

BIROn - Birkbeck Institutional Research Online

Enabling Open Access to Birkbeck's Research Degree output

Investigating the impact of individual differences on candidate reactions to selection processes

https://eprints.bbk.ac.uk/id/eprint/50864/

Version: Full Version

Citation: Welsh, Katy (2023) Investigating the impact of individual differences on candidate reactions to selection processes. [Thesis] (Unpublished)

© 2020 The Author(s)

All material available through BIROn is protected by intellectual property law, including copyright law. Any use made of the contents should comply with the relevant law.

> Deposit Guide Contact: email

Investigating the Impact of Individual Differences on Candidate Reactions to Selection Processes

Katy Welsh

Birkbeck, University of London 2022

Supervised by Dr Jo Yarker, Dr Rachel Lewis, Dr Halley Pontes

Thesis submitted in partial fulfilment for the degree of Professional Doctorate in Occupational Psychology (DOccPsy)

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, in whole or in part, in any previous application for a degree. Except where stated otherwise by reference or acknowledgment, the work presented is entirely my own.

Signed	
Print name	.Katy Welsh
Date	28/09/2022

Acknowledgements

Firstly, I would like to thank my supervisors Jo and Rachel for their support throughout the Doctorate. Thank you for giving me this opportunity, your feedback and guidance has been invaluable. I would also like to thank Halley for your insight and expertise; I have learnt a huge amount from you all.

Secondly, I would like to thank my participants who took the time to take part in my research during what I'm sure were challenging and uncertain times for many. This thesis would not exist without them taking time out of their day to support the study. I would also like to thank the UK Civil Service for their support and interest in my research and the time they have given to enable my study to take place.

Next, I would like to thank all my hugely supportive friends and family who have been there throughout the high and lows of this journey. Thank you for checking in on me when I have hidden myself away to write and for your understanding and acceptance that I have not been able to spend as much time with you as I would have liked while juggling the demands of study and work.

Finally, thank you to my husband Stephen. I could not have done this without you and I'm so grateful for your belief in me.

Abstract

There has been a rapid increase in the number of selection tools and procedures available to organisations looking to identify the most talented individuals in recent years, not all of which are well received by applicants. A poorly received selection procedure can have significant impact for both individuals and organisations. While candidate reactions have been a focus of research for several decades, the role of individual differences in candidate reactions has been less well-researched, despite numerous calls for more research in this area. This thesis aims to answer these calls and address some of the gaps in the evidence base.

The first study of this thesis is a systematic literature review of published research into the impact of individual differences in candidate reactions to selection processes. No known systematic literature review has been conducted in this area, despite several calls for research. The systematic literature review examined how individual differences in candidates going through a selection process related to their candidate experience. The results showed that there is a limited number of studies in the area, with 18 being identified which met the inclusion criteria. There are significant limitations in the replicability of studies in this area, driven by the number of selection methods available. The study findings showed that there is some promising evidence regarding the effects of attitudinal, biographical and personality differences on candidates' reactions, but that these effects are complex and often indirect. Implications for practice are discussed alongside suggestions for future research.

The second study of this thesis addresses some of the gaps identified in the systematic literature review. This study investigated the effects of personality differences on reactions to a low-touch, digitally enabled selection process with internal candidates applying to a talent development scheme. A cross-sectional study was adopted with 625 participants from a public sector organisation. Personality effects were investigated using a Big Five measure of personality, and reactions to both the selection process as a whole and reactions to the four individual exercises which made up the selection process were examined. The results showed small but significant effects for the personality factors of Conscientiousness, Agreeableness and Neuroticism on candidate reactions. Further, the impact of personality factors varied depending on the exercise, suggesting that multi-method approaches could be most beneficial in attracting broad talent pools.

This thesis advances understanding of the role individual differences play in candidate reactions to selection processes. Further, it builds on previous work highlighting the need for more research into the role of technology in selection processes and provides a number of avenues for future research.

Table of Contents

Contents

Declaration	2
Acknowledgements	3
Abstract	4
Table of Contents	6
List of Tables	9
Professional Practice Statement	C
Chapter 1 – Introduction	2
General Introduction	2
Candidate Reactions	3
Impact of candidate reactions	6
Reactions to different selection tools18	3
Individual differences in candidate reactions	C
The role of technology in selection22	2
Summary24	4
Chapter 2 – Methodology	5
Research Epistemology	5
The systematic review methodology	7
SLR Process	9
Study 2	2
Context - Field vs Lab studies	2
Survey design	3
Sample size	4
Reactions Questionnaire	4
Personality Questionnaire	5
Pilot Study	7
Process	7
Ethical Considerations	3
Chapter 3 – Systematic Literature Review: What impact do individual differences have on candidate reactions to selection procedures?	C
Abstract	C
Introduction	C
Rationale for this review43	3
Objectives44	4

Method	
Search Strategy	
Criteria for considering studies	
Sifting process	
Data extraction	
Data synthesis	
Quality assessment	
Results	
Context	
Interventions	53
Mechanisms	
Outcomes	
Assessment of Quality	
Discussion	72
Implications	74
Limitations	75
Conclusion and Recommendations	76
Chapter 4 – Empirical Study: The impact of personality on candidate react	ions to a selection process
for talent development schemes	77
Abstract	77
Introduction	77
Candidate reactions	77
Selection instruments	79
Applicant groups	
Individual differences	
Current study	
Method	
Design	
Sample	
Procedure	
Results	100
Correlations between PQ and RQ variables	100
Hierarchical regression analyses	110
Discussion	118
Discussion of findings	118
Strengths and limitations	

Future research
Implications for theory and practice 123
Conclusion 125
Chapter 5 - Discussion General 126
Synthesis of findings from Study 1 and Study 2 126
Summary of Findings
The role of attitudinal, biographical and personality factors in candidate reactions
Attitudinal factors
Biographical factors
Personality factors
Practical & Theoretical Implications 138
Theoretical implications
Practical implications
Research implications
Limitations
Conclusions 144
References 145
Appendix 1 – Reactions Questionnaire Items 164
Appendix 2 – Personality Questionnaire Items

List of Tables

SLR Inclusion and Exclusion criteria
Context of extracted papers
Intervention used in extracted studies
Details of the mechanisms investigated in the extracted studies
Details of the outcomes measured in the extracted studies
<i>Quality Ratings of extracted papers</i> 71
Evidence statements
Descriptive statistics and correlations between PQ and RQ variables for Total Process 105
Descriptive statistics and correlations between PQ and RQ variables for Candidate
<i>t</i>
Descriptive statistics and correlations between PQ and RQ variables for LAQ 107
Descriptive statistics and correlations between PQ and RQ variables for LPQ 108
Descriptive statistics and correlations between PQ and RQ variables for SJT 109
Hierarchical regression testing associations between Agreeableness, Neuroticism and
tiousness and Candidate Reactions to Total Process
Hierarchical regression model testing the associations between the interaction Personality
and Candidate Reactions to the Total Process
Summary of findings from Study 1 and Study 2 against the CIMO framework 127

Professional Practice Statement

I am a Chartered Occupational Psychologist and registered with the HCPC. I completed the Stage 2 Qualification in Occupational Psychology in 2016, following which I have continued to develop my professional practice through both career development and academic study. As such, I am exempt from the Professional Practice component of this professional doctorate and the work presented in this thesis reflects two years part-time study. Throughout my professional practice I have had a keen interest in evidence-based practice – both drawing on evidence to inform solutions and contributing to evidence creation – and have presented my research at a number of industry conferences.

My first role as a Research Psychologist was in the field of selection and assessment, working for a consultancy providing psychometric tests for use in both selection and development contexts. As I had an interest in and aptitude for statistics, I was tasked with converting a series of single format ability tests employing Classical Test Theory into Item Response Theory tests using large banks of items. In this role I also gained my Test User Qualification in Personality and worked as a tutor for both the Ability and Personality Test User training courses. By observing the positive impact that gaining an understanding of how their own personality affects individuals' experiences in the workplace, I became fascinated with the measurement of traits. Further, by gaining an understanding of the rigour that is applied to the development of assessment tools when best practice is followed, I wished to help other organisations avoid the pitfalls of poorly designed assessment and selection processes and improve their practice.

Building on my initial experience, I then worked for a number of years with an international provider of selection tools, working on modernising their flagship personality questionnaire¹ and developing both off-the-shelf and bespoke ability tests and Situational Judgement Tests (SJTs). I was also fortunate enough to gain a secondment to one of the international offices where I spent six months upskilling the team in the design and development of psychometric tests and delivered training courses to local HR practitioners in best practice selection approaches.

¹ Welsh, K., Tate, L., Mortenson, S. (2015). Cross-cultural development of a personality tool for international use. DOP Annual Conference 2015, Glasgow – Symposium Paper Presentation.

In order to broaden my practice, I then moved into a consultancy role where I worked directly with several large international clients to redesign and modernise their selection and development processes². In this role I focussed mainly on the design of work sample exercises for use in face-to-face and virtual assessment centres³. I worked on projects for a broad range of selection contexts, including selection onto development programmes, senior and executive leadership selection, and assessment for organisational restructures and redundancy.

My most recent roles have been in the UK Civil Service, where I have led on the design of a number of psychometric tests used in the initial stages of the recruitment process across Civil Service departments⁴ and led on the re-design of the selection process for senior leader talent development schemes which resulted in significant improvements in the fairness and rigour of the process, while maintaining positive diversity outcomes⁵. I have seen first-hand the impact that poorly designed selection processes have had on candidates, and I wish to expand my own knowledge of the factors that can drive candidates to react in different ways to the same process in order to help organisations develop a more positive selection experience for all. More broadly I am seeking to move my practice into a research-focused role. By undertaking this Doctorate I have developed my skills in academic methods which I will employ as a practitioner.

² Mescal, M., & Welsh, K. (2018). Audio versus text SJTs: The impact of delivery format. DOP Annual Conference 2018, Stratford-Upon-Avon – Oral Presentation.

³ Welsh, K. (2017). Automation in development centres: A case study from the publishing industry. UK ACG Conference 2017, Old Winsor – Breakout Session

⁴ Welsh, K. (2020). Improving Accessibility with an Alternative Format SJT. DOP Annual Conference 2020, Stratford-Upon-Avon – Oral Presentation.

⁵ Welsh, K., McShea, J. (2022). Selecting for senior leader talent programmes in large organisations: An evidence-based approach. EAWOP Annual 2022, Glasgow (Abstract accepted)

Chapter 1 – Introduction

General Introduction

Competition for talent has never been stronger, with organisations frequently competing for new hires from the same applicant pool. This was illustrated in Eploy's 2020-2021 annual report, which identified 'candidate scarcity' as the number one concern for recruiters in the UK (Eploy, 2021). The Covid-19 pandemic led to even greater challenges in recruitment, as the 'Great Resignation', a term coined by Professor Anthony Koltz to describe the everincreasing number of employees leaving their jobs during the pandemic (Cohen, 2021), resulted in significant increases in job vacancies. For example, the CNBC news site reported a record 11.5 million job openings in the US in 2022 (Smith, 2022) and the Office for National Statistics (ONS) in the UK reported vacancies reaching a record high in April 2022 of 4.3 vacancies to every 100 employee jobs (ONS, 2022). Alongside this, advances in technology mean that organisations can now easily contact potential candidates who are not actively looking for a new role (van Esch & Black, 2019). With this competition, organisations must ensure they keep candidates engaged and feeling positive about the organisation during the selection process, or risk losing them to competitors.

The selection process is often the candidates' first experience of the organisation. As such the consequences of the selection process are far-reaching. A positive experience can mean that applicants are likely to look more favourably on the organisation and potentially increase their future engagement with it (for example increased custom in the case of commercial organisations) regardless of whether they were successful or not (Miles & McCamey, 2018). In contrast, a poor experience can put applicants off further contact with the organisation, perhaps withdrawing custom and passing on their negative experiences to others (Nikolaou, 2021). A stark example of this was reported by Virgin Media, who estimated a loss of £4.4million due to poor candidate experience (Scarborough, 2020). In Viswesvaran and Ones (2004) examples of adverse reactions to selection processes are described that result in litigation and the boycotting of the organisation in question. Despite these far-reaching impacts, research has tended to focus on the design of selection processes (in particular the validity of selection processes, Potočnik et al., 2021), with the impact on candidates receiving less attention.

While 'selection' is often used to describe the process of recruiting applicants to a specific job role, the term can be used to describe any process of defining and measuring a specific set of knowledge, skills, abilities and other attributes (KSAOs) in a given population and identifying those with the desired levels of the KSAOs of interest. Selection processes can be used to identify current employees for further training and development; for access to high potential talent schemes; to take on specific job tasks; to identify people for secondment opportunities or expatriation; or even to help make redundancy decisions.

Candidate Reactions

Understanding the factors that contribute to candidates' feelings about selection processes has been a growing area of interest for many researchers over the years. While early work in selection focused on the organisational perspective, several seminal publications in the early 1990s moved the focus to understanding the individual perspective (Gilliland, 1993; Arvey et al., 1990; Arvey & Sackett, 1993, as cited in Gilliland, 1993).

A paper by Gilliland in 1993 along with other foundational theoretical work carried out at a similar time (e.g., Arvey & Sackett, 1993, as cited in Gilliland, 1993; Schuler, 1993) provided the basis for many researchers interested in the candidate perspective of selection procedures. Gilliland's paper proposed a procedural justice framework against which many later candidate reaction studies were based. This framework identifies a range of components that contribute to the candidates' overall perception of the selection process, specifying ten procedural rules. These rules cover: job-relatedness, opportunity to perform, reconsideration opportunity (for example the opportunity to correct a mistake made during a selection activity), consistency of administration, feedback on performance, selection information (e.g., a clear description of how selection decisions are made from the outset), honesty (of the interviewer/administrator), interpersonal effectiveness of the administrator, two-way communication, and propriety of questions.

Several studies have been conducted which support this model as a framework for understanding candidate reactions. For example, Truxillo et al. (2009) conducted a metaanalysis that examined the impact of selection explanations and based their findings on Gilliland's model. In their analysis they focused on explanations about the structure of the selection process and the interpersonal effectiveness of such explanations. They found that explanations in general impacted fairness perceptions and had other consequences such as impacting perceptions of the organisation, test motivation and test performance. This study had some clear practical applications for organisations as it was found that, in general, explanations do improve reactions, and providing candidates with such explanations is likely to be relatively straightforward for organisations to introduce if they are not doing so already.

Gilliland also tested his model (Gilliland, 1994) to examine the impact of explanations. He found that when explanations around selection decisions were given, rejected candidates were more likely to recommend the project they had been selected for and apply for similar projects again. The study also found that when job-relatedness was high, self-efficacy (defined as individuals' beliefs about their own capabilities and ability to influence events that effect their lives; Bandura, 1994) was increased for successful candidates and decreased for unsuccessful candidates. Gilliland proposed this was not necessarily a negative outcome for candidates as it formed an element of self-awareness that the role may be unsuitable for them. Contrary to the procedural justice model, explanations about the selection process did not affect perceived fairness of the process. However, the explanations given focused on why the procedure was used rather than why the specific individual outcome decision was made. Also, it should be noted that the study was carried out in a university as selection to work on paid projects rather than a specific job role, therefore the stakes may not have been as high as in selection for a job role, on which the model was based.

Other proposed models cover many similar elements as Gilliland. For example, Schuler (1993) proposed a model of four factors that influence how acceptable candidates found a selection process, including information (about the job and organisation), participation or representation in how the selection process was developed, the transparency of the process (i.e., how easy it is to understand how selection decisions are being made), and feedback. Also cited in Gilliland (1993) is the model proposed by Arvey and Sackett in the same year, which suggested that the factors affecting the perceived fairness of the selection process are the content, an understanding of how the process was developed, the administration of the process, and the 'organisational context' (such as the ratio of candidates selected to applicants).

Following Gilliland's work, Ryan and Ployhart (2000) provided a critical review of the candidate reaction literature from the preceding fifteen years. Their review was themed

around the types of perceptions studied, the determinants of perceptions, the consequences of such perceptions and the theoretical frameworks for examining perceptions. In their review, the most studied perceptions were those associated with the job-relatedness of the selection process, various aspects of fairness perceptions, process outcomes, and the applicants' motivations. They concluded that the constructs measured are not always clearly defined and a number of perceptions are likely to be important but aren't routinely measured (referencing Ball et al., 1993, who describe 'procedural pain' or risk of discomfort or humiliation in the selection process as one such example). While providing a comprehensive review of the literature at that time, the technological advances of the 20 years following publication of Ryan and Ployhart's work mean that some of their conclusions are less relevant now. For example, the review stated that there was no evidence of negative candidate reactions leading to negative behaviour such as 'badmouthing' the organisation. However, more recent work has found such evidence. A paper by Nikolaou (2021) gave the example of the online platform Glassdoor, which reaches 41 million users per month. On this platform job applicants can anonymously share reviews of their experience of selection processes, and examples of negative feedback can affect the organisation's branding or even their financial evaluation. These negative behaviours were also illustrated in the Virgin Media example referenced previously (Scarborough, 2020, reporting on the company's estimated loss of £4.4million due to poor candidate experience).

