Perceived interviewee anxiety and performance in telephone interviews

Debora Jeske
University College Cork, Ireland
Edinburgh Napier University, UK

Kenneth S. Shultz
California State University, San Bernardino, USA

Sarah Owen
Northumbria University, UK

Corresponding author:

Debora Jeske
School of Applied Psychology
University College Cork
Cork T23 K208
Republic of Ireland
Email: d.jeske@ucc.ie

This paper was accepted by the editor of "Evidence-based HRM: A global forum for empirical scholarship" on the 26th of July 2018.

Future DOI: 10.1108/EBHRM-05-2018-0033

Funding statement

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Perceived interviewee anxiety and performance in telephone interviews

Purpose: The current research focused on the role of interviewee anxiety as a predictor of perceived hireability (Study 1, $N = 82$) and job suitability (Study 2, $N = 74$).

Design: Using an experimental design, participants were randomly allocated to one of two conditions (an audio recording of either a confident or anxious job candidate with identical scripts) and asked to take the role of an interviewer.

Findings: The anxious interviewee (played by an actor) was consistently rated as less hireable (in a combined sample based on Study 1 and Study 2), less suitable to the job and received less favorable hiring recommendations (as assessed in Study 2) than the confident interviewee (played by the same actor).

Limitations: The study was conducted with students who may have less interview experience than experienced interviewers.

Practical implications: The results suggest that anxiety has a negative biasing effect on perceived hireability and job suitability ratings. In other words, the behavioral manipulation of anxiety affects hireability ratings, independent of any subjective assessment of anxiety.

Originality: The findings provide evidence of an anxiety bias in telephone interview settings. The results highlight the importance of considering anxiety cues when training employment interviewers.

Keywords: anxiety; hireability; job suitability; telephone interviews; verbal cues

Introduction

Interviewing is a ubiquitous part of the employment recruitment process and is still one of the most common forms of human resource selection. Thus, one of the main aims of an employment interview is to gain an accurate assessment of candidates and whether they would be a good fit for the vacancy and the organization. Therefore, it is important to understand the factors which may influence, and potentially bias, an interviewer's perceptions of job candidates. One such variable is perceived interviewee anxiety (see Huffcutt, Van Iddekinge and Roth, 2011; McCarthy and Goffin, 2004). In the context of telephone interviews, such anxiety may have a number of detrimental effects for how interviewees are evaluated in terms of their characteristics and likely performance.

To date, most previous research investigating performance and anxiety in interviews has only focused on the face-to-face method (McCarthy and Goffin, 2004; Sieverding, 2009; Feiler and Powell, 2013). None of the research specifically examined perceived hireability and job suitability in relation to anxiety. Given the increasing prominence of such interviews (e.g., 38% according to a study by the Chartered Institute of Personnel and Development in 2009), it is important to gain a better insight into the potentially biasing effects of anxiety on perceived interviewee performance – in the absence of visual cues.

In this article, we argue that the way in which this anxiety is exhibited via verbal cues by a candidate in telephone interview settings may be potentially linked to performance inferences on the one hand, and therefore potential rating error on the other. We therefore propose that anxiety in telephone interviews will also reduce ratings of hireability and job suitability. We briefly explore each of these suggestions.

Absence of visual cues: Anxiety effects on ratings

In telephone interviews indicators of anxiety may be incorrectly conceived as measures of assertiveness or intelligence or competence, on account of absence of visual cues. This negatively impacts conclusions regarding a candidate's hireability or job suitability negatively. Indeed, some evidence suggests that anxious interviewees are seen as less assertive (Feiler and Powell, 2016). The link to performance is also present. A number of studies provide some insight on how anxiety impacts interview performance (e.g., Feiler and Powell, 2013, 2016; Blacksmith et al., 2016; Poh, 2015). Anxiety has been shown to negatively affect performance in face-to-face interviews among students (e.g., Feiler and Powell, 2013; McCarthy and Goffin, 2004). Interviewers in telephone interview settings will use verbal indicators to form impressions about a candidate's characteristics. These impressions may be based, in the absence of visual information, on interviewees' vocal fluency, voice intensity and pitch, energy, affect, emotional expressiveness, and voice modulation (DeGroot and Gooty, 2009; Frauendorfer and Schmid Mast, 2014; Riggio and Riggio, 2002). Yet many of these verbal indicators are negatively affected by anxiety (Gilboa-Schechtman and Shachar-Lavie, 2013). For example, anxiety may exhibit itself in terms of hesitations, unfinished sentences, stutters, inconsistent pitch, repeated words, use of filler words and intruding incoherent sounds – that is a noise that cannot be identified as a word, stutter or laugh (Kasl and Mahl, 1965; Pope, Siegman, and Blass, 1970; Ragsdale, 1976). The inability to express oneself clearly or answer questions smoothly are verbal indicators of anxiety (Feiler and Powel, 2016). This situation may also further increase applicant anxiety. In the absence of other informaiton, anxiety indicators such as stutters or hesitation marks may be interpreted as a sign that the candidate lacks the confidence or competence for the role. We therefore propose the following when raters encounter an anxious or confident interviewee in a telephone setting (using experimental manipulation):

Hypothesis 1: Ratings for job suitability and hireability will be lower when raters (our participants) listen to the anxious interviewee compared to the raters who listened to the confident interviewee.

