative vs. competitive vs. cooperative*) x Communication Condition (*AV vs. TO*)	NAO robot (teleoperated with motion-capture technology); RGB video cameras to record participants and/or interviewers.
Chapman, D. S. (2001). The impact of videoconference technology, interview structure, and interviewer gender on interviewer evaluations in the employment interview: A field experiment.	**Empirical study** 2 x 2 x 3 mixed model: Interview Medium (*FTF vs. videoconference*) x Interviewer Gender (*M vs. F*) x Amount of Interview Structure (*high structure vs. semi-structured vs. low structure*)	Intel videoconference demonstration system (15-inch SVGA colour monitors, full-screen frame-rate from 12 to 14 frames per second).
Commissaris, C. (2021). Is my mic working? audiovisual quality biases in the virtual residency interview.	**Review**	X

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Dafoulas, G. A. (2002). Business-to-employee cooperation support through online job interviews.	**Review**	X
Fiechter, J. L. (2018). Audiovisual quality impacts assessments of job candidates in video interviews: Evidence for an AV quality bias.	**Empirical studies** Two identical experimental designs: within-subject design with Fluency (*high-quality AV/fluent vs. low-quality AV/disfluent*) as independent variable. The only difference between the experiments is the preliminary warning to discount video quality in assessment during Exp.2	Asynchronous videos edited with Final Cut Pro X to introduce distortions.
Jáuregui, D. A. G. (2021). Design and evaluation of postural interactions between users and a listening virtual agent during a simulated job interview.	**Empirical study** 2x2 between-subject experimental design: Movement (*mirror vs. random*) x Gender (*M vs. F*).	X
Köchling, A. (2021). Highly accurate, but still discriminatory: A fairness evaluation of algorithmic video analysis in the recruitment context.	**Empirical studies** 2 x 3 experimental design: Algorithm1 (*BU-NKU algorithm predictions vs. ROCHCI algorithm predictions*) x Ethnicity (*Asian vs. Caucasian vs. African-American*)	No interview platform, but short clips of videos (15s, 3k video total, all in English) taken from Youtube and then scored by Amazon Mechanical Turk

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
	4 x 5 experimental design: Algorithm2 (*Training data set vs. Test set data vs. BU-NKU algorithm prediction on test set vs. ROCHCI algorithm prediction on test set*) x Ethnicity and Gender (*Asian vs. Caucasian vs. African-American vs. male vs. female*) Experimental design with Algorithm 1 as independent variable (*BU-NKU algorithm predictions on test set vs. ROCHCI algorithm predictions on test set*)	workers on the Big Five personality traits and on an "invite for an interview" dimension (referred to as "job interview score").
Langer, M. (2017). Examining digital interviews for personnel selection: Applicant reactions and interviewer ratings.	**Empirical study** Experimental between-subjects design with Interview modality as independent variable (*videoconference vs. digital*)	Clipchamp (digital interviews); Skype (for videoconference interviews, recorded with ActivePresenter).
Langer, M. (2018). Information as a double-edged sword: The role of computer experience and information on applicant reactions towards novel technologies for personnel selection.	**Empirical study** 2 x 2 quasi-experimental design: Educational background (*computer science students vs. non-computer science students*) x Level of information (*high information vs. low information*)	No platform specified. Screen recording of a female applicant (present only through voice and through a body analysis skeleton on the left side of the screen) being interviewed by a female virtual character.

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Langer, M. (2021). The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers.	**Review**	X
Lukacik, E. (in press). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews.	**Review**	X
Nørskov, S. (2020). Applicant fairness perceptions of a robot-mediated job interview: A video vignette-based experimental survey.	**Empirical study** 2x2 between-subject experimental design: Personality trait (*introvert vs. extrovert personality*) x Interview modality (*FTF vs. robot-mediated*)	No interview platform (vignette study). The robot is Telenoid R1, a teleoperated android robot displaying a minimal human design. It mimics the operator's head movement, emulates the operator's lip movements, and transmits the operator's speech.

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Silvester, J. (2003). Technology and discourse: A comparison of face-to-face and telephone employment interviews.	**Empirical study** 2 x 2 field study: Interview modality (*FTF vs. telephone*) x Applicant gender (*M vs.F*)	Telephone
Srinivasan, R. (2021). Biases in AI systems.	**Review**	X

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Straus, S. G. (2001). The effects of videoconference, telephone, and face-to-face media on interviewer and applicant judgments in employment interviews.	**Empirical study** WITH INTERVIEWERS 3 x 2 mixed-model: Medium (*FTF vs. TEL vs. VC*) x Trial (*first applicant trial vs. second applicant trial*). Each applicant had a compulsory FTF interview and then an interview through TEL or VC, assigned through randomisation. The model is mixed because it contains deliberate manipulation of the independent variable (experimental elements) but also constraints due to the setting, such as having a compulsory FTF interview for each participant (field-study elements). WITH INDEPENDENT RATERS 2 x 2 mixed-model: Medium (*VC vs. FTF*) x Attractiveness (*above median attractiveness vs. below median attractiveness*)	Telephone (TEL) interviews: standard office telephones Videoconference (VC) interviews: PictureTel equipment (48" screen)

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
Torres, E. N. (2018). Hiring manager's evaluations of asynchronous video interviews: The role of candidate competencies, aesthetics, and resume placement.	**Empirical study** Experimental between-subjects design with Presentation order as independent variable: (*AVI first vs. resume first*)	InterviewStream
Torres, E. N. (2017). Asynchronous video interviews in the hospitality industry: Considerations for virtual employee selection.	**Review**	X
Wesche, J. S. (2021). Repelled at first sight? expectations and intentions of job-seekers reading about AI selection in job advertisements.	**Empirical studies** In all 3 studies, information about the use of AI-based technology in the applicant screening (Studies 1,2, and 3) and in the applicant interview stage (Studies 2 and 3) of the selection procedure was manipulated experimentally. STUDY 1: Participants were instructed to imagine that they were seeking a new job and rate six job advertisements.	X

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
	Online experiment with a 3 x 2 within-subjects design: Information on the automation use in the screening stage (*non-automated vs. automated vs. no information*) x Information on employee benefits (*information vs. no information*). STUDY 2 Participants had to imagine that they were seeking a job and rated three comparable but differently worded job advertisements (containing also a description of the selection procedure). Within-subjects design with Level of information about AI functioning in a personnel selection as independent variable: *non-automated (human screening and human interview*) *vs. semi-automated (automated screening and human interview) vs. fully-automated* (*automated screening and automated interview).* STUDY 3 Same procedure and descriptions as STUDY 2, but each participant rated only one job ad.	

Study (Bibliographic info)	Method used (review vs. empirical study)	Interview platform
	Between-subject design: Level of information of AI functioning in a personnel selection context (*non-automated - human screening and human interview*) vs. (*semi-automated - automated screening and human interview*) vs. (*fully-automated - automated screening and automated interview*).	