McCarthy et al.'s (2017) more recent critical review of research on applicant reactions to selection procedures addressed three key areas: the importance and relevance of applicant reactions (described as the "So What" question), recent findings in the field of applicant reactions (the "What's New" question), and future directions for study (the "Where to Next" question). The authors concluded that there is clear evidence that applicant reactions have significant effects on attitudes, intentions, and actual behaviours. Further, there were clear links between reactions and performance in selection tests, and even performance on the job. The authors highlight advances in technology as a key area for the latest research and identified that many studies have expanded on Gilliland's applicant reactions model to include additional determinants of justice perceptions that are specific to technology. With regard to emerging areas, the authors identified four new themes: expansion of the theory, the growing emphasis on new technology, more international research, and new boundary conditions. However, it was not clear from the paper how these themes were identified, and it is likely that there will be other emergent areas to consider. Nonetheless, their review does

highlight the ongoing relevance of research on candidate reactions and acknowledges the need for continued research to develop the field.

Impact of candidate reactions

There can be significant negative impacts of a poorly received selection process for both candidates and organisations. One impact of poor selection processes can be seen in complaints received by organisations from candidates, which can potentially lead to legal challenges (Goldman, 2001). Complaint handling can require significant resources from the organisation and legal challenges can have enormous financial and reputational impacts. It therefore follows that any action that aims to reduce the risks of complaints and subsequent legal challenges will be highly desirable for organisations. Taking a more positive angle, Hausknecht et al.'s (2004) review identified that applicants who held positive perceptions about the selection process had more favourable views of the organisation and had stronger intentions to accept job offers and recommend the employer to others. Collins and Han (2004) also note that organisations that can attract bigger pools of higher quality candidates are more likely to succeed as they are able to be more selective in the hiring decisions that they make.

The selection methods used can also have implications before candidates have even reached the selection process, impacting on the pool of candidates applying to the organisation. For example, Langer et al. (2018) suggested that discomfort felt in digital interviews could lead to candidates self-selecting out of the process, or not applying in the first place.

To assess the long-term impact of selection processes on candidates, Anderson and Goltsi (2006) examined the 'negative psychological effects' (NPEs) of an assessment centre. They measured constructs such as self-esteem, affectivity, and psychological wellbeing prior to the assessment centre, immediately after the centre, and six months after. Differences between accepted and rejected candidates were explored; the authors found that while there were no lasting NPEs for the rejected candidates, this group did report significantly more negative perceptions of the feedback given. The finding that there were no lasting NPEs was unexpected, as previous research (e.g., Fletcher, 1991) had found lasting effects. The authors noted that this study was carried out with a prestigious company and that there was a very

small ratio of selected candidates to unselected, which may have moderated the outcomes. For example, knowledge of the selection ratio is one of the requirements stated in Arvey and Sackett's model (1993, as cited in Gilliland, 1993) and may have contributed to overall fairness perceptions of the assessment centre. The authors also noted that very little research has been carried out concerning such long-term psychological effects on other types of selection processes, suggesting that the results may differ for processes that are seen as less job-related than assessment centres.

More recently, a meta-analysis (Konradt et al., 2020) looked at whether applicants' perceptions of fairness change over time. This study found that fairness perceptions did decrease from initial measurements to post-selection and that this effect was most profound when fairness expectations were initially high and the time gap was small. The authors suggested that identifying applicants most likely to be impacted early in the selection process could prevent negative outcomes, such as them withdrawing their application. However, it is not clear how organisations would practically go about identifying such applicants or how effective the suggestion to supply applicants with further information about the organisation would be. As with Anderson and Goltsi's study (2006), Konradt et al.'s (2020) findings indicate that further research is needed to investigate the effect of the types of selection methods used in more detail.

The experience of the selection process can also have important implications for the individual in terms of the impact on job performance and further outcomes based on performance, such as reward and recognition. To investigate the long-term effects of a selection process, Konradt et al. (2017) carried out a longitudinal study looking at the impact of procedural fairness over a three-year period. Their study showed that acceptance of a job offer was significantly related to justice perceptions, illustrating the link between the procedural justice of the selection process and applicant behaviour. They also showed that job performance 18 months after the selection process was related to justice perceptions. As noted by the authors, this study used an apprentice applicant sample, therefore may not generalise well to other settings. In particular, the findings may be less relevant in more senior or managerial samples, where participants may hold a number of preconceptions about the selection process based on their past experiences. Despite this limitation, this study does highlight the potentially far-reaching consequences of a poorly received selection process.

Reactions to different selection tools

Previous research has provided us with a growing understanding of candidates' reactions to different selection methods (e.g., Hausknecht et al., 2004; Steiner & Gilliland, 1996). However, there is an ever-increasing number of selection tools available to organisations, and candidate reactions to these differ greatly. Anderson et al.'s meta-analysis (2010) found that work sample tests and interviews were most favoured by applicants, followed by resumes, cognitive tests, references, biodata and personality inventories. The least favoured measures were honesty tests, personal contacts (i.e., knowing someone of influence at the organisation) and graphology. Several studies have shown these preferences to be largely similar across countries and samples. For example, Anderson and Witvliet's 2008 study compared candidate reactions in the Netherlands to studies carried out in the United States, France, Spain, Portugal, and Singapore. They found that work samples, interviews and resumes tended to be the most popular methods across these countries, with work samples also being the most popular method in an Italian sample (followed by resumes, written ability tests, interviews and personal references, Bertolino, & Steiner, 2007). These results were also supported by Hausknecht et al. (2004) in their meta-analysis, who found that interviews, work samples, resumes and references were the most favourably rated by applicants.

However, these findings may become less relevant as the types of selection methods used by organisations evolve. For example, very few countries use methods such as graphology as part of their selection process. This was illustrated in a study of Swiss job adverts (searching a database of job adverts from 1950 to 2007) which found that less than 1% requested a handwritten letter as part of the application process (Bangerter et al., 2009), and a UK study (Branine, 2008) found that none of 326 graduate employers surveyed used graphology as part of their selection processes. These reviews also do not look at newer commonly used selection methods such as Situational Judgement Tests (SJTs), which are now used by many large employers in the UK, including the NHS (Petty-Saphon et al., 2017).

Studies have shown some differences between how favourable candidates find different selection methods in general and how valid they perceive them to be. When investigating candidate perspectives on the scientific validity of selection methods, Anderson et al. (2010) found that cognitive ability tests were most highly rated, followed by work samples then personality inventories. Work samples and interviews received the most favourable ratings

regarding their perceived face validity and candidates' belief that they had the opportunity to perform. In relation to other procedural justice rules, interviews and work samples performed best with regards to 'interpersonal warmth', whilst cognitive tests performed worst. With regards to respecting privacy, cognitive ability tests fared much better. As was a consideration for some of the previous studies, it is not clear whether these results would generalise to applicants to more senior roles, as most studies identified for this meta-analysis focused on more junior roles or were with student samples. Further, it is unclear whether any internal candidate groups were considered, as they may have different perceptions from candidates external to the organisation.

Some more recent work compared the validity and fairness perceptions of 'lay people' (defined as those who had no human resources-related qualifications) and practitioners to different selection methods (Jackson et al., 2018). This study found that lay people ranked work experience, interviews and job try-outs as the most valid methods, and biodata such as measures of age and interests as the least valid. The same pattern was found for perceptions of fairness. Unlike Anderson et al.'s (2010) study, personality was next lowest rated; however, this may have been due to it being presented as subjective judgements of an individual's personality, rather than in Anderson's study where it was clearly presented as 'personality tests'. This study had some advantages over earlier studies as it focused on more commonly used and understood selection methods (at least, in the UK). However, the need to ensure that lay people understood the method they were asked to rate means that not all commonly used methods were examined; for example, work samples and SJTs are notably missing from the study.

As illustrated by these studies, the past 30 years have provided several theoretical frameworks against which candidate reactions can be understood, and researchers have sought to provide empirical evidence to support the theories. While the body of research in the area of candidate reactions is substantial, a number of reviews and meta-analyses (e.g., McCarthy et al., 2017; Ryan & Ployhart, 2000) have drawn attention to areas less well understood in the research, such as how new technologies may impact on candidates, and antecedents and determinants of candidate reactions.

Individual differences in candidate reactions

It has been noted that the determinants of candidate reactions have tended to be unexplored in the research. For example, as described in Ryan and Ployhart's review (2000), there have been relatively few studies focusing on individual differences as correlates of candidate reactions.

Individual differences in the field of psychology have typically been related to understanding how and why psychological traits (such as personality and cognitive ability) differ between individuals; more recently the field has been described as 'differential psychology' (Johnson, 2020). The APA defines individual differences as 'traits or other characteristics by which individuals may be distinguished from one another', and while typically referring to psychological traits, biographical characteristics may also be considered under this definition. Sackett et al. (2017) define individual differences as attributes which are generally stable (in contrast to transitory) within individuals. A further related term is that of group differences, for example comparing and contrasting the experiences of males and females, or of older and younger workers (Sackett et al., 2017).

Individual differences are often examined in relation to personality, with the 'Big Five' model of personality being one of the most well-supported in the literature (developed by a number of researchers, most significantly Costa & McCrae, 1985; and Goldberg, 1992). The model describes five broad personality traits against which individuals can be measured. These are typically defined as Extraversion, referring to the individual's tendency to seek out social interactions and how comfortable and assertive they are in social situations; Openness to Experience, concerned with the individual's willingness to try new things and engage with intellectual activities; Conscientiousness, which is broadly concerned with levels of impulse control and persistence; Agreeableness, which describes how co-operative, flexible and friendly individuals are, and Neuroticism; which looks at the overall emotional stability of the individual (Barrick & Mount, 1991).

Other areas of individual difference research have looked at traits beyond those defined in the 'Big Five', such as 'core self-evaluations'. This is a broad dispositional trait comprising the specific traits of self-esteem, generalised self-efficacy, locus of control and emotional

stability, which has been linked to a number of job-related factors (Judge & Bono, 2001). The relationship between core self-evaluations and perceptions of fairness across different selection methods was investigated by Nikolaou and Judge (2007). Some small but significant relationships were found, for example those with a positive self-image had a more favourable view of the interview process and felt it was a fair way of discriminating between candidates. Other researchers have looked at the dispositional trait of 'affectivity', which is typically studied as two separate traits: positive affectivity, which is associated with feelings such as enthusiasm and alertness; and negative affectivity, which is associated with feelings such as guilt, fear and anxiety (Kaplan et al., 2009).

While not specifically focused on individual differences, one meta-analysis by Hausknecht et al., (2004) reviewed 86 studies regarding applicant reactions. This review found limited support for individual differences impacting on applicant reactions, reporting the average correlations for gender, age and ethnicity were near zero. The paper also described average correlations for Conscientiousness and Neuroticism as small. However, this study looked purely at correlational relationships and may not have accounted for interaction effects, mediators and moderators which could play a role in the relationships. This study also focused primarily on more traditional assessment methods (work samples, interviews, cognitive ability tests, honesty tests etc.) and does not account for newer selection methods such as asynchronous video interviews and multimedia selection tests (such as video and audio SJTs). Further, of the 86 studies identified for the meta-analysis, only 6 correlations for Conscientious and Neuroticism were reported, whilst other aspects of the Big Five were not considered. The authors highlight that very little research has been done into the reactions of candidates who drop out of the selection processes (i.e., self-selecting out). This may mean that some individual differences are masked due to individuals with certain traits, or levels of a trait, dropping out of the process before reactions can be captured. They also note the lack of research on reactions to promotion processes, which could yield quite different results due to the pre-existing relationship with the organisation. Regarding personality differences, the authors go on to suggest that there could be a relationship between Openness to Experience and reactions to more innovative selection methods.

To investigate some of the impacts of individual differences, a study by Schinkel et al. (2011) looked specifically at reactions to rejection from selection processes. Their study identified that there was an interaction effect between distributive justice perceptions and self-reported

affective wellbeing, and that the more individuals perceived the selection process as fair, the more positively they felt about the organisation. The specificity of the feedback was also important; for those with an optimistic style, receiving specific feedback lowered their wellbeing. One limitation of this study was the sample; as is the case with many studies in this area the research was conducted using university students. This means that it is not clear whether the results will generalise to a real-world setting. For example, Hausknecht et al. (2004) illustrated that effect sizes tended to be higher in hypothetical settings than for real-world studies. A further study by Tay et al. (2006) examined personality differences in relation to success at job interviews. They found that Extraversion and Conscientiousness had an indirect relationship with interview success, mediated by 'interviewing self-efficacy' (individuals' beliefs about their interview capabilities).

These findings suggest that there are gaps in the research regarding individual differences, likely a reflection on the many different types of individual differences to consider (personality traits, broader dispositional or attitudinal traits, biographical differences). Further, there are some conflicting findings across the applicant reactions literature, indicating that there is much that is not clearly understood regarding the factors affecting candidate reactions.

The role of technology in selection

As referenced in preceding reviews (e.g., McCarthy et al., 2017), a growing area to be considered is the role of technology in selection procedures. Technology is changing the way recruitment is carried out, with technology-enabled selection procedures such as online testing, real-time and asynchronous remote interviews, and virtual assessment centres increasing in use (Woods et al., 2020).

A recent review looking at 'digital selection procedures' (DSPs, a term coined by Woods et al., 2020, to describe a range of technology-enabled selection activities and used hereafter) highlighted the lack of research on such procedures despite their preponderance as a selection method. The authors noted that assumptions are made about how DSPs will function with regards to their reliability, validity and fairness based on findings from historical research on non-digital methods (e.g., paper-and-pencil tests, face-to-face interviews, etc.). It was assumed that these findings would transfer to DSPs, however there is no evidence to support

this assumption. While Woods et al.'s study focuses predominantly on factors around the fairness and validity of DSPs, it is expected that the same issues would apply to candidate reactions, in that it is not known whether previous findings on candidate reactions to selection procedures are applicable to digital selection procedures.

One study prior to Woods et al.'s review (Oostrom et al., 2012) sought to compare reactions to a paper-and-pencil 'in-basket exercise' (a traditional assessment centre exercise where candidates are asked to respond to simulated communications such as letters and memos) with a computerised version. In their field study, the authors found that while there were no pre-test differences in perceptions of face validity, predictive validity, and fairness between the paper-and-pencil and computerised version, post-test candidates had more positive perceptions of the paper-and-pencil version than the computerised exercise. Chapman et al. (2003) also found that reactions to face-to-face interviews were more positive than reactions to interviews delivered via videoconference or telephone. These studies support Woods et al.'s suggestion that it cannot be assumed that DSPs can be considered as the same as, or preferable to, more traditional selection methods. In another earlier study investigating computerised testing, Wiechmann and Ryan (2003) found that level of comfort and experience in working with computers had a significant impact on reactions to the selection process. This also supports the assertion that reactions to DSPs will differ from reactions to traditional methods; with DSPs there are likely to be additional factors that can impact on reactions to the process.

One type of DSP increasingly studied in the research is the selection interview. Blacksmith et al. (2016) carried out a meta-analysis of 12 studies and found consistently less favourable applicant reactions to technology-enabled interviews, with the authors suggesting that this could be due to perceptions of unfairness, feelings of frustration and technology-enabled interviews being less personal than face-to-face interviews. The interviewers also tended to give lower ratings than with face-to-face interviews, which the authors suggested could result in lower fairness perceptions of the process.

In addition to technology enabling virtual interviews, further advances now allow for interviews to be carried out asynchronously, adding an additional layer of complexity to understanding how candidates' perceptions may differ from traditional face-to-face interviews. 'Asynchronous' refers to the fact that interviewees and interviewers do not need

to be present at the same time. Instead, the candidate records a video of themselves answering set questions which are assessed by the interviewer at a later point in time. In a 2018 study, Langer et al. looked at the impact of these asynchronous interviews on both candidates and interviewers. The researchers compared asynchronous with real-time video conference interviews and found that the interviewees found the asynchronous interview less pleasant (experiencing more 'creepiness') than the real-time interviews and experienced greater privacy concerns. With regard to procedural justice, interviewees in the asynchronous interview as much lower (despite the real-time interview also being highly structured) but opportunity to perform and organisational attractiveness was not impacted. The authors note that the discomfort felt during asynchronous interviews could in fact cause potential candidates to refrain or withdraw from the selection process.

A further study by Brenner et al. (2016) investigated candidate reactions to asynchronous interviews in relation to other selection tools. In terms of favourability, asynchronous video interviews were rated as favourably as tools such as personality inventories and online tests. This is lower than interviews are typically rated in such studies, offering further support to the argument that existing research on traditional selection methods does not necessarily apply to newer DSPs.

Taken collectively, these studies suggest that there are some real differences in how candidates react to DSPs compared to traditional selection methods. This suggests that many of the findings of earlier studies may not translate to the current selection environment, and further research is needed to understand what factors affect candidates' reactions when organisations are using DSPs.

Summary

The previous reviews and meta-analytic studies regarding applicant reactions have given limited focus to individual differences in applicants. Where focus has been given, the findings have been very mixed, not least made increasingly complex by the sheer number of selection tools available (with virtually all selection tools having scope for multiple delivery methods) and myriad aspects of individual difference to explore. In their 2004 article, Chan and Schmitt propose an agenda for future research to be carried out regarding candidate reactions to selection processes. They recommend carrying out more research to look at the impact of personality on reactions to selection methods (specifically to selection tests), noting that the majority of research in this area has focused on race differences. This recommendation to study individual differences further has been echoed in a number of other papers, such as Brenner et al. (2016), who noted that a surprisingly small amount of research has been carried out looking at the role of personality on candidate reactions (particularly with regard to reactions to new technologies), and Oostrom et al. (2010), who noted that there has been relatively little exploration of the relationships between individual differences and applicant reactions.