Absence of visual cues: Potential for rater bias

Since in telephone interviews only 'verbal behaviors' such as fluency, hesitations, stutters, pitch, and long silences, are observed, the anxiety-driven verbal cues will weigh more heavily as there is no means to correct for these via visual cues (e.g., firm conduct, looking interviewer in the eyes, or smiling) as is the case in real-life interviews. While face-to-face interviews usually provide interviewers with visible anxiety cues, interviewers are not exposed to such indicators in telephone interviews. Not surprisingly, several authors suggest that generalizing findings across different interview types is inappropriate (Blacksmith, Willford, and Behrend, 2016; Poh, 2015).

Indeed the potential for rater bias is substantiated by other research. A number of researchers moreover noted that anxiety in interview settings may operate as a an extraneous variable that affects test scores and may thus make it harder to discern the interviewee's actual ability (see Arvey et al., 1990; Ryan and Ployhart, 2000). Due to the lack of visual feedback, interviewers are unable to see and correct for anxiety cues (such as blushing, see Bögels et al., 2010) and may therefore make inferences about the ability of the candidate to perform on the job without being able to take into account their level of anxiety during the interview. The lack of visual feedback or confirmation may contribute to this impression among raters (Cavanaugh et al., 2000). By linking anxiety to potential competence, raters will also generate lower performance ratings. As a result, we hypothesize:

Hypothesis 2: Perceived interviewee anxiety (as reported by the rater) is a negative predictor of their ratings (job suitability and hireability), even controlling for which candidate (anxious or confident) they evaluate.

Contribution

Research on virtual interviews has looked at nonverbal behavior (Langer et al., 2016) and nonverbal cues in relation to verbal cues (Kwon et al., 2013; see also face-to-face interviews in Feiler and Powell, 2016). However, no research has specifically examined verbal anxiety cues in telephone interviews. It is this research gap that we focus on in our study. The next section outlines what we know to date before proceeding to our research rationale: To examine anxiety effects in the context of telephone interviews. Finding answers to our two hypotheses is important for two reasons. First, while telephone interviews are quite established today, most of the research on anxiety is still based on face-to-face research. Second, the perceived anxiety of interviewees may close doors long-term. Competent but anxious interviewees may also be less likely to be invited to a second interview (Cook, Vance, and Spector, 2000). Using an experimental design involving a simulated interview task, we therefore aimed to provide answers to these two questions about how anxiety (experimentally manipulated and perceived by raters) effects hireability and job suitability ratings in telephone interviews.

Materials and Methods

Design

The study featured a between-subjects groups design with two conditions. In the experimental condition, participants listened to a telephone interview recording with an anxious job candidate. In the control condition, participants listened to the same interview recording but with a confident

job candidate (in both Study 1 and 2). These interviews were based on one and the same interview (transcribed from an actual interview with a computer science professional) that was read by two actors (the same two for each of the interviews, one taking the role of the interviewer and one taking the role of the interviewee). The decision to use professional and experienced actors was based on ethical concerns about recording of a particularly anxious interviewee and the issue of poor comparability if we used another and more confident interviewee. The script also referred to a number of programs that the job candidate was expected to use and know in the new job. The script for both conditions was identical across both conditions. This decision was taken to control for narrative effects (Borges et al., 2015) and very similar to the script procedures used for video and telephone interviews in Poh (2015). The male actor playing the interviewee was instructed to include a number of anxiety cues while reading the interview script in the experimental condition. The anxiety cues were taken from the literature on anxiety cues and included coughs, filler sounds (e.g., 'ah', 'er', and 'um'), brief hesitations, stutters, and voice changes as reported in the literature (Kasl and Mahl, 1965; Pope, Siegman, and Blass, 1970; Ragsdale, 1976). The equipment used to create the recordings was Pro Tools recording Software and a C24 Mixing Desk at a university recording studio under the supervision of a qualified sound technician. Our design therefore included measures of both actual (experimentally manipulated) anxiety and perceived interviewee anxiety (as rated by the rater = participant). All materials were included in the format of an online survey (one including the confident and the other the anxious interviewee recording).

Participants

For Study 1, 94 participants were recruited to act as raters of the two recorded telephone interviews. Of these, only 82 provided complete data used in the analysis (this was assessed based on the time they spent to complete the study and the number of sections they had completed).

There were 41 participants in each condition. Data collection was closely monitored to ensure the number of participants in each condition was equal. There was an overall age range of 18 to 49 years of age (M = 22.08, SD = 6.54). The final sample included 14 men and 68 women. Around half of the sample were also working at the time of the survey (n = 44, 53.7%).

For Study 2, 88 participants were recruited. Following the exclusion of 14 reports due to missing values and drop-outs, the final sample included the data from 74 participants (with 37 participants in each interview condition). Participants were between 18 and 53 years old (M = 25.44, SD = 7.19). The sample included 22 men and 49 women (3 missing values). Just under two third of the sample was working at the time they participated (64.9%).

Procedure

The participants (raters) in Study 1 were recruited through an opportunity sample via email and the University participant pool of social science students in exchange for course credit (UK sample). The participants in Study 2 were recruited through email and in-class announcements by instructors and included a mix of business and social science students (combined UK and US sample). Participants in Study 2 received a small incentive for participating in the research (£3 in the UK and $5 in the US). All participation was voluntary. The only exclusion criterion was that they must be over 18. Participants who indicated that they had been diagnosed with an anxiety disorder in the past were excluded, as such history may impact their ratings (i.e., either increasing their sensitivity to such cues or potentially raising their own anxiety while completing the study). Both studies received ethical approval from all participating universities.