Despite these reviews and calls for further research, research concerning the impact of individual differences in candidates on their reactions to selection procedures remains limited. An understanding of individual differences could influence how practitioners go about designing their selection processes. If a selection method is poorly received by individuals with a certain group membership (e.g., females) or a certain trait (e.g., introversion), the organisation could risk losing applications, narrowing their pool of applicants and missing out on candidates with highly desirable traits. In a similar vein, candidates could be self-selecting out of selection procedures that are off-putting to them, despite being potentially well-suited to the role in question.

Chapter 2 – Methodology

Research Epistemology

This thesis employs a Systematic Literature Review (SLR) and quantitative study, both of which follow well-established processes and procedures as developed and tested in the Industrial and Organisational (I/O) psychology literature.

Both methods were approached from the positivist perspective, a determinist philosophy which aims to explain associations or casual relationships in order to verify theories. In the positivist approach, concepts are reduced to discrete phenomena which can be tested; this is typically done though studies with large sample sizes, using paradigms which separate the researcher from the researched (Creswell, 2014; Park et al., 2020).

There are a number of features of this approach which make it suitable for the current research topic. The positivist approach allows for generalisations, meaning that it is often used in I/O psychology where practitioners are looking for study results that can be generalised to their own practice and inform policies and procedures. Johnson and Cassell (2001) provide a review of epistemological debates in I/O psychology; at time of writing they noted that I/O psychology is ingrained in the positivist paradigm. While their review was aimed at highlighting the need to consider a range of approaches in the field, the authors acknowledged that positivism has been crucial in establishing I/O psychology as a credible discipline.

It is acknowledged that there are some challenges in the positivist approach. For example, the reliance on large sample sizes when conducting statistical analysis can make it challenging to access certain participant groups, and generalisations may miss nuances in the data, such as the intention behind observed actions (Alharahsheh & Pius, 2020). Due to the separation between the researcher and the participants, the approach does not allow for reflexivity in practice or for subjectivity in interpretations (Johnson & Cassell, 2001).

The positivist approach can be contrasted with the interpretive/constructivist approach, which does not aim to identify rules or 'laws' that can be generalised across groups, but rather seeks

to understand the individual experience and considers different cultures and circumstances. It is assumed there is no single reality, but rather what is experienced as reality is unique to the individuals (Alharahsheh & Pius, 2020). A core principle of this approach is that the act of studying something influences the observation (Heisenberg, 1958, as cited in Johnson & Cassell, 2001), therefore reflexivity is critical in interpretation.

Sitting between these two approaches is the epistemology of critical realism, which is based on the notion that reality is socially constructed through influences such as politics, culture, gender and economics (Scotland, 2012). This approach can be appropriate for mixed-method research (Scott, 2005) and places an emphasis on describing and explaining observed phenomena (Sousa, 2010).

The systematic review methodology

Study 1 employs a systematic literature review (SLR), described by Briner and Denyer (2012) as a process which addresses a specific research question and uses clear, structured methods to produce a detailed literature search. It provides a critical appraisal of studies identified and draws conclusions about what is known (and unknown) about the given topic. The process was first used in the field of medical research (Leucht et al., 2009), and has been conducted in a number of different ways depending on the field. In medical science the systematic review process has typically been focused on synthesising cause and effect relationships, whereas in the field of I/O psychology and related disciplines (such as Human Resources Management, and Management and Organisation research) it has wider applications to address evidence from a variety of research questions (Rojon et al., 2011). The guidance from medical science for conducting an SLR was adapted by Denyer et al. (2008) to be more appropriate to the field of management and organisation studies. Denver et al. (2008) suggested that systematic reviews should be tested for their transparency, inclusivity, explanatory, and heuristic natures. With regard to transparency, they posit that the reviewer must be open and explicit about the process and methods used within the review to ensure readers are clear on the scope of the study, and to provide clarity in reporting the findings to make clear the links between the research identified and the conclusions and recommendations of the review. The principle of inclusivity suggests that rather than having a hierarchy of evidence to include (such as in medical research, where double blind,

randomly controlled trials (RCTs) are seen as the most desirable), all research that adds to understanding is included. In terms of explanation, rather than aggregating data, the review should seek to find a feasible explanation of the study findings and finally the review should provide some heuristics (such as rules, suggestions or guidelines) that make progress towards a solution (but do not necessarily need to provide detailed solutions as an outcome of the review).

The systematic review process has become more common in the field of I/O psychology and management research over the past decade, with a study by Rojon et al. in 2011 identifying no published reviews on a search of the 13 leading journals in the field, compared to a study in 2021 (Rojon et al., 2021) which identified 391 systematic reviews in this field. A further study by Snyder (2019) noted that the field of business research is expanding at an increasing rate, making it challenging to remain up to date with current thinking and evidence. This means that a structured process for capturing and evaluating research is particularly important. Snyder described systematic reviews as being used to provide an overview of a certain topic, or to evaluate the current state of knowledge, which can then be built on to identify future research agendas. However, the systematic review is not the only approach to reviewing previous research. Rojon et al. (2011) compares the advantages of SLR with traditional narrative reviews and meta-analysis, noting that SLRs provide rigour and objectivity due to the need to adhere to prescribed stages of the review and allow for consideration of all relevant evidence due to the inclusivity principles of the approach, meaning they can allow for both meta-analytical and narrative synthesis of findings. By comparison, the traditional narrative review does not have such a formal methodological structure, which means it can be more difficult to replicate and the researcher may be biased in deciding what to include and exclude. Snyder (2019) also compared the SLR process with some similar review processes, such as an integrative review. They noted that the purpose of the integrative review is to enable new theoretical frameworks to emerge, whereas the SLR process is better suited to areas where the research question is narrow.

An alternative to the SLR for Study 1 could have been to carry out a meta-analysis. This technique is most suitable where there are a range of comparable quantitative datasets and can be challenging when there are variations in the study design and nature of the data collected (Briner & Denyer, 2012). Given the scope of the current study, a meta-analysis would not have been appropriate due to the range of study designs required to investigate all

of the individual differences in scope. For example, had the study focused purely on categorical differences such as gender, a meta-analysis may have been suitable as it is likely that effect sizes would be the primary method of investigation. However, the individual differences comprised categorical and scale data investigated through group differences and effect sizes, direct correlational relationships, and more complex indirect relationships, making aggregation of findings very difficult.

Compared to meta-analysis, SLRs can be more exploratory in nature and do not require as large a body of preceding research as meta-analysis does, making them particularly suited to emerging or under-researched areas (Briner & Rousseau, 2011). The meta-analytical approach can also result in researchers being highly selective in what is included, narrowing the pool of evidence, and is typically focussed on synthesising quantitative data rather than allowing for narrative interpretation (Rojon et al., 2011). For these reasons, an SLR was the chosen research method for Study 1.

SLR Process

The current SLR followed the protocol described in Briner and Denyer (2012) which involves a number of stages, each of which has considerations, risk factors and mitigations to ensure the study is as replicable and robust as possible.

The initial stages of an SLR involve the development of the review question, which can include different techniques such as using an advisory group, testing the question by searching for relevant studies, or using a framework to focus the research question. Briner and Denyer (2012) caution that the review question is crucial, and that a poor-quality question will result in a poor-quality review. Liberati et al. (2009) recommend that the PICOS framework is used (referring to the Population, Intervention, Comparator group, Outcome and Study design) as a framework for structuring the research question; however, this is more suited to medical intervention studies than I/O psychology. Denyer et al. (2008) developed the CIMO framework as an alternative for the social sciences, which describes the Context, Interventions, Mechanism and Outcomes. The current SLR used a combination of techniques to develop the research question, first by wider reading around the topic areas and testing similar research questions to identify some existing research, followed by refinement through

discussion with the researcher's supervisors and formulation of the review question against the CIMO framework.

The next stage of the process involved the development of search terms and selection of databases to ensure that the review could be replicated and updated by other researchers. If the search terms are too narrow research could be missed, but if they are too broad the volume of research returned could become unmanageable. Snyder (2019) acknowledges that there may be gaps in the literature collected or even the wrong conclusions drawn as the outcomes are dependent on the search terms, suggesting that while there is not a definitive mitigation for this, transparent documentation of all steps taken and decisions made will allow future researchers to understand the process and develop the review further if warranted. To develop robust search terms for the current review, significant reading around the subject area was undertaken to identify the terminology used by researchers in the space. After selection of search terms, a second researcher reviewed the terms before the protocol was finalised. An alternative approach could have been to gather a group of subject matter experts to agree on search terms, however this would have been more resource intensive and may not have yielded significantly different results. To ensure appropriate databases were selected the subject area librarian was consulted, and to mitigate against the risk of missing key papers due to poorly defined or narrow search terms, hand searching of key journals was carried out which aimed to pick up any papers that had not been identified through the searches.

To identify which research studies to include in the review, inclusion and exclusion criteria need to be developed. The process of developing clear inclusion and exclusion criteria from the outset means that there is significantly reduced risk of researcher bias (either deliberate or unconscious), as there is limited risk of the researcher 'cherry picking' which papers to include in the review (Briner & Denyer, 2012). Rojon et al. (2011) also notes that the use of inclusion/exclusion criteria can help ensure quality while avoiding the risk of digression into other research areas. To ensure the inclusion and exclusion criteria were suitable for the current study, they were developed against the CIMO framework and reviewed by the researcher's supervisors.

After developing the search terms and inclusion and exclusion criteria, and searching in the identified databases, the next step of the process involved reviewing the papers for their

inclusion or exclusion in the final study. To manage the volume of papers, a review of the paper titles was carried out first, followed by a review of the abstracts. This has some risks, as unclear or poorly written titles and abstracts could result in a relevant paper being missed. To mitigate against this risk, as recommended in Briner and Denyer (2012) a second researcher reviewed a random selection of 10% of the papers and abstracts to check for agreement on their inclusion/exclusion status.

Once the papers were identified, the study details were extracted and assessed for quality. Liberati et al. (2009) detail that all study characteristics extracted should be reported, giving the example of the PICOS framework to structure the information. The current study used the CIMO framework as this was more suited to the research area, however it follows the same principles in terms of structuring the extraction to aid interpretation.

As Briner and Denyer (2012) note, a quality appraisal is needed to understand whether any differences found between studies could be due to variations in quality, and to guide the reader in how much confidence can be placed in the findings. Rojon et al. (2021) go on to suggest that using pre-defined quality criteria will increase the ease of carrying out the quality appraisal and will result in the SLR reporting evidence of a higher quality. The current study used Hong and Pluye's (2019) framework to look at the quality of individual studies, and Snape et al.'s (2017) checklist was used to develop questions to assess methodological quality. As noted previously, earlier SLRs have looked at hierarchies of studies, placing the most importance on RCTs, however this is less well suited to the I/O psychology field where RCTs are rare (Rojon et al., 2021). Hong and Pluye's framework looks more broadly at the methodological, conceptual and reporting quality of the studies, making it more suited to the breadth of studies in this field.

The final stage of the SLR was to analyse and synthesise the results. Briner and Denyer (2012) describe the analysis process as the examination of individual studies, whereas the synthesis brings the findings from the individual studies together. There are a number of techniques for synthesis, which can be grouped into four categories: aggregation, integration, interpretation and explanation (Rousseau et al., 2008). Aggregative approaches aim to combine effects to increase sample sizes and reduce bias; typically these are used with quantitative studies where the methods used are homogeneous. Integrative approaches combine qualitative and quantitative studies, often to explore the appropriateness of an

intervention. Interpretative synthesis involves bringing together the research findings to build theoretical constructs, and explanatory synthesis can be used to create both explanation and theories, and tends to be pragmatic, aiming to inform decisions. Briner and Denyer (2012) observe that in the I/O psychology and management fields, narrative synthesis is most common, used to combine different aspects of a phenomena into a bigger picture or story. As the current SLR aimed to identify patterns from a range of mixed-method studies, an exploratory-narrative synthesis approach was selected as the most appropriate method.

Study 2

The second study employed a quantitative, observational study using a cross sectional survey design. The following section provides a discussion of the merits and considerations of the different design elements of the study.

Context - Field vs Lab studies

In candidate research, it can be challenging to gain access to potential participants as selection processes can be disjointed, with different systems or areas of the human resources function administering different stages (CIPD, 2021). For example, candidate details may be collected by the organisation, but the selection process, or elements of the process, may be administered by third parties. This means that accessing participants in a way that is ethical and compliant with the relevant data protection rules can be difficult. For this reason, much research is carried out in lab settings, often with university students, however this can come with its own challenges. For example, Konradt et al. (2017) compared findings from hypothetical study settings with real-world settings and found that effects tended to be higher in the hypothetical setting, suggesting that these may be the upper bounds of what could be expected to be found in the field. Conversely, Wheeler et al. (2014) compared student and non-student samples. While the differences were small, they suggested student recruited samples may result in smaller effects being identified and noted that these types of samples could have an over- or under-representation of certain demographic characteristics. This lack of consistency suggests that the appropriateness of generalising from lab studies to field studies could be highly dependent on the phenomena being investigated. There are limited studies in the field of candidate reactions that have directly compared real-world settings with lab studies and/or student samples, however where such studies have been located, differences have been found. Hiemstra et al. (2019) found differences in the favourability of selection methods between real job applicants and paid participants, suggesting that when there is more at stake, candidates are more sensitive to the choice of selection methods. Similarly, in their study on the effects of explanations on fairness perceptions, Truxillo et al. (2009) found stronger effects in field settings than lab settings, and suggested that in applicant reactions, research lab studies may underestimate effect sizes.

To balance the ethical and access considerations of using external applicants with the generalisability considerations of using lab participants, an opportunity sample of internal candidates applying to talent development schemes was selected. This group comprised of middle and senior managers within a single organisation. Internal candidates tend to be under-represented in applicant reactions research, as do participants in more senior job roles (Hausknecht et al., 2004; Konradt et al., 2017), meaning that there is clear justification for expanding the research to broader participant groups. Additionally, there were practical considerations in terms of the timing of the selection process and expected volumes, making the selected population highly appropriate for this study.

Survey design

In order to assess candidate reactions, a survey design was used. An online survey was used rather than alternative methods (such a face-to-face, telephone, or postal) as this was the most efficient method for contacting participants. Further, participants had provided an email address as part of their application to the talent development schemes; using this as the contact method meant that no further details were required (such as telephone numbers or home/work addresses), minimising the risk of data protection issues. There are some disadvantages online surveys, for example Jones et al. (2013) highlights a higher risk of non-responses compared to face-to-face or personal surveys, and a recent meta-analysis by Daikeler et al. (2020) showed that online surveys tend to have 12% lower response rates than other methods. Despite this, it was concluded that the benefits of access to participants, plus further benefits such as lower costs, ease of data compilation and fewer resources being required, outweighed any risks.

A cross sectional design was used, meaning that all of the data for the sample was collected at a single point in time; this can be contrasted with a longitudinal study where each participant is observed at multiple time points (Sedgwick, 2014). Longitudinal studies are concerned with measuring trends over time, whereas the current study investigated reactions to a specific selection process, making a cross sectional design most appropriate. As Sedgwick (2014) suggests however, cross sectional studies can be prone to non-response bias, meaning that participants who agree to take part in the study may differ from those who don't. As a mitigation to this, the demographics of the sample can be compared to the population demographics (if known) to check how representative they are.

Sample size

While the sample was an opportunity sample, it was necessary that the size was sufficient to carry out the proposed statistical analysis techniques. Based on previous selection processes carried out for the talent development schemes of interest, it was expected that approximately 3,000 individuals would apply for the schemes. Feedback surveys on the selection process are carried out each year and in previous years approximately 10% of applicants have completed the feedback survey. With regard to acceptable response rates, previous meta-analysis has put typical response rates at 34% for online surveys (Shih & Fan, 2008, as cited in Wu et al., 2022). A more recent meta-analysis of 1,071 studies (Wu et al., 2022) put the figure at 44.1% on average, however this study also found that response rates tended to reduce as sample sizes increased, with samples larger than 2,500 having a mean response rate of 20.3%. In their study of responses from higher education studies, Fosnacht et al. (2017) suggested that surveys with samples of fewer than 500 needed a 20-25% response rate to be reliable, but with larger samples reliable estimates could be achieved with 5-10% response rates. Based on previous years' data, it was expected that response rates for the current study would be at the lower end of these estimates but would be sufficient for the planned analysis methods.

Reactions Questionnaire

Study 2 used the Selection Procedural Justice Scale (SPJS, Bauer et al., 2001) to measure reactions to the selection process. This questionnaire was designed using a robust methodology with a clear theoretical basis and was designed specifically for use by

organisations. Further, it was an expectation of the questionnaire designers that researchers would amend the questionnaire to suit their specific selection contexts, making it suitably flexible for the purposes of this study. This questionnaire has been used in a large number of previous studies (including Butucescu et al., 2019; Konradt et al., 2013; LaHuis, 2005; Reeder et al., 2012; Truxillo et al., 2006) and meets accepted standards for reliability and validity (EFPA, 2008).

The SPJS was developed to measure Gilliland's (1993) procedural justice rules and measures a number of subscales: Job-relatedness Predictive; Information Known; Chance to Perform; Reconsideration Opportunity; Feedback; Consistency; Openness; Treatment; Two-way Communication; Propriety of Questions and Job-relatedness Content. The scales of Consistency, Openness, Treatment, Two-way Communication and Propriety of Questions are loaded on to one factor referred to as the 'Social' factor by Bauer et al. (2001); Job-relatedness Predictive, Information Known, Chance to Perform, Reconsideration Opportunity and Feedback are loaded onto a separate factor labelled 'Structure'. Job-relatedness Content was retained as a third factor.

The scales of Reconsideration Opportunity, Consistency, Openness, and Treatment were not included in the current study. These scales refer largely to interactions with the test administrator (for example, an item in the consistency scale asks '*Test administrators made no distinction in how they treated applicants*'); the candidates completed the selection process remotely and in an unsupervised setting, therefore there was no opportunity for them to interact directly with a test administrator. Reconsideration opportunity examines participants' perceptions of being able to challenge the decision making (an example question was '*There was a chance to discuss my test results with someone*'); this was excluded as participants did not have their outcomes at the time of completing the survey.

The current selection process comprised four exercises; in order to understand the reactions to the process as a whole and the individual elements of the selection process, the questionnaire was split into two sections. The first section covered the subscales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way Communication, Job-relatedness Content, and a new scale looking at the contribution of each element to the overall outcomes called 'Contribution to Process'. The second section covered the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way element to the overall outcomes called 'Contribution to Process'. The second section covered the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform, Feedback, Two-way the scales of Job-relatedness Predictive, Information Known, Chance to Perform Predictive, Informa
Job-relatedness Content, and Contribution to Process. This section did not include 'Two-way Communication' as all of the communications presented to candidates covered all elements of the process, therefore it would not have made sense to ask about communications for the individual elements. The second section also included the Propriety of Questions scale from the SPJS; this was not included in the first section as asking candidates about the content of the selection process as a whole would not have provided useful information. To ensure items were selected and adapted appropriately, a second researcher reviewed the adapted questionnaire prior to its use with participants.

Personality Questionnaire

A 'Big Five' questionnaire was used to measure personality as this is one of the most wellestablished personality frameworks in the literature (developed by a number of researchers, most significantly Costa & McCrae, 1985; and Goldberg, 1992). The IPIP (International Personality Item Pool, a collection of personality questionnaire items in the public domain) was selected as a source of personality questionnaire items as there were no restrictions on using these items for research purposes and no restrictions on the platform on which they could be used. Access to proprietary personality questionnaires would likely have been restricted and require a fee for use, and there may have been restrictions in the candidates' ability to access test publishers' online platforms due to data security considerations within the supporting organisation. To measure the Big Five factors of personality, a short (50 item) questionnaire containing a sub-set of IPIP items most similar to the factors as measured by the NEO-PI-R were selected (Goldberg, 1999, as cited in Goldberg et al., 2006). The NEO-PI-R (Costa & McCrae, 2008) is an established questionnaire that has been used for measuring personality in the workplace and is certified by the BPS psychological testing centre as a valid and reliable tool. The short-form questionnaire used for Study 2 has slightly lower (though still acceptable) reliability than longer questionnaires and it only allows for interpretation at the factor level rather than an exploration of sub-scales. While a longer questionnaire may have provided richer and slightly more reliable data, this needed to be balanced against the potential non-completion of participants due to the survey being overlylong. Previous studies have found that survey length can negatively impact completion numbers, particularly with online surveys (Church, 2001) and it has been recommended that surveys should not take more than 20 minutes to complete (Revilla & Ochoa, 2017). As the personality questionnaire immediately followed the reaction survey, it was concluded that the 50-item questionnaire would provide sufficient data while minimising the risk of losing participants.

Pilot Study

Prior to the main study, a pilot study was carried out with current participants of the development schemes for which the participants were applying. The purpose of the pilot was to expose both the survey and the personality questionnaire items to a sample representative of the study participants. Jones et al. (2013) note that a pilot study in survey research can be beneficial for highlighting any errors or ambiguities in the items, as well as any other issues that could impair completion of the survey. van Teijlingen and Hundley (2001) suggest that pilot studies are crucial to the design of research and increase the chances of a successful main study. They discuss some of the reasons for conducting a pilot, such as assessing the feasibility of the study, testing the adequacy of the research tools, identifying any logical problems which may occur, and assessing the data analysis techniques proposed. While it would have been beneficial to use the pilot to check the psychometric properties of both the personality questionnaire and survey, it was not expected that sufficient pilot participants would be available. However, the pilot did allow the researcher to check for issues such as any challenges accessing the survey platform and any ambiguities or problems with the questionnaire items.

Process

A further consideration for the study methodology was when to collect the reaction feedback, as fairness perceptions can be impacted by the results of the selection process. This was illustrated by Van Vianen et al. (2004); their study investigated reactions at three time points during the selection process (before the process, immediately after the process but prior to receiving feedback, and after the process was completed and feedback had been received). Their study showed that fairness perceptions are influenced by different determinants at different stages in the process, and that fairness perceptions were positively affected by feedback content (i.e., actual performance). As actual test performance may have been a confounding factor in the perceptions of candidates in the current study, the survey was administered prior to any feedback being received. This also had the benefit of ensuring that

participants did not feel compelled to take part in the study due to the outcomes of the process.

Ethical Considerations

There were a number of ethical considerations to be made during the design and execution of Study 2. The BPS code of ethics (2021) specifies four core ethical principles for practising psychologists against which the study can be considered.

With regard to the principle of Respect, due to the quantitative nature of the study there were limited issues in terms of power balance between the researcher and the participants, however communication with participants needed to be carefully considered so as not to suggest an unequal power balance.

The study used a personality questionnaire which was not dissimilar in format to the questionnaires used as part of the selection process, this meant that participants may have had concerns that this data would be used as part of the selection decision. To mitigate against this, the study was designed so that the survey was only sent out to participants after the relevant stage of the selection process had closed for all participants, which also had the benefit of reducing fatigue had the participants been requested to complete the survey alongside the selection process. However, participants still may have had concerns that this data would be used to inform later stages of the selection process, therefore it was important to provide clear messaging about how the data would and would not be used. As the survey also involved asking participants about their reactions to the selection process, they may have had concerns that providing negative feedback would negatively impact their application in some way. This risk was reduced by ensuring the voluntary nature of the study was made clear, and it was also stated explicitly that the study responses would not be linked to the application process.

In terms of the design, it may have been desirable to ask participants to complete further selection methods in addition to those required for the talent development schemes for research purposes. However, this may have caused undue additional stress for participants, and it would have been unethical to ask candidates to complete tasks that were irrelevant to

the program to which they were applying. Therefore, the study was designed around the actual selection process put in place by the supporting organisation.

With regard to the ethical principle of Competence, the lead researcher was experienced in survey design and trained in the use of many psychometrics, therefore had the required level of competence to conduct the study. However, to further ensure good practice had been adhered to, the researcher's supervisor reviewed all survey items and test instructions and all participant-facing materials, and the research plan gained ethical approval from Birkbeck, University of London prior to any contact with participants.

The ethical principle of Responsibility was particularly important when discussing the study with the supporting organisation, as it was necessary to make them aware of what they would likely gain from taking part in the study but also what the limitations were. In terms of the principle of Integrity, it was important to be open and honest with both the supporting organisation and the participants regarding the purpose of the study. To achieve this, the supporting organisation was presented with the research proposal prior to starting the study, and participants were provided with the participant information sheet and were required to provide their informed consent before beginning the survey. With regard to potential conflicts of interest, the research was carried out in the UK Civil Service, where the lead researcher was also employed. This meant that participants may have felt compelled to take part or have concerns that the responses would be used outside of the study. However, the researcher was not working for the department responsible for the talent development schemes at time of carrying out the study, and the size of the organisation made it highly unlikely that study participants would have any personal connection with the researcher.

Chapter 3 – Systematic Literature Review: What impact do individual differences have on candidate reactions to selection procedures?

Abstract

Recent years have seen a great increase in the number of selection tools and procedures available to organisations looking to identify the most talented individuals, however the impact of a poorly received selection procedure can be significant for both individuals and organisations. No known systematic literature review has been conducted to examine the impact of individual differences on candidate reactions to selection procedures, despite several calls for research in this area. This systematic review examines how individual differences in candidates going through a selection process relate to their candidate experience. 18 studies met the inclusion criteria. The results showed that there are significant limitations in the replicability of studies in this area, driven by the number of selection methods available. The study findings showed that there are some significant effects of attitudinal, biographical and personality differences on candidates' reactions, but that these effects are complex and often indirect. Implications for practice are discussed alongside suggestions for future research.

Introduction

For many decades, the focus of research in the recruitment and selection space was from the perspective of the organisation, typically aimed at identifying methods to select the most talented employees and thus improving organisations' performance. It wasn't until the early 90s that researchers began to consider the perspective of the individual taking part in the recruitment and selection process as the focus of their research. Further, the past decade has seen a huge increase in the number of possible selection methods that organisations can use as part of a selection process (Viswesvaran & Ones, 2018; Woods et al., 2020). Traditionally the most commonly used selection methods were CVs/resumes (Nikolaou & Foti, 2018; Ryan

et al., 1999), references (Viswesvaran & Ones, 2018) and interviews (Searle & Al-Sharif, 2018; Viswesvaran & Ones, 2018), then psychometrics and assessment centres became more popular (Nikolaou & Foti, 2018). Now technology has allowed for elements of automation in the selection journey (van Esch et al., 2021); interviewees and interviewers can take part in a selection interview at different times, supported by technology to pre-record responses (Suen et al., 2019). There are psychometrics being designed to measure a much wider range of knowledge, skills, abilities and other characteristics than just personality and cognitive ability (Ryan & Ployhart, 2014), and more recently even the option for virtual reality assessment centres (Dahl, 2017). These options are increasing at an accelerating rate, and we have very little research to understand what impact all this is having on candidates (Woods et al., 2020) and what the wider implications for organisations may be.

How candidates react to selection processes is an important area of study as there can be significant impacts for both the individuals themselves and the organisation. Negative reactions to selection processes can severely impact the organisations' brands and performance, as the sharing of candidate experiences is now easy to access through online platforms such as Glassdoor (Nikolaou (2021, Miles & McCamey, 2018). Negative experiences can affect candidates' confidence in applying for future job roles (McCarthy et al., 2017) and can even have an impact on their future job performance (Konradt et al., 2017; McCarthy et al., 2017).

There are also huge potential impacts for the organisation; in extreme cases, dissatisfaction can lead to legal action (Viswesvaran & Ones, 2004). There can also be risks in narrowing the pool of potential applicants if there is something in the process that is off-putting to certain groups, or people with particular traits or attributes; candidates may be put off from joining the organisation even if they are successful in the selection process. This was illustrated by Lievens and Slaughter (2016); their review of employer branding reported that impressions of the organisation measured early on in the recruitment process were related to job acceptance decisions. Similarly, Hausknecht et al.'s (2004) review of 86 applicant reaction studies concluded that candidate reactions are related to many factors associated with the organisation including organisational attractiveness, recommendation intentions and offer acceptance intentions; and that improving selection processes could help with attracting and retaining more diverse candidate groups.

When selection is with internal candidates, such as for promotion or selection for development, there can be negative impacts for the organisation such as losing talented people (Giumetti & Sinar, 2012) and reducing the engagement of employees who remain (Ford et al., 2009). This was illustrated by Ambrose and Cropanzano (2003); in their longitudinal study of university professors applying for tenure they found that negative procedural and distributive justice perceptions impacted on attitudes towards the organisation such as turnover intentions, organisational commitment, and job satisfaction.

Factors influencing candidate reactions

Gilliland (1993) laid the foundation for many researchers interested in the candidate perspective of recruitment and selection procedures, proposing a model that identifies a range of components contributing to the candidates' overall perception of the selection process. Gilliland (1993) drew on discussions of both the procedural and distributive justice of selection processes and the impact these can have on applicant perceptions. Distributive justice is concerned with the outcomes, or decisions made by the organisations (such as how many people to hire) whereas procedural justice is focused on the process leading to those decisions or outcomes. Gilliland proposed 10 procedural justice rules that must be followed in order for the selection process to be deemed fair by candidates; these rules cover aspects around formal characteristics of the process itself, factors associated with explanations given during the process, and interpersonal factors associated with the applicants' interactions with the organisation during the process.

The majority of these factors are still relevant in the current selection landscape; however, it is not uncommon for researchers to adapt the framework to suit their own context and areas of interest. For example, studies involving remote or 'low-touch' selection methods such as ability tests (as opposed to 'high-touch' methods such as face-to-face assessment centres) would typically not consider the procedural justice rule of 'assessor friendliness' as a factor, as this would clearly not be relevant when applicants have had no direct contact with an assessor. A popular measure of candidate reactions is the Selection Procedural Justice Scale (SPJS), developed by Bauer et al. in 2001 as a measure of Gilliland's procedural justice rules. Following a comprehensive development process, the SPJS measures 11 dimensions of procedural justice grouped under 3 higher-order factors, which map to the 10 procedural justice rules proposed by Gilliland (with one of Gilliland's rules, job relatedness, being split in to two subscales looking at predictiveness and content-relatedness).

In recent years there has been a growing body of research examining the determinants of reactions as well as investigating why reactions are important. In 2000, Ryan and Ployhart conducted a review of the preceding 15 years of applicant reactions research and found that there were two major streams to the research, the first looking at the influence of fairness and other related characteristics of the selection process on candidate attraction to the organisation, and the second looking at the impact of perceptions on candidates' performance in the selection process. These can be considered as the determinants of reactions and the outcomes of reactions, with the research most commonly studying perceptions of the validity, fairness and/or job relatedness of the selection process, fairness of the outcomes of the process, and test taking motivation. The authors noted that while studying the determinants of reactions is of interest, it is challenging to synthesise research findings as some studies treat the variables as correlates or predictors whereas others look at them as outcomes. They offer some observations about the determinants of perceptions, including that both the type of selection method used (e.g., ability tests, interviews) and the method of assessment (e.g., video, paper-and-pencil) influence perceptions of procedures, as do perceptions of performance, the type of job applied for, and the information given about the process. They also suggested that racial differences are sometimes found but there is a need to better understand the context of when these are found, and that social influence and individual differences were largely unexplored. A number of suggestions for areas of future research are made, in particular noting that there is a lack of research regarding the impact of individual differences.

Rationale for this review

As illustrated above, the field of applicant reaction research is complex and provides a number of avenues for further research. A recent review by McCarthy et al. (2017) sought to bring together this existing research to understand the ongoing relevance of applicant reactions and identify directions for research in the future. This review supported the need for further research in this field and suggested that new research needed to look at factors such as the impact of technological advances and to expand understanding of boundary conditions (or moderators) of applicant reactions. This could include individual differences; for example, the authors notes that there is emerging consideration of the effect of personality on reactions. While an earlier review by Hausknecht et al. (2004) found limited results for the impact of individual differences, they also suggested that "Careful study of the conditions under which

diverse populations might react differently to selection is warranted" (p25) and highlighted that further research regarding personality differences is needed. In line with McCarthy et al. (2017), an earlier review by Chan and Schmitt (2004) provided an agenda for research into applicant reactions and highlighted reactions to new technology and understanding the determinants of reactions (including applicant characteristics) as two key areas. Following McCarthy et al. (2017), a review by Woods et al. (2020) looked at the research, or lack thereof, into selection tools which make use of technology, coining the term 'digital selection processes' (DSPs). The authors of this paper note that much of what we think we know about selection is based on older processes and research needs to keep up with the newer technologies.

It is clear from these and other papers that applicant reactions continue to be an important area of study, but despite being an established concept and having numerous calls for further research there are still significant gaps in the evidence base, with both the impact of technology and the role of individual differences being two clear gaps. To understand how technology-enabled selection procedures may be received by candidates, we first need to understand the state of the research regarding individual differences and their impact on candidate reactions to selection procedures more generally.

To date, no known systematic review has brought together the existing evidence examining the impact of individual differences in candidate reactions to selection procedures. This review evaluates our current understanding in this area to enable those working in a selection context to better predict the likely impact of new selection processes on their own candidate pools.

Objectives

The primary objective of this SLR is to understand existing research regarding the impact that individual differences such as personality and biographical factors have on candidates' reactions to selection procedures. This will be addressed by the research question: what impact do individual differences have on candidate reactions to selection procedures?

Method

A systematic literature search was used to identify relevant papers, following the methodology outlined in Briner and Denyer (2012), and as applied in Donaldson-Feilder et al. (2019).

Search Strategy

During spring 2021, a literature search was carried out using databases selected for their relevance to the fields of occupational psychology and human resources, and for their breadth of literature coverage. These databases were PsychINFO, Business Sources Premier and Web of Science. In addition to the database searches, hand searching of selected journals of interest was also carried out. These were the Journal of Occupational and Organizational Psychology; Personality and Individual Differences; Journal of Personality and Social Psychology and International Journal of Selection and Assessment. Finally, forward and backward referencing of key papers was carried out to identify any further sources. The title search terms were developed through reviews of the existing literature and discussion between the research team. Search terms were developed to describe the mechanisms of interest relating to individual differences, the context, and the outcomes related to candidate perceptions. The parameters were as follows: Individual diff* (difference, differences) OR Personality OR Age OR Gender OR Ethnic* (Ethnicity, ethnic background) OR Socio* (Socio-economic status, socio-economic background) OR Disability OR Neurodivers* (Neurodiversity, neurodiverse) OR Ability AND Selection OR Applicant OR Candidate OR Test tak* (taker, taking) AND Perception OR Attitude OR Perspective or Opinion OR Reaction.