As soon as participants read the participant information sheet, they were asked to give consent and a code word (in case they wish to withdraw data within a specific timeframe). Allocation to interview condition was random, although attention was paid to ensure similar numbers were

recruited at the end of the data collection effort. All participants were asked to act as raters and evaluate the interviewee on the audio recording in terms of their hireability and job suitability. We use the term 'rater' here as the participants did not actually interact with the interviewee. While listening to the interview, participants were asked to identify all those software programs which were mentioned in the telephone interview and asked to complete a series of questionnaires on hireability (Study 1), as well as job suitability and hiring recommendation (Study 2). Other measures followed, including perceived anxiety of the interviewee, participants' anxiety and interview experience, followed by demographic questions. The debriefing statement concluded the survey.

Measures

The studies included several measures, some of which were shortened to reduce cognitive load due to the length of the study and recording (25 minutes). Study 2 featured a number of additional measures (indicated below). The scale characteristics and reliability of all measures for both Study 1 and Study 2 are provided in Table 1.

Attention check. This measure was created to ensure participants were attentively listening to the audio, this included a list of programs related to the role of a web developer (Asp.NetMVC, C Sharp, HTML, Sequel, JavaScript, LINQ, XML, AJAX, jQuery and CSS). Participants had to identify which of these programs were mentioned in the telephone interview. The right answers were: ASP.NetMVC, C#, JavaScript, HTML, CSS, and SQL.

Hireability questionnaire. Hireability was based on eight of 16 items of the hireability measure developed by Howard and Ferris (1996). The items that we retained were those which related to perceived competence of the applicant and job suitability. The items not included were those pertaining to perceived similarity and affect towards applicants as they were not considered

to relate to hireability per se. We added "in web development" to several questions to contextualize the items to the position being advertised. An example item is: "This applicant appears to have a good understanding of web development." The response range was 1= *Totally disagree* and 7 = *Totally agree*.

Job suitability (added in Study 2). Job suitability was measured using eight items taken from Krueger (2011) and based on measures used by Hosoda, Stone and Stone-Romero (2003), Stone and Stone (1987). An example item is: "This job applicant is well suited for the job." The response options ranged from 1 = *Strongly disagree* to 7 = *Strongly agree*.

Hiring recommendation (added in Study 2). The assessment of hiring recommendation involved 7 items used by Krueger (2011), based on the measures by Hosoda *et al*. (2003). An example item is: "I highly recommend that this job applicant be hired for the job." The response options were 1 = *Strongly disagree* and 7 = *Strongly agree*.

Anxiety (interviewee). Perceived interviewee anxiety was used to gain an impression as to how anxious our raters viewed the interviewee to be (as portrayed by the actor playing the interviewee), keeping in mind that the participants were not aware of there being two conditions. In order to assess participant's perception of these impressions, we utilized 12 modified items from the Measure of Anxiety in Selection Interview (MASI) including twelve of the 16-item scale by McCarthy and Goffin (2004). The content of the items was slightly changed to refer to 'the interviewee' whose interview participants listened to. The four excluded items all referenced thoughts or comparisons to others that the participants had no means of assessing, given that they only listened to, but did not interact with the interviewee. An example item is: "The interviewee seemed very nervous about whether or not they would be a good candidate for the job." The items were scored on a five point scale where 1 = *Strongly disagree* and 5 = *Strongly agree*.

Control variables. Four characteristics of our participants are considered as potential control variables. This included participant's own anxiety in interviews, their previous interview experience, age and gender. The rationale and form of measurement are outlined next.

First, research has shown that more anxious individuals believe themselves to be less likeable and competent (Voncken, Dijk, de Jong, and Roelofs, 2010) and tend to self-evaluate themselves more negatively (see Brozovich and Heimberg, 2011). In addition, anxious interviewees tend to focus on more negative information (e.g., Clerkin and Teachman, 2010), which may drive bias in terms of how they evaluate others. We used 16 items from the Measure of Anxiety in Selection Interview (MASI) by McCarthy and Goffin (2004) to assess participant's own interview anxiety. We excluded items referencing behavioral anxiety and appearance as they also asked about behaviors that would not be observable in a telephone interview. An example item is: "I get so anxious while taking job interviews that I have trouble answering questions." The items were scored on a five point scale where 1 = *Strongly disagree* and 5 = *Strongly agree*.

Second, personal interviewee (or managerial) experience may influence the reliability ratings. Third, experience tends to increase with age, making age an important control variable. And fourth, gender differences tend to be frequently observed in anxiety research (see also Boyer et al., 2017; Egloff and Schmukle, 2004; Feeney, McCarthy, and Goffin, 2015; Zalta and Chambless, 2012), if not always (see Feiler and Powell, 2013). Interview experience was measured with the help of three newly developed items. The three items were as follows: (1) "Have you been interviewed in the last 6 months?"; (2) "Do you currently hold a managerial positon that requires you to conduct interviews?"; (3) "Have you interviewed anyone in the last 6 months?". All answering options included two response options: *Yes* and *No*. These items were used to code interview experience. Age and gender were part of the demographic section that was presented at the end of the survey.

Age was recorded as a numerical variable while gender included options such as 1 = *male*, 2 = *female*, and 3 = *prefer not to say*.

Results

Descriptive Statistics and Scale Characteristics

The scale descriptive statistics for all scales in the combined sample are listed in Table 1, while correlations are reported in Table 2.

As early indicators suggested very similar correlation patterns between the independent and dependent variables of interest in both studies, we combined the samples from Study 1 and Study 2. This step increased statistical power for subsequent regression analysis. And lastly, in line with the analysis run by Borges et al. (2016) to control for group differences in participant anxiety in their study, we also checked and found that participant anxiety did not differ significantly across the two conditions ($p > .05$).