The time frame selected was post-1993; this date was selected as it was following Gilliland's seminal work developing the procedural justice model for candidate reactions (Gilliland, 1993). Mendeley was used to store and manage all studies identified in the search.

Criteria for considering studies

The CIMO framework (Context, Intervention, Mechanisms, Outcomes; developed by Denyer & Tranfield, 2009) was used to develop the inclusion and exclusion criteria. This offers a systematic framework for understanding research and is particularly well-suited to an organisational context.

The context concerned the populations of interest and the types of studies used (e.g., field/lab studies, surveys). The intervention was concerned with the selection process; this could include psychometric testing, examinations, skills testing (online or offline), interviews (groups or individual; live, telephone or online; real-time or asynchronous), assessment centres (as per interviews, any delivery method) or application forms. The mechanism related to the individual differences examined, including biographical, psychological and attitudinal traits. The outcomes were candidates' self-reported reactions to the selection procedure. This included scales and metrics based on Gillian's procedural justice frameworks such as the Test Attitude Survey (TAS, Arvey et al., 1990), the SPJS (Bauer et al., 2001), and the Social Process Questionnaire on Selection (Derous et al., 2004). It could also include qualitative data from surveys, interviews and focus groups.

The detailed inclusion and exclusion criteria for each area of the framework are shown in Table 1.

Table 1

SLR Inclusion and Exclusion criteria

	Inclusion	Exclusion:				
Context	Adult population (18+)	Under 18s				
	Can include lab studies as well as field work	Outside date-range				
	Work-based or education/training settings	Not related to work or education/training settings				
	English language only as this is the only language spoken fluently by the lead researcher.	Unpublished theses were not included				
Intervention	Selection for job roles (can include promotions or internal moves) or onto education or training programmes	Candidate is not actively involved in the selection procedure (e.g., selection through reserve lists or pre-existing information) or study is only conducted in a descriptive context (as described in Hausknecht et al., 2004)				
Mechanisms	Predictors or mechanisms are individual differences including biodata factors (gender, age, ethnicity, disability, neurodiversity, socio-economic background) and personal factors (personality traits and types, affectivity, cognitive ability/general intelligence, self-evaluations).	Does not look at individual differences as predictors, mediators or moderators.				
	Individual difference variables are predictors, mediators or moderators rather than controls.	Individual differences are only considered as controls.				
Outcomes	Any quantitative or qualitative data based on candidate reactions. Candidate reactions may be assessed through closed questions in surveys, qualitative data may also be collected through open-ended survey questions, follow-up interviews and focus groups with candidates.	Not related to reactions, e.g., success/fail rates only				

Sifting process

Papers identified in the initial search were first reviewed based on their titles to identify if they were largely related to reactions to selection processes and whether they considered any individual differences. A second researcher reviewed a random selection of 10% of the papers selected based on their titles to ensure agreement with the selection and exclusion of papers; inter-rater reliability at this title screening stage was 83%. Following this, a review of the abstracts and then full papers was carried out against the inclusion and exclusion criteria. The second researcher again reviewed a selection of 10% of the papers and at this stage 100% agreement was reached (calculating using the kappa statistic as described in McHugh, 2012). Following this initial systematic review, a further five papers were identified through wider reading and are included in the full paper extraction.

Data extraction

The data extraction was organised following the CIMO framework as described in Denyer and Tranfield (2009). Basic information about the papers was included, followed by details of the study (e.g., aim, design, population), the selection methods used, the individual differences studied and the results, including any limitations and implications for research and practice.

Data synthesis

While all the studies identified were quantitative, the variation in selection methods used and individual differences studied meant that a meta-analysis would not be suitable to synthesise the results. Rather, an explanatory narrative was used to bring together the findings, by extracting the data against the CIMO framework to understand the findings in terms of the individual differences studied.

Quality assessment

To assess individual studies, Hong and Pluye's (2019) framework was used. This looks at methodological quality, conceptual quality and reporting quality, resulting in a 1-3 score for each. As all the studies were quantitative, questions to assess methodological quality were

adapted from Snape et al., (2017) to assess factors such as sample size, quality of measures used, moderators and mediators examined, and whether the studies were field studies (rather than lab studies).

To understand the overall quality of data in relation to each evidence statement, the Nesta Standards of Evidence framework (Puttick & Ludlow, 2013) was used, with single findings being lower quality and replicated findings with alternative samples being higher quality.

Results

The initial database searches identified 298 records; following deduplication, 192 records remained for screening. The results of the searches and screening process can be seen in Figure 1, below.

Figure 1 - Flow Chart showing number of papers at each stage of the search and screening process.



Context

Participant population

On average there was a slight under-representation of females, with them making up 45% of the sample across all the studies. Seven of the studies made use of university student samples, with a range of other sectors represented across the other studies, including Front Line Services (Police, Fire and Rescue, Army), Government, Financial, Manufacturing, Teaching and Law. Sample sizes ranged from 45 to 3,984; the median number of participants was 218. Note, two studies (Bernerth, 2005; and Bernerth, et al., 2006) make use of the same sample for both studies.

Design

There was a relatively even split between lab and field studies, with 8 field studies (Bye & Sandal, 2016; Honkaniemi et al., 2013; LaHuis, 2005; Macan et al., 1994; Merkulova et al., 2014; Truxillo et al., 2006; Van Vianen et al., 2004; Visser & Schaap, 2017), 9 lab studies (Bernerth, 2005; Bernerth et al., 2006; Brenner et al., 2016; Chan et al., 1997; Hiemstra et al., 2012; Oostrom et al., 2010; Ployhart et al., 2003; Reeder et al., 2012; Saks & McCarthy, 2006;) and one that contained both (Hiemstra et al., 2019).

Location

The largest number of studies were from the US (44%; Bernerth, 2005; Bernerth et al., 2006; Chan et al., 1997; LaHuis, 2005; Macan et al., 1994; Ployhart et al., 2003; Reeder et al., 2012; Truxillo et al., 2006), followed by the Netherlands (22%; Hiemstra et al., 2012; Oostrom et al., 2010; Van Vianen et al., 2004), with the remainder coming from Central and Northern Europe (Brenner et al., 2016; Bye & Sandal, 2016; Honkaniemi et al., 2013; Merkulova et al., 2014), Canada (Saks & McCarthy, 2006) and South Africa (Visser & Schaap, 2017).

Assessment context

2,771 participants were students completing the studies for either payment or course credit, 312 were adults of working age who were paid for their participation, 445 completed as part of training, and 1,494 completed the studies as part of their application for actual job opportunities.

A summary of the context of the identified studies is shown in Table 2.

Table 2

Context of extracted papers

Paper	Context (e.g., Work including Organisation if specified)	Country	Field study or lab	N of participants	N of participants in	Gender (%
no.				invited	final study	female)
1	Applicants to teaching roles at a Norwegian public high school	Norway	Field study	Time 1:129	97	56%
				Time 2:106		
2	Applicants to clerical roles at a federal agency	US	Field study	423	291	70%
3	Applicants to manufacturing jobs in a south-eastern U.S. plant	US	Field study	Study 1: 5,056	Study 1: 3,984	Study 1: 38%
				Study 2: 210	Study 2: 194	Study 2: 28%
4	Paid participants, told study was to develop employment tests for entry-level management selection	US	Lab study	210	180	70%
5	Applicants to a vocational school for the fire and rescue service	Finland	Field study	258	218	16%
6	Applicants to a South Africa-based financial services organisation	South Africa	Field study	175	160	57%
7	Selection of career officers for the Swiss Armed Forces	Switzerland	Field study	294	294	4%
8	Students taking part as part of MBA course	Canada	Lab study	116	116	41%
9	University students at large Eastern university	US	Lab study	394	394	53%
10	Study 1: respondents from Amazon's Mechanical Turk,	Study 1: Mixed	Study 1: Survey	Study 1: 160	Study 1: 160	Study 1: 47%
	Study 2: Applicants to a Dutch entry-level legislative lawyer traineeship position	Study 2: Netherlands	Study 2: Field	Study 2: 103	Study 2: 45	Study 2: 60%
11	Undergraduate business school students	US	Lab study	Study 1: 715	Study 1: 612	38%
				Study 2: 524	Study 2: 503	
12	Applicants to the police	US	Field study	240	120	18%
13	Unemployed job seekers completing application training	Netherlands	Lab study	-	445	42%
14	Undergraduate business school students	US	Lab study	715	503	38%
15	Applicants to a range of jobs in the Netherlands	Netherlands	Field study	437	282	38%
16	Undergraduate and Masters psychology students	Netherlands	Lab study	153	153	66%
17	Paid participants from the community and psychology students	USA	Lab study	356	356 (community n=	65%
					152, student n=204)	
18	Paid participants from German universities	Germany	Lab study	106	106	61%

1 Bye & Sandal, 2016; 2 LaHuis, 2005; 3 Macan et al., 1994; 4 Chan et al., 1997; 5 Honkaniemi et al., 2013; 6 Visser & Schaap, 2017; 7 Merkulova et al., 2014; 8 Saks & McCarthy, 2006; 9 Ployhart et al., 2003; 10 Hiemstra et al., 2019; 11 Bernerth, 2005; 12 Truxillo et al., 2006; 13 Hiemstra et al., 2012; 14 Bernerth et al., 2006; 15 Van Vianen et al., 2004; 16 Oostrom et al., 2010; 17 Reeder et al., 2012; 18 Brenner et al., 2016.

Interventions

Selection Method

Cognitive ability tests were the most well-researched, with nine studies investigating this selection method. Four studies investigated personality tests and four investigated interviews (two of which were asynchronous video interviews and one a group interview). Two studies investigated assessment centres, two investigated SJTs (one of which was a multimedia SJT), two investigated CVs/ resumes, and one study investigated written tests. The full details of the selection methods used can be seen in Table 3.

Table 3

Intervention used in extracted studies

								Interventions
Paper no.	Interviews	Cognitive ability tests	Assessment centre	Personality test	Written test	SJT	CV/ Resume	Details of selection methods used
1	х							Group interviews: 2–3-hour interviews with 3-5 participants
2		х						Cognitive ability test: Speed test looking at verbal comprehension, letter/number comparisons and numerical computations
3		x	X					Cognitive ability test, assessment centre: Three cognitive ability tests: Reading comprehension, arithmetic and forms checking. Assessment centre exercises: group discussion and hands-on assembly group exercise
4		х						Cognitive ability test: Paper and pencil cognitive ability test battery, parallel forms looking at verbal comprehension, grammar, quantitative reasoning, deduction, induction, distraction.
5		х						Cognitive ability tests: Timed ability tests
6		х		х				Personality tests and cognitive ability tests: Supervised online administration of forced choice personality questions and verbal and numerical cognitive tests.
7			х					Assessment Centre: Two group exercises (leaderless group discussion and a group debate), two role plays, two oral presentations Cognitive ability test: paper-and-pencil test of verbal reasoning, numerical reasoning and abstract reasoning.
8	Х							Interviews: Interviews containing varying number of discriminatory questions
9		x						Cognitive ability test: Test described as diagnostic (of ability) non-diagnostic (related to job performance) or control (no explanation), Contextualised questions in the face valid condition to a retail role, Items were based on established measures of analytical reasoning and quantitative reasoning.
10	х							Asynchronous video interview: 3 interview questions, recorded via webcam.
11				х				Personality test: 70-item personality inventory developed for the study, measured various scales related to the Big Five
12					х			Written test: Skills test focused on police work
13							Х	Video resume & Paper resume: Filmed 46 - 60 second video resume
14				х				Personality test: 12 items from the NEO-FFI measuring Agreeableness, Openness to Experience and Neuroticism
15		х		X		х		Cognitive ability, personality test and SJT. Cognitive ability and personality measures were based on standard validated tests, SJTs varied depending on the job role.
16		X				X		Cognitive ability test Multimedia SJT: Cognitive test measuring verbal reasoning, number series and abstract reasoning, SJT: 17 video scenarios measuring managerial skills.
17		х						Cognitive ability tests: Paper and pencil tests of mechanical reasoning and numerical reasoning.
18	X							Asynchronous video interview: 3 mock interview questions representative of typical employment interviews.

Mechanisms

Three broad categories of individual differences were examined in the papers included in this review: attitudinal, biographical and personality.

Attitudinal factors

Attitudinal factors are aspects of the individual related to their thoughts, feelings and attitudes. These may have some overlap with personality traits, however for the purposes of this research they are treated as distinct from personality factors. Typical attitudinal factors include core self-evaluations, motivation, commitment and affectivity.

Biographical factors

Biographical factors typically include gender, age, ethnicity and disability. For the purpose of this review, other factors such as educational attainment, job roles applied for, and general mental ability are also considered as biographical factors.

Personality factors

Personality factors are personality types or traits that have been recognised as valid and reliable aspects of an individual to measure in existing literature (typically the Big Five factors of Neuroticism/Emotional Stability, Conscientiousness, Agreeableness, Openness to Experience and Extraversion).

Attitudinal factors were explored in six of the paper, biographical factors were explored in seven of the papers, and personality in nine papers. In the majority of papers, the individual differences studied were predictors of candidate perceptions (14 out of 18 papers). Individual differences were investigated as mediators or moderators in five of the papers. Table 4 shows the details of the mechanisms investigated, alongside any other predictors, mediators, moderators and controls.

Table 4

Details of the mechanisms investigated in the extracted studies

			Mechanis	sms	
Paper no.	Type of individual difference (personality, biographical or attitudinal)	Details of individual differences studied	Individual differences as predictors or mediators/ moderators	Other predictors, mediators and moderators	Controls
1	Personality	Personality questionnaire prior to taking part in the interview, NEO-FFI (Norwegian translation, Costa and McCrae, 1992	Predictors		Applicant age
2	Attitudinal	Employment commitment measured by one questionnaire item from Rowley and Feather (1987) Job-search self-efficacy measured with three questionnaire items developed by the Institute of Social Research (Van Ryn & Vinokur, 1991) Motivational control measured with three questionnaire items developed by Wanberget et al. (1999).	Moderators	Perceived procedural fairness as a predictor of job pursuit intentions	Pre-test attitudes Perceived performance
3	Biographical	Sex, race, employment status	Predictors		Pre-test attitudes
4	Biographical Attitudinal	Race Test Taking motivation: 7 item measure adapted from the Motivation Scale in Arvey et al.'s (1990) Test Attitude Survey	Race is both a predictor and a control Test taking motivation is a mediator	First test performance was a mediator.	Race Previous test performance
5	Personality	Personality using the Finnish version of the PRF (Niitamo,1997), translated from Jackson (1999)	Predictors		Self-evaluated performance, gender, age

6	Biographical	Ethnic origin, gender, educational level,	Predictors	Position applied for as a predictor	
		age			
7	Personality	Big Five personality measured by a	Predictors		Self-rated and actual performance,
	Attitudinal	shortened version of Schallberger and			candidates' language
		Venetz's minimal redundant scales (1999)			
		Core self-evaluation questionnaires			
		measuring locus of control and self-			
		efficacy (Krampen 1991), and self-esteem			
		(Bandura, 1987).			
		Trait affect measured by modified version			
		of the PANAS (Watson et al., 1988)			
		General mental ability measured by tests of			
		verbal, numerical, and abstract non-verbal			
		reasoning (SHL, 2006)			
8	Biographical	Gender	Mediator	Interviewer gender as a moderator	
9	Biographical	Race	Moderator	Self-reported stereotype threat; racial	Level of race identification
				identity	
10	Personality	Study 1: Personality using 50 items from	Predictors		
		the International Personality Item Pool			
		(Goldberg, 1999)			
		Study 2: Personality using Dutch 224-item			
		personality questionnaire (G5R; Oostrom,			
		et al., 2010)			
11	Biographical	Gender	Predictor	Outcome of phase 1 (randomly assigned	Grade point average
				'met' or 'failed')	
12	Personality	Big Five measured by the Mini-Markers	Predictors		Gender, Test Taking experience, test
		scale (Saucier, 1994)			score, process fairness and outcomes
					fairness
13	Biographical	Ethnicity	Predictor	Ethnic identity and language proficiency	Education
14	Personality	Personality, 3 scales of the FFI	Predictor	Selection decision as a moderator	
	Attitudinal	(Agreeableness, Openness to Experiences,			
		Neuroticism, Costa & McCrae, 1992),			
		Test taking self-efficacy, measured by 4			
		item scale (Bauer et al., 1998)			

15	Personality	Openness to experience, measured by Dutch version of the NEO-PI-R (Hoekstra et al., 1996)	Predictor	Perceived performance as a predictor Test beliefs as a mediator	Previous test experience
16	5 Personality Attitudinal	Personality: 224-item computer-based personality questionnaire developed by Dutch consultancy GITP (Koch, 1998), Core self-evaluations: 12-item Core Self Evaluation Scale (Judge et al., 2003) Subjective well-being: 5 item Satisfaction With Life Scale (Diener et al.,) Test taking self-efficacy: 3 items from Motivated Strategies for Learning Questionnaire (Pintrich and De Groot, 1990)	Predictors	Computer anxiety as a predictor: 5 items, adopted from Heinssen et al. (1987) Test anxiety as a predictor: 7 items, adopted from Cassady and Johnson (2002)	Age, gender, job experience, test experience, self-assessed test performance
17	7 Attitudinal	Work locus of control: 8 items from Spector's (1988) 16-item Work Locus of Control measure	Moderator	Predictors: Self-assessed performance: 2 items from Brutus and Ryan (1998) Test experience: two items adapted from Wiechamn and Ryan (2003) Prior success: two items adapted from Wiechamn and Ryan (2003) Job experience: yes/no question Job familiarity: 4 items Moderator: Implicit theories: Six items adapted from Dweck (2000), assessing individual beliefs that one's ability is malleable and capable of change	
18	Personality	Big Five personality measure using unipolar descriptors (Ostendorf, 1990)	Predictor	Predictors: Computer self-efficacy: 6 items scale adapted from Cassidy and Eachus (2002) and Spannagel and Bescherer (2009) Job interview self-efficacy: 5 items scale selected from Sieverding and Ortner (unpublished manuscript)	Self-rated performance, gender, age, mother tongue, prior experience with selection methods, and how seriously the participants took the interview

Outcomes

The outcomes examine both the way in which candidate reactions were measured and the relationships between individual differences and candidate reactions.

With regard to the measures used, a number of studies based their measures of candidate reactions on the TAS (Arvey et al., 1990) and the SPJS (Bauer et al., 2001). Four studies based some of their measures on the TAS and four on the SPJS. The work of Smither et al. (1993, e.g., as cited in Chan et al., 1997) was drawn on in three studies (one of which also used the TAS and one also used the SPJS) and two studies drew on the work of Chan et al. (1998). The work of Elkins and Phillips (2000) was drawn on for the two Benerth studies (Bernerth, 2005; Bernerth et al., 2006); however as noted previously these studies made use of the same sample and methodology for both. The remaining studies drew on a range of other previous literature or developed their own measures.

Table 5 shows the outcomes reported by the extracted studies. The relationships reported are then discussed, grouped by type of individual differences to aid interpretation. As all the studies identified were quantitative, the outcomes are discussed in terms of statistical relationships found, which includes direct correlational relationships, interaction effects, tests of difference and effect size, and the variance explained by the individual differences studied.

Table 5

Details of the outcomes measured in the extracted studies

		Outcomes
Paper	Candidate reactions measured (models and scales)	Relationship between individual differences and outcomes (such as correlations, group differences, predictive models)
no.		
1	New 20 item questionnaire measuring fairness perceptions, grouped as 'social fairness' and 'structural fairness'. Some items adapted from Bauer et al.'s SPJS scale (2001)	Significant interaction effect of high Neuroticism and low Extraversion on social justice and structural justice ($b = .21, p < .05$) Significant interaction effect of high Extraversion and high Agreeableness on social justice but not structural justice ($b = .21, p < .05$) < .05) Openness to experience was not a significant predictor of social justice or structural justice.
2	Perceived performance measured by three questionnaire items (LaHuis & MacLane, 2001) Perceived procedural fairness measured by four items questionnaire adapted from Chan et al. (1998), Job-pursuit intentions measured with one questionnaire item from Smither et al., (1996)	Perceived procedural fairness correlated positively with test-taking motivation ($r = .39$, $p < .05$); employment commitment ($r = .13$, $p > .05$) and job search self-efficacy ($r = .21$, $p > .05$). Job-search self-efficacy and motivational control moderated the relationship between perceived procedural fairness and job- pursuit intentions (product terms were .14 and16; 95% CI ranged from .01 to .26 and29 to04 respectively); there were stronger relationships for those with higher levels of job search self-efficacy and lower motivational control. Employment commitment was not a moderator (product term estimated at05, 95% CI ranged from17 to .07) Higher levels of job-search self-efficacy were associated with stronger relationships between perceived procedural fairness and job-pursuit intentions.
3	Four questionnaire variables to assess applicants' perceptions of cognitive ability adapted from previous research (Kluger & Rothstein, 1993; Rynes & Connerley, 1993; Smither et al., 1993) Three variables measuring responses to the selection process in general, the job and the organisation.	Applicants' background characteristics were not strongly associated with their satisfaction with the selection process, liking of the job, or organisational attractiveness, accounting for less than 1% of the variance. Background characteristics accounted for 3% or less of variance in study 2 (characteristics were not reported separately due to practical non-significance)
4	Face validity, measured with a 4-item questionnaire developed from Smither et al. (1993). Test taking motivation measured with 7 items adapted from the Motivation scale in Arvey et al.'s TAS (1990).	Black participants reported lower face validity ($t = 1.85$, $p < .05$) and lower test-taking motivation than white participants ($t = 2.98$, $p < .05$). Effect sizes for face validity and test-taking motivation were28 and45. The relationship of race and motivation was mediated by first test performance (beta =14), as was the relationship between race and face validity (beta =11). The correlation between face validity perceptions and motivation was .45 ($p < .05$) and motivation fully mediated the relationship between face validity and test performance.

5 Applicant reaction scale adapted from Chan et al. (1998), measuring face validity, predictive validity and fairness. Ratings given on a 7-point rating scale.	Latent profile analysis identified 4 personality types. Personality type explained fairness perceptions when controlling for gender ($F(3) = 4.097$, $p < .01$) but not face validity or predictive validity perceptions. 'Overcontrolled' personality types (characterised as above average Neuroticism, low Extraversion and Agreeableness) had less favourable fairness perceptions ($M = 4.02$, $SD = 1.08$) than Bohemian (characterised as low Extraversion and low Conscientiousness; $M = 4.80$, $SD = .99$) or Resilient (characterised as below average Neuroticism, high Conscientiousness, above average Extraversion and Agreeableness; $M = 4.78$, $SD = 1.09$). Bohemian had the most favourable fairness perceptions.
 6 Short version (20 items) of the TAS developed by Smith (1997) measuring motivation (5 items); lack of concentration (3 items); belief in tests (4 items); comparative anxiety (5 items); external attribution (1 items); future effects (2 items) 	No significant effect of Age on TAS scores for the cognitive ability test [Test Motivation (TM) = $\chi 2(1,160) = 1.44$, $p = .23$; Test Attitude (TA) = $\chi 2(1,160) = 2.9$, $p = .09$] or the personality test [TM = $\chi 2(1,160) = .44$, $p = .51$; TA = $\chi 2(1,160) = 3.17$, $p = .08$] No significant effect of Education level on TAS for cognitive ability tests [TM = $\chi 2(1,160) = 5.43$, $p = .07$; TA = $\chi 2(1,160) = 3.34$, $p = .19$] and the personality test [TM = $\chi 2(1,160) = 2.72$, $p = .51$; TA = $\chi 2(1,160) = 1.52$, $p = .47$] The African group had significantly more positive general attitude toward both tests {TA= $\chi 2(1,160) = 7.16$, $p = .007$; TA = $\chi 2(1,160) = 7.73$, $p = .005$].
7 Questionnaire concerning reactions to the AC; five dimensions each measured with four items (face validity, measurement quality, controllability, absence of strain, quality of administration) using a 6-point rating scale (Kersting, 2010).	Agreeableness correlated significantly with face validity, measurement quality, controllability, and quality of administration ($r = .18, .22, .19$ and .15 respectively, all $p \le .1$) Conscientiousness correlated positively with face validity ($r = .12, p < .05$) and measurement quality ($r = .18, p < .01$). Emotional Stability was significantly related to controllability ($r = .12, p < .05$) and measurement quality ($r = .18, p < .01$). Emotional Stability was significantly related to controllability ($r = .12, p < .05$) and measurement quality ($r = .18, p < .01$) and quality of administration ($r = .19, p < .05$). No significant correlations were found for Extraversion or Openness to Experience. GMA was not related to applicant perceptions. Locus of control was related to all perceptions: face validity ($r = .21, p < .01$), measurement quality ($r = .26, p < .01$), controllability ($r = .13, p < .05$), absence of strain ($r = .16, p < .01$), and quality of administration ($r = .21, p < .01$). Self-efficacy was significantly related to all perception variables: face validity, measurement quality, controllability, absence of strain, and quality of administration ($rs = .25, .21, .31, .25,$ and .16, respectively, all $ps < .01$), and quality of administration ($r = .27, p < .01$). Positive affectivity was positively related to face validity ($r = .14, p < .05$), measurement quality ($r = .18, p < .01$), and quality of administration ($r = .27, p < .01$). Positive affectivity was positively related to face validity ($r = .14, p < .05$), negative affectivity was negatively related to controllability ($r = .13, p < .05$), absence of strain ($r = .19, p < .01$), and quality of administration ($r =13, p < .05$). Agreeableness was a significant predictor of perceived measurement quality ($\beta = .24, p < .01$) and a marginally significant predictor of perceived controllability. Emotional Stability was a significant predictor of perceived controllability ($\beta = .23, p < .01$). Locus of control was the strongest predic

8	Reactions to the interviewer and the interview, measured by two questionnaire items each and based on items used in previous studies (Liden & Parsons, 1986; Turban & Dougherty, 1992) Intentions to pursue employment measured by two items, job acceptance intention measured by two items, intention to recommend measured by two items.	Females had lower expectations of receiving a job offer ($p < .05$) and lower perceptions of the treatment of employees ($p < .05$) but there was no effect of gender on reactions to interview questions.
9	Five item face validity scale (Smither et al., 1993); test taking motivation adapted from the TAS (Arvey et al., 1990); 10 item comparative anxiety scale (Arvey at al., 1990); domain identification, 3 items (modelled after Steele & Aronson, 1995)	Stereotype threat interacted with face validity and race, when individuals highly identified with their racial group; three-way Stereotype Threat × Face Validity × Race interaction was significant, $F(2, 176) = 3.57$, $p < .05$. There was a significant difference in test performance for Black participants, when they completed the generic test (non-face valid) they performed significantly better in the control condition than the diagnostic ($t(21) = 2.16$, $p < .05$), whereas in the face valid condition, they scored significantly better in the non-diagnostic than the diagnostic condition ($t(22) = 1.96$, $p > .05$). There were no significant differences in test performance across the conditions for White participants.
10	Items measuring fairness (4 items adapted from Kluger & Rothstein,1993) and procedural justice dimensions (adapted from Smither et al.,1993; 4 items measuring face validity, 5 items measuring perceived predictive validity, 4 items measuring opportunity to perform) using 1-5 rating scale.	Study 1: Extraversion related positively to perceptions of video applications predictive validity ($r = .19$, $p = .02$) and opportunity to perform ($r = .25$, $p = .002$). No support was found for an interaction effect of Neuroticism and Extraversion. Study 2: Extraversion was only related to opportunity to perform ($r = .32$, $p = .03$). No support was found for the interaction effect between Neuroticism and Extraversion.
11	Five item procedural justice measure and four item distributive justice measure (adapted from Elkins and Phillips, 2000)	Gender*selection decision accounted for variance in both procedural (r sqr =.01, $p <.001$) and distributive justice (r sqr = .02, $p < .001$). The relationship between selection decision favourability and organisational justice perceptions was moderated by applicant gender. Rejected female applicants rated both procedural and distributive justice significantly lower than rejected males ($t = 2.33$, $p < .05$; $t = 2.61$, $p < 0.01$ respectively).
12	Social Fairness (20 items from Bauer et al.'s SPJS (2001)), Structure Fairness (17 items from the SPJS (Bauer et al., 2001)), Outcome Fairness (2 items); test taking self-efficacy (3 items), likelihood of job offer (1 item); perceived employee relations (3 items), turnover perceptions (3 items).	Neuroticism was negatively correlated with all the applicant perceptions variables except for likelihood of getting the job ($rs = .22$ to 36 , $ps < .05$). Agreeableness was correlated with social fairness ($r = .30$, $p < .01$), turnover perceptions and employee relations ($rs = .18$ and .40, respectively, $ps < .05$). Openness to experience was related only to social fairness ($r = .25$, $p < .01$). Agreeableness was positively related to perceived employee relations, ($b = .43$, t = 3.02, $p < .01$). Unhypothesised correlations were also found, Conscientiousness was related to social fairness ($r = .21$, $p < .05$) and perceived employee relations ($r = .28$, $p < .01$). Hierarchical regression results: Neuroticism was negatively related to social fairness ($b = .26$, $t = 2.24$, $p < .05$). Agreeableness was positively related to perceived employee relations ($b = .43$, $t = 3.02$, $p < .01$).

13	Overall fairness adapted from Kluger and Rothstein (1993; 4 items), face validity (Smitheret al., 1993; 4 items), perceived predictive validity (Smither et al., 1993; 5 items), opportunity to perform (Bauer et al., 2001; 4 items).	No ethnicity effects we found for the perceived overall fairness of paper and video interviews. Ethnic minority applicants perceived the predictive validity of both resumes higher than majority applicants (F(2, 413) = 4.58, $p = .01$) There was a significant ethnicity effect for face validity of the paper resumes (F(2,413) = 6.33, $p = .00$), Turkish/Moroccan applicants rated the face validity of paper resumes as lower compared with Dutch applicants. Significant correlations were found for minorities' identity and fairness perceptions (r(video) = .11, $p < .05$; r (paper) = .11, $p < .05$), the perceived predictive validity of paper resumes ($r = .19, p < .01$), face validity of video resumes ($r = .19, p < .01$), and opportunity to perform of paper and video resumes (r (paper) = .12, $p < .05$; r (video) = .15, $p < .01$) but ethnic identity did not significantly moderate the ethnic minorities' perceptions of paper versus video resumes. Ethnic minorities who identified more strongly with their ethnic group had more positive applicant perceptions of both video and paper resumes (F(16, 648) = 1.83, $p = .02$).
14	5-item measure of procedural justice (adapted from Elkins & Phillips, 2000) and 4-item measure of distributive justice (also adapted from Elkins & Phillips, 2000)	Procedural justice was correlated with Agreeableness ($r = .14$, $p < .01$) and Openness to Experience ($r = .11$, $p < .05$). Distributive justice was also correlated with Agreeableness ($r = .12$, $p < .01$) and Openness to Experience ($r = .09$, $p < .05$). Neuroticism was negatively related with distributive justice ($b =07$, $p < .05$), but no association was found with procedural justice. There was an interaction of test-taking self-efficacy and selection decision for both procedural ($r \ sqr = .01$, $p < .05$) and distributive ($r \ sqr = .02$, $p < .001$), indicating that individuals with high test-taking self-efficacy had more positive justice reactions when they were accepted and more negative reactions when they were rejected.
15	Test beliefs measures with 5 items adapted from the TAS (Arvey et al, 1990); 3 items measuring perceived job relatedness; 1 item measuring perceived performance; pre- and post-feedback measures adapted from Truxillo and Bauer (1990); 4 items measuring perceived feedback treatment; 1 item measuring feedback content and 3 items measuring job attractiveness (Truxillo & Bauer, 1999).	Openness to experience was positively related to test beliefs ($r = .13$, $p < 0.01$). Indirect effects from Openness to Experience through test beliefs to perceived performance, pre- and post-feedback perceptions, and perceived feedback treatment were significant (effect sizes ranging from .02 to .04, $p < .05$).

16	Face validity and perceived predictive validity measured with 3 items each	No significant relationship between test anxiety and job relatedness of either test.
	from Smither et al. (1993)	Computer anxiety negatively related to face validity of the SJT ($r =20$, $p < .05$)
	Self-assessed test performance: 4 items from Wiechmann and Ryan (2003)	Core self-evaluations was positively related to the perceived predictive validity of the cognitive ability test ($r = .19, p < .05$)
		and the face validity of the SJT ($r = .20$, $p < .05$),
		Subjective well-being positively related to the face validity and perceived predictive validity of the of the SJT ($r = .17, p < .05$;
		r = .17, p < .05)
		No significant correlations between test-taking self-efficacy and face validity or perceived predictive validity of ability test and the SJT.
		Ability test: Agreeableness positively related to validity ($r = .20, p < .05$) and perceived predictive validity ($r = .22, p < .05$),
		Emotional Stability positively related to face validity ($r = 27$, $p < .01$) and perceived predictive validity ($r = .26$, $p < .01$), and
		Openness to Experience was positively related to face validity ($r = .27$, $p < .01$) and its perceived predictive validity ($r = .29$, $p < .01$).
		Openness to experience was also significantly related to the face validity of the SJT ($r = .19, p < .05$).
		Regression: for face validity of the cognitive ability test, Openness to Experience ($b = .20$, $t = 2.18$, $p < .05$) and Emotional
		Stability ($b = .19, t = 1.99, p < .05$) remained following a stepwise procedure. For perceived predictive validity of the cognitive
		ability test, only Openness to Experience ($b = .19$, $t = 2.11$, $p < .05$) explained additional variance up to and beyond the control
		variables. For face validity of the SJT, Openness to Experience ($b = .19$, $t = 2.10$, $p < .05$) and core self-evaluations ($b = .19$, $t = 0.19$, $t = 0.10$,
		= 2.03, p < .05) explained additional variance. For perceived predictive validity of the SJT, only subjective well-being ($b = .19$,
		t = 2.18, p < .05) explained additional variance.
17	Perceived predictive validity and face validity items adapted from Bauer et	Neither implicit theories ($b =05$, $p = .39$) nor self-assessed performance ($b = .10$, $p = .07$) was significantly related to face
	al.'s SPJS (2001)	validity perceptions of the numerical test. For incremental theorists, there was a positive and significant relationship with face
		validity ($b = .18, p < .01$).
		Neither implicit theories ($b =05$, $p = .39$) nor self-assessed performance ($b = .10$, $p = .07$) was significantly related to face
		validity perceptions for the numerical tests.
		For incremental theorists, the relationship between face validity perceptions and self-assessed performance was positive and $\frac{1}{2}$
		significant $(b = .18, p < .01)$.
		For entity theorists, self-assessed performance and face validity perceptions were not related to one another ($b =01$, $p = .91$)
		None of the interaction terms involving locus of control were significant. Face validity-numerical ($b = .05$, $p = .30$), perceived
		mechanical $(b =01, n =00)$, $p = .27$, race validity-inechanical $(b =10, p = .07)$, perceived predictive validity-
		$\frac{1}{100}$
18	Perceived usefulness and perceived ease of use (6 items each inspired by Davis	Individual differences (i.e., personality) accounted for 14% of the variance in attitudes ($p < .01$).
- 0	1989)	The interaction between perceived usefulness and openness accounted for another 3%.
	12 items questionnaire looking at appropriateness, chance to perform, fairness,	Higher Conscientiousness was related to positive attitudes towards the interview ($r = .25$, $p < .01$), however this was explained
	and informativeness, validated against Steiner and Gilliland (1996)	by perceived usefulness and ease of use.