Table 1: Scale properties (Study 1 and 2)

Scale		Total sample (N = 156)			Control (conf.) (n = 78; $n_{§}$ = 37)			Experimental (anx.) (n = 78; $n_{§}$ = 37)		
	Items	α	*M*	*SD*	α	*M*	*SD*	α	*M*	*SD*
Attention check	6(10)	--	4.54	1.54	-	4.36	1.60	-	4.67	1.70
Hireability (I)	8	.91	4.84	0.79	.79	5.30	0.51	.93	4.40	1.33
Anxiety (I)	12	.91	3.13	0.52	.90	2.68	0.46	.85	3.54	0.59
Anxiety (P)	16	.90	2.68	0.46	.93	2.72	0.55	.85	2.65	0.36
Job suit.(I) §	8	.94	4.64	1.41	.91	5.06	1.18	.95	4.24	1.51
Hiring rec. (I) §	7	.94	4.17	1.61	.89	4.78	1.25	.96	3.59	1.71

Note. Study 1 N = 82. Study 2 N = 74. I = Interviewee. P = Participant. Missing values were replaced with mean (maximum number of missing cases was 2). The attention check included four red herrings (only six items were

correct). Skew and kurtosis were unremarkable (below 1) for all scales. § Added to Study 2 = values for job suitability and hiring recommendation are based on $n_{§} = 37$ participants each in the control (confident) and the experimental (anxious) condition.

Table 2: Correlations between scales (combined sample)

Scale	1)	2)	3)	4)	5)	6)
1) Hireability	--					
2) Job suit. §	.83**	--				
3) Hiring recomm. §	.79**	.91**	--			
4) Interviewee anx.	-.53**	-.50**	-.53**	--		
5) Participant anx.	.11	.05	.11	.12	--	
6) Age	.04	.02	-.08	-.12	-.03	--

Note. $N = 156$. § Added to Study 2 ($N = 74$).

Experimental vs. Control Conditions (Interview Condition)

The first step was to check that our audio recordings were indeed successful in evoking different ratings of the interviewee on tape. Rather than assuming the conditions are effective, we have to demonstrate that the participants in the anxious condition also considered the interviewee more anxious than the participants who listened to the interviewee in the confident condition. As expected, the anxiety of the interviewee was rated as significantly higher in the anxious than in the confident condition in the combined sample ($t(148) = 28.57$, $p < .001$, $\eta^2 = .16$). The results were largely identical ($F(1, 145) = 29.45$, $p < .001$, $\eta^2 = .17$) when taking into account potential covariates (age, gender, the attention check, and whether participants had listened to the end of the recording; all covariates were found to be not significant at $p < .05$).

Hypothesis Testing

Interviewee anxiety (condition and measured) and job suitability/hireability (H1).

Hypothesis 1 proposed that ratings for job suitability and hireability were predicted by whether or not raters listened to the anxious or confident interviewee. The effect of interview condition was tested using a means test. Participants who rated the anxious interviewee less hireable ($t(154) = -4.40$, $p < .001$, $\eta^2 = .11$, $M = 4.59$, $SD = 1.10$) compared to the participants who rated the confident interviewee ($M = 5.27$, $SD = 0.82$). Participants also rated the anxious interviewee as less suited for the job ($t(72) = -2.56$, $p = .012$, , $\eta^2 = .08$, $M = 4.24$, $SD = 1.51$) compared to the participants who rated the confident interviewee ($M = 5.04$, $SD = 1.15$). Participants who rated the anxious interviewee were less likely to recommend hiring him ($t(72) = -3.37$, $p = .001$, , $\eta^2 = .13$, $M = 3.58$, $SD = 1.70$) compared to the participants who rated the confident interviewee ($M = 4.75$, $SD = 1.22$).

Evidence of attribution bias in hireability/job suitability ratings (H2).

In the next step, we wanted to assess if raters' perception of interviewee anxiety (reported for interviewee, not manipulated) explain ratings – above and beyond the experimental design. In theory, the manipulated variable (interview condition) should capture all the variance in ratings, unless our raters make further attributions about the anxiety of their interviewee. We considered participants' gender, age, and experience as covariates. However, these variables were not significant variable here, we excluded it from the results (as reported in Table 3).

Table 3: Regression results for job suitability and hiring recommendations

	Study 1 & 2: main outcome		Study 2 only: additional outcomes			
	Hireability		Job suitability		Hiring recommendation	
	$R^2\Delta$	β	$R^2\Delta$	β	$R^2\Delta$	β
Step 1						
Interview condition	.07***		.08*		.14**	
Step 2						
Perceived interviewee anxiety (as rated by participants)	.22***	-0.51	.16***	-0.48	.15***	-0.46
Total R^2	.28***		.24***		.29***	
N	156		74		74	

Note. Participants were not made aware that there were two interview conditions. Their ratings of the actor's perceived interview anxiety is therefore based on the impressions the participants acting as raters formed about the interviewee. Control variable included interview condition (manipulated, explained 14% in perceived interviewee anxiety alone). No other potential covariates were significant. ^{t} $p < .10$; ** $p < .01$; *** $p < .001$.

The results for the combined sample from Study 1 and 2 revealed that the perceived anxiety of the interviewee as rated by the interviewer had a significant influence on their perceived hireability (β = -0.51; $p < .001$; Table 3), even after we controlled for our manipulation variable (interview condition). Very similar results were obtained for job suitability and hiring recommendation. In the absence of rater bias, the manipulated variable should have been the only one to explain variance in ratings. While 6.8% of variance in hireability was due to the dichotomous (anxious or confident) interview condition, 21.1% of variance in ratings were due to how the raters perceived

the interviewees' anxiety. This means that 21% of 27% of variance can be considered bias on behalf of our untrained raters.