Notes: SPJS = Selection Procedural Justice Scale, CI = Confidence Interval, TAS = Test Attitude Survey, M = mean, SD = standard deviation, AC = Assessment Centre, GMA = General Mental Ability

Attitudinal differences

There are a broad number of factors that could be considered as 'attitudinal' in the papers; these covered measures of employment commitment, self-efficacy, motivational control, test taking motivation, core-self evaluations, trait affectivity, well-being, and locus of control. Collectively, these can be considered to measure positive attitudes and emotions experienced by candidates (e.g., being motivated to succeed, feeling in control, being committed to gaining employment etc.). Six papers examined attitudinal differences.

In LaHuis (2005), job-search self-efficacy and motivational control were found to moderate the relationship between perceived procedural fairness and job pursuit intentions. Chan et al., (1997) found that test taking motivation was positively correlated with face validity and mediated the relationship between face validity and test performance. Several attitudinal factors were investigated in Merkulova et al. (2014); they found that locus of control, self-efficacy and affectivity were related to the majority of perception variables studied (face validity, measurement quality, controllability, absence of strain, quality of administration). In Bernerth et al. (2006), test-taking self-efficacy was found to interact with selection decision on perceptions of procedural and distributive justice. Oostrom et al. (2010) found that coreself evaluations and subjective well-being were related to perceived predictive validity and face validity (but these results varied depending on the selection method used). They did not find a significant relationship for test taking self-efficacy. Reeder et al. (2012) found that attitudes about how malleable attributes were felt to be were related to face validity, but they did not find any significant relationship between locus of control and candidate perceptions.

There was a reasonable degree of consistency in the results for attitudinal differences, with most showing some significant relationship between the attitudinal variables investigated and candidate perceptions. However, direct replicability was limited due to the number of different attitudinal factors investigated and the different selection methods used. Further, one study (Oostrom et al., 2010) found a different pattern of results depending on the selection method used (in this case a cognitive ability test and an SJT).

Biographical differences

To interpret the findings regarding biographical differences, the results are discussed by the two most studied areas first (race/ethnicity and gender), followed by a summary of other areas investigated.

Race/Ethnicity

Five papers investigated the impact of race on reactions, however some caution must be taken in comparing these findings as the comparison groups identified in the studies differed depending on the country where the research took place. Macan et al. (1994) did not find any significant associations between background characteristics, which included race, and candidate satisfaction with the selection process. Chan et al. (1997) found that Black candidates had more negative perceptions of the face validity than White candidates and that this relationship was mediated by test performance. In Visser and Schaap (2017), African candidates had significantly more positive general attitudes towards the selection tools used compared to other ethnic groups. Ployhart et al. (2003) found a three-way interaction effect with race, face validity and stereotype threat, but only when individuals identified strongly with their racial group. In Hiemstra et al. (2012) predictive validity perceptions were more favourable for ethnic minority applicants than the majority group, and ethnic minorities who identified more strongly with their ethnic group had more positive perceptions.

These studies collectively do not suggest that there is a strong direct relationship between candidate ethnicity and reactions to selection processes, however there may be some more complex relationships associated with the specific ethnic group and/or the selection method used. The consistent findings between Hiemstra et al. (2012) and Ployhart et al., (2003) suggest that identification with race is likely to play a role, however it is not clear from the papers identified in this review whether this is a consistent finding across a range of selection processes.

Gender

Four studies investigated the impact of gender. As previously noted, Macan et al. (1994) examined the impact of background characteristics collectively (which included gender) and did not find any significant relationships with candidate reactions. Saks and McCarthy (2006) found that females had lower expectation than males of receiving a job offer and rated the perceived organisational treatment of employees as lower than males, however there were no

gender differences on the impact of discriminatory interviewer questions. Bernerth (2005) found that there was an interaction effect between gender and selection outcomes, with rejected female applicants reporting lower perceptions of procedural and distributive justice than males. Visser and Schaap (2017) reported investigating gender differences, but no results were stated for this.

Other areas examined were Age and Educational level (Visser & Schaap, 2017) and Employment status (Macan et al., 1994). The papers did not report any significant relationships between these variables and candidate reactions to the selection processes.

To summarise the findings for biographical factors, there is limited evidence to suggest that there are direct relationships between variables such as ethnicity, gender, age, education level and employment status and reactions to selection processes. There may be some indirect relationships, however these were not consistently reported in the papers identified.

Personality

All the papers looking at personality used a Big Five, or Big Five-related, measure to assess personality traits. This means that the findings can be reviewed in relation to each of the five factors of the model.

Agreeableness

In Merkulova et al. (2014), Agreeableness was found to correlate significantly and positively with a range of procedural justice perceptions (face validity, measurement quality, controllability and quality of administration). In the same study, it was also found to significantly predict perceived controllability and perceived measurement quality. Agreeableness was positively related to social fairness perceptions and perceptions regarding turnover and employee relations in Truxillo et al. (2006). In Bernerth et al. (2006) Agreeableness was correlated with short measures of both distributive and procedural justice. Positive relationships were found between Agreeableness and face validity and perceived predictive validity in Oostrom et al. (2010). Interaction effects of Agreeableness as shown in Bye and Sandal (2016) and Honkaniemi et al. (2013) are described below in relation to Extraversion

Conscientiousness

Conscientiousness was found to correlate positively with face validity and measurement quality in Merkulova et al. (2014). In Truxillo et al. (2006) Conscientiousness was related to perceptions around social fairness and perceived employee relations. In Brenner et al. (2016) Conscientiousness was related to positive attitudes, however this was explained by other perception variables in the regression analysis. Honkaniemi et al. (2013) found that Conscientiousness played a role in the fairness perceptions of different 'types' along with other elements of the Big Five (described below in relation to Extraversion).

Extraversion

A significant interaction effect of low Extraversion and high Neuroticism was found in relation to perceptions of social justice and structural justice in Bye and Sandal (2016). There was also a significant interaction effect of high Extraversion and high Agreeableness on social justice perceptions (Bye & Sandal, 2016). Interaction effects were also found by Honkaniemi et al. (2013) who described personality 'types' made up of combinations of different trait scores. They found that types characterised by above average Neuroticism, low Extraversion and low Agreeableness had less favourable fairness perceptions than types characterised as low Extraversion and low Conscientiousness; or types characterised as below average Neuroticism, high Conscientiousness and above average Extraversion and Agreeableness. In Merkulova et al., (2014) no correlations were found between Extraversion and various aspects of procedural justice. In Hiemstra et al. (2019) Extraversion was positively related to predictive validity perceptions and opportunity to perform, but the predicted interaction with Neuroticism was not found.

Neuroticism

Interaction effects of Neuroticism as shown in Bye and Sandal (2016) Honkaniemi et al. (2013) and Hiemstra et al. (2019) are described above in relation to Extraversion. In Merkulova et al. (2014), Neuroticism (Emotional Stability) was found to correlate with controllability, absences of strain and quality of administration factors of procedural justice and was a significant predictor of perceived controllability. Truxillo et al. (2006) found Neuroticism was negatively correlated with various procedural and distributive justice perceptions; in their regression analysis Neuroticism was negatively related to social fairness. Neuroticism was found to correlate negatively with distributive justice but not with

procedural justice in Bernerth et al. (2006). In Oostrom et al. (2010) Neuroticism (Emotional Stability) was related to face validity and perceived predictive validity.

Openness to Experience

Openness to Experience was not found to be a significant predictor of social justice or structural justice in Bye and Sandal (2016), nor was it found to correlate with various aspects of procedural justice in Merkulova et al. (2014). In Truxillo et al. (2006) Openness to Experience was related to perceptions of social fairness, but not to any of the other perception variables investigated in the study. In Oostrom et al. (2010) Openness to Experience was positively related to face validity and perceived predictive validity, and explained additional variance beyond control variables in the regression analysis. In Brenner et al. (2016) Openness to Experience interacted with perceptions of usefulness to explain significant variance in the regression analysis.

To summarise, Agreeableness and Neuroticism (sometimes measured as the reverse of Neuroticism, Emotional Stability) had significant relationships with numerous perception variables, both through direct relationships and interaction effects. While relationships were also found for Extraversion, Conscientiousness and Openness to Experience, these did not appear in as many of the studies as Agreeableness and Neuroticism. However, taken as a whole, the results show that there are significant relationships between all of the Big Five factors of personality and candidate perceptions of selection processes.

Assessment of Quality

The results of the quality analysis are shown in Table 6 below and Table 7 presents a summary of the evidence statements.

In general, the papers tended to use reliable measures to assess candidates' reactions and employed appropriate statistical techniques with adequate sample sizes. The vast majority of papers used acceptable quality measures to look at reactions, drawing from established questionnaires such as TAS (Arvey et al., 1990) and the SPJS (Bauer et al., 2001). Where individual differences were measured (as opposed to differences that were categorical, such as ethnicity or gender), the measures tended to also be of acceptable quality. All papers

looking at personality used an established measures of the Big Five (e.g., the NEO- FFI used in Bye & Sandal, 2016; the PRF used in Honkaniemi et al., 2013; the Mini Markers scale used in Truxillo et al., 2006). Likewise, papers exploring attitudes used measures of acceptable reliability and validity (e.g., the PANAS in Merkulova et al., 2014; and the Core Self Evaluation scale in Oostrom et al., 2010).

The generalisability of the results to real-world workplace settings was a factor for the majority of the papers with several being lab studies and/or using student samples, meaning the results may not always generalise to real, high stakes selection contexts. Where field studies were conducted, these tended to be with groups applying to specific job roles such as Teachers (Bye & Sandal, 2016), Fire and Rescue Service (Honkaniemi et al., 2013), Armed Forces (Merkulova et al., 2014), again meaning that the findings may not generalise to all applicant groups.

Table 6

Quality Ratings of extracted papers

	Methodological quality								quality	Reporting Quality				Score	Score	Score	Total
													Methodologic	Conceptual	Reportin	Score	
														al (1-3)	(1-3)	g Quality	1
																(1-3)	
Paper	Was	Was the	Can the	Were the	Were the	Were the	Were the	Are there	Does the study	Is the	Are the	Is there	Are the				ł
no	the	sample	results be	measures of	reaction	moderators	statistical	clear	provide new	hypothesis/aim/objecti	finding	adequate	findings				ł
	study	size	generalised	individual	measures used	and	tests used to	translatabl	understanding	ve of the study clearly	s made	discussion	discussed				ł
	a field	adequate	to other	difference (if	of acceptable	mediators	assess the	e	?	described?	explicit	of the	in				ł
	study	for the	populations	used) of	quality? (e.g.,	examined	main	concepts?			?	evidence	relation				ł
	(rathe	statistica	?	acceptable	reliability of	(if	outcomes					both for	to the				ł
	r than	1		quality? (e.g.,	questionnaires,	included)	appropriate					and against	original				ł
	a lab	methods		personality	appropriatenes	appropriate	?					the	research				ł
	study)	used?		questionnaires	s of questions	?						researcher'	question?				ł
)	asked)							s					ł
												arguments?					I
1	Y	N	Y	Y	Y	-	N	Y	Y	Y	Y	Y	Y	2	3	3	8
2	Y	Y	Y	N	N	Y	Y	N	Y	Y	N	Y	Y	1	2	1	4
3	Y	Y	Y	-	Y	Y	Y	Y	N	Y	Y	Y	Y	3	2	3	8
4	Ν	Y	N	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	2	2	3	7
5	Y	Y	Y	Y	CT	Y	Y	Ν	Y	Ν	Y	Ν	Y	2	2	1	5
6	Y	Y	Y	-	Y	-	Y	N	Y	Y	Ν	N	N	3	2	1	6
7	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	3	3	3	9
8	Ν	Y	N	-	N	Y	Y	N	N	Y	Y	N	Y	1	1	2	4
9	Ν	Y	N	-	Y	Y	Y	N	Y	Y	Y	Y	Y	2	2	3	7
10	Y	N	Y	Y	Y	-	Y	Y	N	Y	Y	Y	Y	3	2	3	8
11	Ν	Y	Y	-	Y	Y	Y	Y	Y	Y	Y	N	Y	2	2	2	6
12	Y	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	Y	Y	3	2	3	8
13	N	Y	Y	-	Y	Y	Y	N	Y	Y	Y	Y	Y	2	2	3	7
14	Ν	Y	Y	Y	Y	Y	Y	Y	N	Y	Y	N	Y	2	2	2	6
15	Y	Y	Y	Y	СТ	Y	Y	Y	Y	Y	Y	Y	Y	3	3	3	9
16	Ν	Y	N	Y	Y	-	Y	Y	N	Y	Y	Y	Y	2	2	3	7
17	Ν	Y	Y	Y	Y	Y	Y	N	N	N	Ν	Y	Y	2	1	1	4
18	Ν	Y	N	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	2	3	3	8
Evidence statements

Evidence Statement	Rating	Reason
There is a significant impact of attitudinal factors on candidates' reactions to selection processes	Promising	Several papers of mixed quality. The majority of results are in the same direction but there is very little replicability of methodologies
There is a significant impact of biographical factors on candidates' reactions to selection processes	Promising	Mix of quality papers, however most are fair or good. The majority of results are in the same direction but there is very little replicability of methodologies
There is a significant impact of personality on candidates' reactions to selection processes	Promising	Lots of papers, the majority are of good quality. The majority of results are in the same direction but there is very little replicability of methodologies

Discussion

This review examined the impact of individual differences on candidates' reactions to selection processes. Despite the preponderance of research focused on selection procedures (Nikolaou & Foti, 2018), only 18 papers met the inclusion criteria since Gilliland (1993) first proposed his model of procedural justice in selection processes almost 30 years ago. Previous calls for research identified the need to understand the impact of determinants of candidate reactions alongside the role new technology is starting to play in this field. This review has focused on the role of individual differences, however, by including more recent studies which have made use of newer selection methods such as asynchronous interviews, (Hiemstra et al., 2019), SJTs (Oostrom et al., 2010; Van Vianen et al., 2004) and video resumes (Hiemstra et al., 2012) it is also starting to expand understanding of the determinants of reactions to these technology-enabled methods. It was clear from previous reviews that the role of individual differences in candidate reactions was not well understood, with preceding research often focusing on specific biographical differences such as race/ethnicity (Ryan & Ployhart, 2000) or focussing on a narrow range of personality traits (e.g., Conscientiousness and Neuroticism in Hausknecht et al., 2004). This review has expanded the knowledge base by looking broadly across the spectrum of individual differences to synthesis findings and identify directions for future research.

As 'individual differences' are a broad area of investigation, it was necessary to split the area to look at more discrete types of individual differences. Therefore, this review assessed the evidence for the impact of biographical differences, attitudinal differences, and personality differences. These categories were aligned with work published by Evertz and Süß (2017), who sought to develop a framework for individual differences research in candidate attraction. While these authors proposed a larger number of categories, some of these would be considered outcomes rather than determinates in the current study (such as 'perceptions'). The three categories in the current study were split relatively evenly across the studies included, with slightly more focus on differences driven by personality factors.

A significant issue with the research in this area is caused by the number of different selection methods available, as was highlighted by Ryan and Ployhart (2000); it may not be possible to generalise findings even between highly similar selection methods. The 18 papers covered a broad range of selection methods from individual, low-touch activities such as cognitive ability tests, to high-touch group activities such as group interviews and assessment centres. This meant that there were no two studies which looked at the same selection procedure and type of individual difference.

Despite this issue, there were some consistent patterns emerging from the data. There were several significant relationships found for attitudinal factors, with more positive attitudes tending to be related to more positive candidate reactions. With regard to personality differences, lower Neuroticism and higher Agreeableness tended to be related to more favourable reactions. The results for biographical factors were more mixed; there was some limited evidence that females may have more negative perceptions of selection processes, particularly if they were not selected. With regard to ethnicity, it appears that race identification plays a role and that there were more favourable reactions particularly when identification was high.