Discussion

The purpose of the present research was to examine the effect of effect of manipulating anxiety and interviewer perceptions to understand the effect of anxiety generally, and potentially interviewer perceptions of interviewee anxiety, on telephone outcomes. In our design, we considered the extent to which verbal anxiety cues could convey anxiety in the absence of visual indicators. The work thus builds on a number of recent studies on anxiety that has emphasized the role of more studies on verbal cues (Feiler and Powell, 2016), interview medium, rater bias and errors (Huffcutt et al., 2011).

Our first hypothesis proposed that ratings will be lower when raters (our participants) listen to the anxious interviewee compared to the raters who listened to the confident interviewee. Our results demonstrated a statistically significant group effect with a small to medium effect sizes. Participants who rated to the anxious interviewee were rated as lower on all three job outcomes (hireability, job suitability, and hiring recommendation), with the suitability outcome having a slightly lower effect than the two hireability outcomes. In other words, they were considered less hireable, job suitable, and less likely to receive an endorsement to be hired. This provides support for our first hypothesis of anxiety. Our results match the results of Feiler and Powell's (2013) mock interview study. In this study, the authors also observed a negative correlation between interview anxiety (as reported by the participant) and ratings. In addition, this finding provides support for McCarthy and Goffin (2004) who found that interviewee anxiety negatively predicted interviewee ratings in face-to-face interviews.

One additional reference provides more information on context effect, specifically, the Masters dissertation by Poh (2015). In this research, Poh (2015) gave her student participants two interview assignments involving both a telephone and a one-way video interview, the order of which was counter-balanced for all participants. In contrast to our findings, Poh (2015) anxiety was only negatively correlated with performance in the one-way video interview, but not the telephone interview. A number of factors may explain this, including novelty of video interviews and lack of training in this type of interview. First, Poh's (2015) as well as Feiler and Powell's (2013) raters received interview training and had prior interview experience, training that our participants did not receive (they also had limited previous experience). And second, Poh's (2015) trained raters evaluated interviewees during the interview rather than having access to a recording as our sample did. This leads us to the discussion of our second hypothesis on bias.

In the words by Boyer et al. (2017, pg. 1), "it is important for the recruiters to understand the level at which anxiety can influence selection outcomes". We took this suggestion further by considering interviewer perceptions on anxiety as a source of bias. The second hypothesis investigated the possibility of bias by also considering the role of perceived interviewee anxiety as rated by our participants and the influence of this perception above and beyond the experimental condition. In other words, ratings about a candidate may be subject to experimental condition (anxious/confident candidate, H1), perceived anxiety (rater bias, H2) and rating error. This means that unlike McCarthy and Goffin (2004), our experimental manipulation also means that we can discern a significant effect of bias in terms of how perceived anxiety might impacts ratings in an experimental setting when the information presented in the interviews is one and the same.

This meant any differences in interviewee ratings our research could only be attributed to the anxiety manipulation. And when this effect is controlled for, the differences in ratings can only be

attributed to participant rating bias when judging the interviewee's (actor's) level of anxiety. In other words, rater variance is more attributable to the rater than to the interviewee (see also Scullen, Mount, and Goff, 2000). Our results showed a strong negative effect of how anxious the interviewees were perceived on ratings, with lower ratings being given the more anxious the interviewers perceived the interviewee to be. These findings are in line with the negative association between anxiety and assertiveness, which in the context of an interview, may be to the disadvantage of the interviewee. But we should also note that our results are dependent on a number of factors, such as the managerial position of the rater. The meta-analysis by Viswesvaran, Ones and Schmidt (1996) showed that rater-specific factors are smaller with managers. Gender differences may also come into play, which is why interview training is essential for interviewers.

Contribution

Our results contribute to the existing research on interviewee anxiety in two important ways. First, we capture the effect due to interviewee (or situation-specific/actor) anxiety. Second, we examined the anxiety perceptions our raters gained from listening to the interviewee as potential source of bias. Our results suggest that anxiety perceptions may bias interviewers above and beyond what we expected based on our carefully controlled experimental design.

Our findings contribute to the growing research on observer accuracy in employee selection (Zimmerman, del Carmen Triana, and Barrick, 2010). At the same time, we need to be careful when interpreting findings. First, as our discussion of anxiety and context indicates, different forms of anxiety may come into play depending on the type of interview settings. This may also then influence the generalizability and transferability of findings. Blacksmith, Willford and (2016) propose that it is inappropriate to generalize findings from face-to-face to telephone interviews.

Poh (2015) similarly cautioned that different interview techniques may also result in varying outcomes.

Second, interviewer experience may have more or less detrimental effects in the interpretation of anxiety in telephone interviews. Context (e.g., type of interview) as well as interviewer experience may therefore jointly influence ratings. Anxiety cues (e.g., stuttering, long pauses and repeated words/phrases) may be misinterpreted as indicators of insecurity and apprehension (e.g., Cook *et al*., 2000), leading raters to conclude that the person may be unlikely to perform well on the job. Yet, the influence of these cues may be given less weight when less visual information is available. Maybe a combination of telephone and video interviewing could help to avoid bias due to missing anxiety cues. However, the question of whether or not our raters showed evidence of bias due to inexperience or misattributions is not easily answered. We attempt an answer next. The question then is how reliable interview ratings can be expected to be. If anxiety in interviews does predict performance on the job (and there is some evidence of this by McCarthy et al., 2013), it will be important to understand to what extent anxiety is perceived as anxiety, and heuristics interviewers employ, that may also result in rater bias.

Practical Implications

A number of practical implications arise from our findings and previous discussion. First, it is in human resource managers' best interest to ensure better observer accuracy in employee selection (Zimmerman *et al*., 2010). If video or telephone interviews are the most likely forms of interviews to be conducted in the future, interviewer training should be revised accordingly to prevent bias due to lack of experience with such interviews. Even interviewers experienced with face-to-face interviews may need training to assess candidates in other forms interview settings. One practical way to accomplish this is through the use of behavioral interviewing techniques (Barclay, 2001),

where interviewers are trained to focus their attention on key behavioral indicators, rather than irrelevant factors such as interviewee anxiety. This might require training the managers who conduct interviews in their organization to recognize anxiety cues in telephone interviews and to not let the applicant's anxiety bias their competency rating. Consequently, training to detect anxiety cues may be helpful (see work by Givens, 2014) for both face-to-face and telephone interviews, as detecting anxiety appears to be a concern for both interview methods.

The second implication relates to recent technical innovations. New software running alongside the interview may be able to identify anxiety cues better than a human interviewer, recognizing stutters, voice changes, and other indicators. Several free speech analysis software programs exist that could be modified for this purpose, while new programs are available to test voice stress and even indicators for depression (see Chu, 2009). Several recent press reports also emphasize the importance of speech analyses and voice in hiring (e.g., Heller Zaimont, 2014; Morrison, 2017). This suggest some resources already exist that may be used as a starting point for the development of such software to identify anxiety cues. Such software may also be a useful training tool for interviewer training, that is, such software could be used to teach individuals with interview anxiety how to control their voice and reduce the frequency or occurrence of anxiety indicators. From a human resource managers' perspective, such software and training may ensure better accuracy in employee selection (Zimmerman *et al*., 2010).

And third, an established review process for interview recordings and a scoring system for anxiety cues could help interviewers to retroactively (rather than in the moment) check for indicators for anxiety and adjust on their ratings accordingly. Again, software could make this particularly simple and help to assess anxiety in both contexts. Training to detect anxiety cues may be helpful when interviewers do not have access to such software. The benefit of such procedures

may extend from telephone to face-to-face interviews as well (see work by Givens, 2014) as detecting anxiety appears to be a challenge in both cases. In terms of the generalizability we ought to note that anxiety may also arise as a matter of traits and performance concerns, making it difficult to tease apart the origin of anxiety. Not all anxiety is indicative of trait anxiety or low assertive. Interviews are not the best instruments to measure ability, so additional selection tools such as a works samples test may be more effective to help differentiate anxiety and ability.

Limitations and Future Research

Several limitations apply to the present studies. First, hesitation markers, stutters and similar verbal interrupters may not be due to anxiety alone. Please note that we assessed the effect of participants needing less time than anticipated (calculated by audio clip length and average completion rate of questions). No significant effect emerged – length of time was not a significant covariate in any analyses. From a methodological perspective, our outcome measures also correlated very highly with one another. Although we used different measures, several items in each scale tapped into similar constructs such as job suitability (e.g. Howard & Ferris, 1996; Krueger, 2001). This suggests that the three different measures performed remarkably similar, although hireability was meant to focus also on perceived competence, while the hiring recommendation was meant to focus on rater's behavioral intentions. While the results suggest consistent findings, future researchers may wish to reduce the number of measures.

Second, the current research looked at telephone interviews alone, examining the findings from cross-sectional student populations. In addition, the interview involved two male actors. The decision to use male actors was based on the fact that males have been noticed to show greater impairments in job interviews as a result of interview anxiety (Feiler and Powell, 2013), a finding

we were mimicking in the anxious condition. It would be interesting to examine how well the trends are replicated when the interview involves two women instead of two men. General trait ratings often show greater bias than more concrete task (competence) ratings. More work in this area may also reduce rater bias and inform ways to improve interviewer performance.

Third, all reports were based on self-reports (Schwarz, 1999), did not differentiate the impact of specific anxiety cues as attempted by Feiler and Powell (2016), and featured a study design that included only one male candidate. In realistic settings, raters interact with interviewees and will interview multiple candidates that are male or female. This means our results may not generalize across all settings and jobs. The importance of verbal performance may also need to be weighted and considered. Computer scientists may not often be in customer-focused role, but work in designated teams. Anxiety may therefore not necessarily affect performance on the job itself. As noted above in the "Practical implications" section, seeking additional ability indicators may help recruiters identify, by comparing the performance on two selection tools, those candidates who may perform more poorly in telephone interviews due to anxiety rather than a lack of ability. If performance of interviewees improved over subsequent interviews, recruiters may also be able to narrow down whether or not anxiety stems from ability concerns or anxiety during the selection process.

In addition, we had selected a computer science job with only a certain degree of social interaction and leadership responsibility. It would be interesting to examine the extent to which anxiety in interviews for jobs requiring an increasing degree of interpersonal interactions and confidence in interaction with others (e.g., hospitality, education, or retail) is even more detrimental to how interviewers rate potential candidates. In line with Woods *et al.*'s (2013) work, we would encourage researchers to conduct interview studies with more age-diverse samples. This

may also enable researchers to examine the dynamic influence of age as well as situation strength (Bortfield *et al*., 2001).

Acknowledgements

We gratefully acknowledge the support of our two professional actors, the Department of Psychology and Phil Capewell at Northumbria University in the production of the interview materials. We would also like to thank our participants for their support of this study. Finally, we appreciate the feedback and helpful suggestions we received from Sonia Lippke, Donna Garcia and the reviewers of Evidence-based HRM.

References

Anderson, N., Lievens, F., van Dam, K., and Ryan, A.M. (2004), "Future perspectives on employee selection: Key directions for future research and practice", *Applied Psychology: An International Review,* Vol. 53 No. 4, 487-501.

Arvey, R. D., Strickland, W., Drauden, G., and Martin, C. (1990), "Motivational components of test taking", *Personnel Psychology,* Vol. 43 No. 4, 695-716.

Barclay, J.M. (2001), "Improving selection interviews with structure: Organizations use of "behavioural" interviews", *Personnel Review,* Vol. 30 No. 1, 81-101.

Blacksmith, N., Willford, J.C., and Behrend, T. S. (2016), "Technology in the Employment Interview: A Meta-Analysis and Future Research Agenda", *Personnel Assessment and Decisions,* Vol. 2 No. 1, pp. 12-20.

Bögels, S.M., Alden, L., Beidel, D.C., Clark, L.A., Pine, D.S., Stein, M.B., and Voncken, M. (2010), "Social anxiety disorder: questions and answers for the DSM-V", *Depression and Anxiety,* Vol. 27 No. 2, 168–189.

Borges, T., Ramosa, S., Vilar, E., Noriega, P., and Rebelo, F. (2015), "Interview anxiety narrative validation for a virtual reality-based study", 6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015). *Procedia Manufacturing,* Vol. *3* No. 2015, 5934 – 5940

Bortfield, H., Leon, S. D., Blom, J. E., Schober, M. F., and Brennan, S. E. (2001), "Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender", *Language and Speech,* Vol. 44 No. 2, 123-147.

Boyer, L., Carden, L., Johnson, L., and Boyd, R. (2017), "Establishing an interview anxiety baseline: assessing applicants' readiness", *Business and Professional Communication Quarterly,* Vol. 80 No. 3, pp. 365-378.

Brozovich, F., and Heimberg, R.G. (2011), "The Relationship of Post-Event Processing to Self-Evaluation of Performance in Social Anxiety", *Behavior Therapy,* Vol. 42 No. 2, 224–235.

Cavanaugh, M. A., Boswell, W. R., Roehling, M. V., and Boudreau, J. W. (2000), "An empirical examination of self-reported work stress among U.S. managers", *Journal of Applied Psychology,* Vol. 85 No. 1, 65–74.

Chu J. (2009), "Software Listens for Hints of Depression. A large-scale trial will test whether software can identify depressed patients", *MIT Technology Review.* Blog post 4 November 2009. Available at: https://www.technologyreview.com/s/416131/software-listens-for-hints-of-depression/ (accessed 21 February 2017)

CIPD (2009), "Annual survey report 2009. Recruitment, retention and turnover", Retrieved on 10/11/2015 from https://www.cipd.co.uk/binaries/recruitment_retention_turnover_annual_survey_2009.pdf

Clerkin, E.M., and Teachman, B.A. (2010), "Training Implicit Social Anxiety Associations: An Experimental Intervention", *Journal of Anxiety Disorders,* Vol. 24 No. 3, 300–308.

Cook, K. W., Vance, C. A., and Spector, P. E. (2000), "The relation of candidate personality with selection interview outcomes", *Journal of Applied Social Psychology,* Vol. 30 No. 4, 867-885.

DeGroot, T., and Gooty, J. (2009), "Can nonverbal cues be used to make meaningful personality attributions in employment interviews?", *Journal of Business and Psychology,* Vol. 24 No. 2, 179–192.

Egloff, B. and Schmukle, S.C. (2004), "Gender differences in implicit and explicit anxiety measures", Personality and Individual Differences, Vol. 36 No. 2004, pp. 1807–1815.

Feeney, J.R., McCarthy, J.M., and Goffin, R. (2015), "Applicant Anxiety: Examining the sex-linked anxiety coping theory in job interview contexts", *International Journal of Selection and Assessment,* Vol. 23 No. 3, 295-305.

Feiler, A. R., and Powell, D. M. (2013), "Interview anxiety across the sexes: Support for the sex-linked anxiety coping theory", *Personality and Individual Differences,* Vol. 54 No. 1, 12-17.

Feiler, A. R., and Powell, D. M. (2016), "Behavioral expression of job interview anxiety", *Journal of Business and Psychology,* Vol. 31 No. 1, 155-171.

Frauendorfer, D., and Schmid Mast, M. (2014), "The Impact of Nonverbal Behavior in the Job Interview", In A. Kostić and D. Chadee (Eds)., *The Social Psychology of Nonverbal*

Communication (pp. 220-247). Houndmills, Basingstoke / New York, NY: Palgrave/Macmillan.

Gilboa-Schechtman, E., and Shachar-Lavie, I. (2013), "More than a face: a unified theoretical perspective on nonverbal social cue processing in social anxiety", *Frontiers in Human Neuroscience*, Vol. *7*, Article 904, 1-14.

Givens, D. B. (2014), "Measuring Gestures", In A. Kostić and D. Chadee (Eds). (2014), *The Social Psychology of Nonverbal Communication* (pp. 66-91). Houndmills, Basingstoke / New York, NY: Palgrave/Macmillan.

Heller Zaimont, R. (2014), "Can Your Voice Get You Hired? New Software Helps People Find Jobs By Analyzing The Sound Of Their Speech. Jobaline's new technology is designed to help combat bias in the hiring process", Blog post from 13 November 2014. Available at: https://www.fastcompany.com/3038463/innovation-agents/can-your-voice-get-you-hired-new-software-helps-people-find-jobs-by-analyz (accessed 21 February 2017)

Hosoda, M., Stone, D. L., and Stone-Romero, E. F. (2003), "The interactive effects of race and gender on job suitability ratings and job assignment decisions", *Journal of Applied Social Psychology*, Vol. 33 No. 1, 145-178.

Howard, J. L., and Ferris, G. R. (1996), "The employment interview context: Social and situational influences on interviewer decisions", *Journal of Applied Social Psychology*, Vol. *26* No. 2, 112-136.

Huffcutt, A. I., Van Iddekinge, C. H., and Roth, P. L. (2011), "Understanding applicant behavior in employment interviews: A theoretical model of interviewee performance", *Human Resource Management Review*, Vol. 21 No. 4, 353-367.

Kasl, S.V., and Mahl, G.F. (1965), "Relationship of disturbances and hesitations in spontaneous speech to anxiety", *Journal of Personality and Social Psychology,* Vol. 95 No. 5, 425-433.

Krueger, D. C. (2011), "*The effects of Applicant Weight, Conscientiousness, and Nurturing Job Demands on Job Suitability Ratings and Hiring Recommendations,"* PhD diss., San Antonio, USA: University of Texas at San Antonio.

Kwon, J.H., Powell, J., and Chalmers, A. (2013), "How level of realism influences anxiety in virtual reality environments for a job interview", *International Journal of Human-Computer Studies,* Vol. 71 No. 10, 978–987.

Langer, M., König, C.J., Gebhard, P., and André, E. (2016), "Dear Computer, Teach Me Manners: Testing Virtual Employment Interview Training", *International Journal of Selection and Assessment,* Vol. 24 No. 4, 312-323.

Leary, M. R. (1983), "Social anxiousness: The construct and its measurement", *Journal of Personality Assessment,* Vol. 47 No. 1, 66–75.

McCarthy, J. M., and Goffin, R. (2004), "Measuring job interview anxiety: Beyond weak knees and sweaty palms", *Personnel Psychology,* Vol. 57 No. 3, 607-637.

McCarthy, J. M., Van Iddekinge, C. H., Lievens, F., Kung, M.-C., Sinar, E. F., and Campion, M. A. (2013), "Do candidate reactions relate to job performance or affect criterion: related validity? A multistudy investigation of relations among reactions, selection test scores, and job performance", *Journal of Applied Psychology,* Vol. 98 No. 5, 701-719.

Morrison, L. (2017), "Speech analysis could now land you a promotion", BBC. Blog post from 9 January 2017. Available at: http://www.bbc.com/capital/story/20170108-speech-analysis-could-now-land-you-a-promotion (accessed 21 February 2017)

Poh, W. Y. F. (2015), "*Evaluating candidate performance and reaction in one-way video interviews*", Masters of Science in Psychology dissertation, San Francisco, USA: San Francisco State University, California.

Pope, B., Siegman, A. W., and Blass, T. (1970), "Anxiety and speech in the initial interview", *Journal of Consulting and Clinical Psychology,* Vol. 35 No. 2, 233-238.

Ragsdale, J. D. (1976), "Relationships between hesitation phenomena, anxiety, and self-control in a normal communication situation", *Language and Speech,* Vol. 19 No. 3, 257-265.

Riggio, H. R., and Riggio, R. E. (2002), "Emotional expressiveness, extraversion, and emotional stability: A meta-analysis", *Journal of Nonverbal Behavior,* Vol. 26 No. 4, 195–218.

Ryan, A. M., and Ployhart, R. E. (2000), "Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future", *Journal of Management,* Vol. *26* No. 3, 565-606.

Schwarz, N. (1999), "Self-reports. How the questions shape the answers", *American Psychologist,* Vol. 54 No. 2, 93-105.

Scullen, S.E., Mount, M.K., and Goff, M (2000), "Understanding the latent structure of job performance ratings", *Journal of Applied Psychology*, Vol. 85 No. 6, 956-970.

Sieverding, M. (2009), "'Be Cool!': Emotional costs of hiding feelings in a job interview", *International Journal of Selection and Assessment,* Vol. 17 No. 4, 391-401.

Stone, D. L., and Stone, E. F. (1987), "The effects of missing application blank information on personnel selection decisions: Do privacy protection strategies bias the outcome?" *Journal of Applied Psychology*, Vol. 72 No. 3, 452-456.

Viswesvaran, C., Ones, D. S., and Schmidt, F. L. (1996), "Comparative analysis of the reliability of job performance ratings", *Journal of Applied Psychology*, Vol. 81 No. 5, 557-574.

Voncken, M. J., Dijk, C., de Jong, P. J., and Roelofs, J. (2010), "Not self-focused attention but negative beliefs affect poor social performance in social anxiety: An investigation of pathways in the social anxiety–social rejection relationship", *Behaviour Research and Therapy*, Vol. 48 No. 10, 984-991.

Woods, S. A., Lievens, F., de Fruyt, F., and Wille, B. (2013), "Personality across working life: The longitudinal and reciprocal influences of personality on work", *Journal of Organizational Behavior*, Vol. 34 No. S1, S7-S25.

Zalta, A.K., and Chambless, D.L. (2012), "Understanding gender differences in anxiety", *Psychology of Women Quarterly*, Vol. 36 No. 4, 488-499.

Zimmerman, R. D., del Carmen Triana, M., and Barrick, M. R. (2010), "Predictive criterion-related validity of observer ratings of personality and job-related competencies using multiple raters and multiple performance criteria", *Human Performance*, Vol. 23 No. 4, 361-378.