All the papers identified in the study were quantitative, which was to be expected as reactions tended to be measured using survey methods aligned to an existing framework (such as Gilliland, 1993). Nine of the papers were in a lab setting only, which could be an issue for the applicability of findings to a real-world setting; the implications of succeeding or failing a real selection process are likely to be far more significant than in a lab setting, which could result in the selection process being viewed very differently (Hausknecht et al., 2004).

Implications

It was not always clear from the research what contribution the findings had made to practice. Some made suggestions about what to include in the selection process; for example, Bye and Sandal (2016) and Macan et al. (1994) found more favourable justice perceptions towards group interviews and assessment centres. While potentially useful for practitioners choosing between selection methods, these did not add significant new insights, as interviews and assessment centres have been shown to be well received in a number of previous studies (e.g., Anderson et al., 2010). Somewhat more useful was the suggestion from Visser and Schaap (2017) that using different methods in combination may help mitigate differences in reactions, and they noted that it is important to have good processes to attract diverse candidate groups.

Several of the papers focused on raising awareness. For example, Oostrom et al. (2010) noted the need to be aware there may be different reactions to different selection methods and Bernerth et al. (2006) noted that some individuals with different traits might be more likely to perceive processes as unfair, but they did not go on to offer any practical advice regarding what can be done to mitigate this. Truxillo et al. (2006) even suggests that when differences are caused by dispositional traits practitioners could be limited in what can be done, but they also note that these differences should be considered. Merkulova et al. (2014) also highlights the need to consider the nature of applicant pools and provides some suggested actions, such as providing more information about the process if candidates have an external locus of control. However, it is unlikely that practitioners would be able to predict or measure such traits as 'locus of control' in their applicant pools.

Some papers did address the more practical implications of the findings. For example, Ployhart et al. (2003) discussed the need to ensure tests are face valid to help with ethnicity differences in perceptions, whilst Hiemstra et al. (2012) suggested that video resumes might be beneficial for certain groups, which could be advantageous if organisations are looking to attract more diverse application pools. Hiemstra et al. (2019) suggests thinking about the job requirements, noting that if Extraversion is not a key trait for the role, practitioners may wish to avoid video applications, which were off putting to more introverted candidates. Van Vianen et al. (2004) was one of the first papers to look at how perceptions change across the selection journey and provides some practical considerations regarding when and how to give feedback messages.

Limitations

There are limitations in both this review and in the studies found. With regard to the review, the results have been synthesised in narrative form rather than carrying out a meta-analysis. While carrying out a meta-analysis may have made the interpretation of the findings clearer, the variation in methodologies would have made a meta-analysis process challenging. While many of the relationships explored were correlations, direct correlational relationships between individual differences and the candidate reactions measured were not always reported clearly. Further, some individual differences were based on categories (such as "male" and "female") meaning that statistics such as effect size are considered instead. Many of the papers presented reasonably complex statistical models, meaning that model fit was the primary outcome of the analysis.

This review focused on papers that had been published following Gilliland's (1993) work on procedural and distributive justice as this, along with other work published at the same time (Arvey & Sackett, 1993, as cited in Gilliland, 1993; Schuler, 1993), provided a model against which reactions could be measured. However, there may have been research prior to this which was not identified by the review, and the existence of theoretical models and early development of scales to measure the constructs may be the reason why no qualitative papers were identified.

The papers themselves were limited in the samples they considered, with only 2 papers (Bernerth et al., 2006; Hiemstra et al., 2012) looking at reactions of actual job seekers to a range of different roles. As the other field studies all considered applicants to only one type of role, the samples may have been more homogenous in their expectations and therefore their reactions to the selection process than with more varied samples. While not always clearly specified, it can also be reasonably assumed that the majority of participants did not have previous experience of the organisations in question (with the exception of Merkulova et al., 2014). This could have restricted the findings, as the reactions of internal candidates can be significantly different to the reactions of external candidates (Giumetti & Sinar, 2012).

Finally, none of the studies looked at selection contexts other than for a job role. In the workplace, there can be selection processes run for other reasons, such as for secondment opportunities, selection onto talent development schemes, to run specific projects, and for promotions.

Conclusion and Recommendations

A number of significant relationships between individual differences and candidate reactions were identified across the papers in this review, however these relationships were often complex. It appears that relationships are often indirect and can be mediated or moderated by a number of other factors, particularly in the case of personality, where it appears that interactions of different factors of personality can have a greater impact on candidate reactions than direct relationships with single factors.

Despite the focus that has been given to selection processes and candidate reactions in the literature, there is a clear need for more research to assess whether the findings of the studies identified in this review can be replicated with different candidate populations. In particular, internal candidate populations have been given very little attention in the research, therefore research with these populations should be a focus of future studies.

Chapter 4 – Empirical Study: The impact of personality on candidate reactions to a selection process for talent development schemes

Abstract

Competition for talented employees is increasing and the selection process is often an individual's first point of contact with the organisation. Organisations therefore need to ensure that their selection processes are viewed positively by potential candidates or risk losing them to competitors. This study investigates the effects of personality on candidate reactions to a technology-enabled selection process for talent development schemes. A crosssectional survey design was adopted with 625 internal candidates from a public sector organisation. Personality effects were investigated using a Big Five measure of personality, and reactions were measured using an adaption of the Selection Procedural Justice Scale (Bauer et al., 2001), based on Gilliland's (1993) model of Procedural Justice. Reactions to both the selection process as a whole and the four individual exercises which made up the selection process were investigated. The results showed that Conscientiousness, Agreeableness and Neuroticism had small but significant effects on candidate reactions, with Conscientiousness explaining the largest proportion of variance in reactions. The impact of personality varied depending on the exercises, which suggests that multi-method approaches could be highly beneficial in developing a selection process which will attract a broad talent pool. Further implications for theory and practice are discussed and a number of avenues for further research are suggested.

Introduction

Candidate reactions

With the continuing 'War for Talent' (Chambers et al., 1998) and the 'Great Resignation' (Cohen, 2021) following the Covid-19 pandemic it is becoming increasingly critical that

organisations can attract and retain talented employees. Previous research has shown that selection processes can have a significant impact on both potential and current employees' perception of the organisation (Hausknecht et al., 2004). Therefore, it is important that organisations consider how their candidates may react to any selection process they are putting in place in order to mitigate the potential negative outcomes of a poorly received process.

Interest in candidate reactions as a research area within the field of selection grew following Gilliland's (1993) work to develop a model to understand the factors affecting perceived justice or fairness of a selection process. Gilliland's framework (described in full detail in Chapter 1) proposed ten rules which a selection process must follow in order to be perceived as fair, clustered around three factors regarding the formal characteristics or structure of the process, explanations about the process and the interpersonal effectiveness of process administrators. These 'procedural justice' rules provided a strong theoretical foundation on which later researchers (e.g., Chan & Schmitt, 2004; Hausknecht et al., 2004;) have developed to provide insight into the factors affecting candidate reactions to selection processes. However, despite candidate reactions being a focus of the selection literature for over 30 years, a number of gaps in the evidence base remain. For example, Potočnik et al.'s (2021) recent meta-analysis of recruitment and selection research highlighted a number of findings suggesting that there is increasing interest in the role of technology in selection research, and that further research needs to provide conceptual insights to better understand the underlying mechanisms of applicant reactions. Potočnik et al. (2021) also identified the need for practitioners to understand how to design selection practices that are suitable for different demographic groups (e.g., different age groups, ethnicities, genders, and parental status).

Several researchers have noted that while technology has had a significant impact on selection and recruitment, research has not kept pace. In their recent review of the role of technology in recruitment and selection, Nikolaou (2021) noted the importance of applicant reactions as an area that has been significantly impacted by technology. Furthermore, Woods et al. (2020) reviewed selection in the 'digital age', coining the term 'Digital Selection Procedures' (DSP) to describe selection processes supported by emerging digital technologies. Woods et al. concluded that practice has moved faster than research and that a more comprehensive evidence base is needed to support the application of DSPs, and

78

specifically note the changing nature of procedural justice and call to re-define this in the digital age. These technological developments have also made it easy for candidates to share their experiences with others, leading to increased risk to organisations if selection processes are poorly perceived. For example, website Glassdoor, a platform that attracts approximately 41 million users per month, allows employees to post reviews of companies' selection processes. Research has shown that negative reviews on platforms such as this can significantly impact on the organisation's brand and even its financial valuation (Woods et al., 2020; Nikolaou, 2021).

In this current landscape of candidate scarcity is it clear that further investigation of the use of technology-enabled selection processes is needed. This move away from more traditional selection methods means that further interrogation of the determinants and antecedents of candidate reactions, such as individual differences, is needed.

Selection instruments

Several studies have investigated reactions to different types of selection instruments with cross-cultural studies typically reporting consistent results. For example, work samples and interviews are often rated favourably by candidates (Anderson et al., 2010; Anderson & Witvliet, 2008; Hausknecht et al., 2004; Rynes & Connerley, 1993) and honesty (integrity) tests, personal contacts (i.e., knowing someone in a position of influence in the organisation) and graphology are rated least favourably (Anderson et al., 2010; Anderson & Witvliet, 2008; Hausknecht et al, 2004; Bertolino, & Steiner, 2007). However, these studies have tended to look at more traditional selection methods and tend not to include newer selection methods such as SJTs, or questionnaires looking at more discrete knowledge, skills, abilities and other attributes (KSAOs), such as learning agility. Further, personal statements are not included in any of these studies; while most common in selection to higher education institutions (Traynor, Neill & Roulston, 2022) they are also used by a number of employers for selection and are particularly common in the non-profit sector (Blunt, 2022). This suggests that earlier research is becoming less relevant as organisations move away from traditional selection methods.

Applicant groups

While applicant reactions have been studied in a range of contexts, very few researchers have examined the reactions of internal candidates (such as those going through selection for promotions or to access development schemes) or candidates in more senior management roles. This is despite the stakes potentially being much higher for internal candidates than those external to the organisation (Hausknecht et al., 2004) and the concerning finding by Sharf and Jones (2000) that current or former employees may be up to seven times more likely to take legal action than external applicants. One study by Giumetti and Sinar (2012) found differences between the reactions of internal and external applicants, with internal applicants having less positive reactions regarding the level of information they received about a selection test, but being more positive about the predictive validity, user friendliness, and general satisfaction with the process. This suggests that findings regarding external candidate reactions cannot be generalised to internal candidates. Additionally, Ford et al (2009) reviewed the antecedents, justice theories, and outcomes of applicant reactions in a promotion context. They note that a major barrier to research into promotions is the lack of reasonably sized applicant pools to allow statistical analysis (with many promotions being offered on an individual basis) and note that organisations often do not employ a formal process for promotions. It has been suggested that there may be greater impacts on the organisation from reactions to promotion processes than there are associated with external job applicants (Ford et al., 2009; McCarthy et al., 2009).

Individual differences

One of the areas that has had some attention with regard to determinants of applicant reactions is individual differences. In their meta-analysis, Hausknecht et al. (2004) found that biological characteristics such as age, gender, and race had near zero correlations with applicant perceptions, however they noted that some studies found stronger relationships, therefore consideration of different contexts is needed. It was concluded that existing findings from research on personality differences were promising, as small relationships were found between Conscientiousness and Neuroticism and applicant perceptions. Hausknecht et al. (2004) suggested that Openness to Experience in particular may be related to reactions to newer selection processes. When looking at characteristics that were important to individuals

going through a selection process, Viswesvaran and Ones (2004) found relationships between personality traits and the level of importance put on certain characteristics of the process. For example, a combination of Emotional Stability and Extraversion was associated with importance being placed on the development of the process, and there were negative relationships between Emotional Stability and Conscientiousness with context variables (for example the attractiveness of the job, the selection ratio and company reputation).

With regard to the implications of individual differences in applicant reactions, Oostrom et al. (2010) suggest that an understanding of the role personality plays in applicant reactions could be beneficial from a conceptual point of view in understanding the nature of reactions, and from a practitioner point of view it could help identify sources of dissatisfaction in selection processes. Oostrom et al. (2010) suggest that adjusting the test content or test administration medium would have little effect if reactions were due to individual differences. Other researchers (e.g., Hiemstra et al., 2019; Ryan & Huth, 2008) have taken a more positive view, suggesting that understanding the impact of individual differences on candidate reactions could identify if the selection method is ill suited to individuals with traits desirable for the job role, potentially reducing the number of candidates who self-select out of the selection process

Despite several calls from researchers to investigate individual differences (e.g., see McCarthy et al., 2017), particularly with regard to reactions to newer, technology-enabled processes, the systematic literature review (SLR) described in Chapter 3 identified only 18 journal articles that explicitly sought to investigate the role of individual differences in candidate reactions to selection processes.

Current study

As reported in Chapter 3, there are some clear gaps with regard to the impact of personality differences on candidate reactions, in particular the reactions of internal candidates in more senior roles and reactions to selection processes using newer technologies. This study aims to address these gaps by investigating the impact of personality on candidate reactions to an online selection process employing novel selection methods for talent development schemes for senior roles. Gilliland's (1993) procedural justice rules are used as a model for candidate reactions.

As noted by Ford et al. (2009), it can be challenging to find sufficient sample sizes for research with internal candidates and access to these populations can be an issue for researchers. Further, it is unusual to have access to sufficient numbers of senior candidates, as most high-volume selection processes are run for more junior populations, meaning that senior applicants tend to be under-researched in the field of selection (see Hausknecht et al., 2004). An exception to this is the United Kingdom (UK) Civil Service; as an employer of over 480,000 people (Institute for Government, 2022) they have access to much larger groups of internal and external applicants at all levels, and as a public sector organisation they have a duty to be transparent in their processes.

As described in further detail in the method section, an opportunity sample was available for this study to investigate reactions of internal candidates to a selection process, specifically for selection on to prestigious senior leader talent development schemes within the UK Civil Service.

The selection process for this study comprised four different exercises or elements. There was little consistency in the studies identified by the SLR investigating personality with regard to the type of selection methods used, with them covering a number of different selection methods such as interviews, cognitive ability tests, assessment centres, written tests, personality tests and SJTs. As suggested in Truxillo et al. (2006), certain selection methods may be more attractive to some applicants with higher levels of specific personality traits than others, therefore the effect of personality on different selection methods should be considered. As the focus of this study is an online selection process using low-touch selection methods (i.e., where candidates do not have direct contact with assessors), hypotheses will draw upon previous research findings regarding similarly low-touch methods and other relevant theoretical work. Due to the limited existing data on the types of exercises used in the current selection process, no specific hypotheses are made about reactions to the different exercises beyond the hypothesis relating to reactions to the process as a whole, however they will be explored for each of the four exercises.

In the SLR, all the studies identified used a 'Big Five' (Goldberg, 1992; Costa & McCrae, 1985) or Big Five-related measure to assess personality, which includes five personality factors (Agreeableness, Neuroticism, Openness to Experience, Conscientiousness, and

Extraversion). These study findings and the resulting hypotheses to test in the current study are presented below against each of the Big Five factors.

Agreeableness

The 'Big Five' personality trait of Agreeableness has been associated with being kind and trusting, with low Agreeableness being associated with competitiveness and arrogance (Costa & McCrae, 1992). Other traits associated with higher Agreeableness include being flexible, co-operative, forgiving, and tolerant (Barrick & Mount, 1991). In previous research, Agreeableness has been reported to have some of the most significant relationships with candidates' reactions. Truxillo et al. (2006) investigated reactions to a multiple-choice written test and found positive relationships between Agreeableness and social fairness, and in their study investigating reactions to personality questionnaires Bernerth et al. (2006) found Agreeableness to be positively associated with measures of procedural and distributive justice. Finally, Oostrom et al. (2010) investigated reactions to a cognitive ability test and an SJT and found positive relationships with face validity and perceived predictive validity. Based on these previous findings, is it hypothesised that:

Hypothesis 1a:

Agreeableness will predict overall candidate reactions, such that as scores on the trait of Agreeableness increase, total candidate reactions will be more positive.

Hypothesis 1b:

Agreeableness will predict reactions to the social factor of procedural justice, such that as scores on the trait of Agreeableness increase, candidate reactions to the social factor will be more positive.

Hypothesis 1c:

Agreeableness will predict perceived predictive validity of the selection process, such that as scores on the trait of Agreeableness increase, candidates will rate the perceived predictive validity of the selection process more positively.

Hypothesis 1d:

Agreeableness will predict the face validity of the selection process, such that as scores on the trait of Agreeableness increase, candidates will rate face validity of the selection process more positively.

Neuroticism

Neuroticism has been associated with being predisposed to experiencing emotional distress, compared to being more emotionally stable (Costa & McCrae, 1992). Those who score highly on Neuroticism are more likely to experience feelings of anxiety, anger, embarrassment, and insecurity (Barrick & Mount, 1991). Based on the likelihood of those higher in Neuroticism to experience more negative emotions in general, it is hypothesised that Neuroticism would be negatively related to reactions to the selection process overall. Oostrom et al. (2010) examined Emotional Stability (as the inverse of Neuroticism) and found a positive relationship with face validity and perceived predictive validity. Neuroticism has been reported to be negatively related to social fairness in Truxillo et al. (2006), and a negative relationship with distributive justice (but not procedural justice) was found in Bernerth et al (2006). Thus, the following hypotheses have been devised:

Hypothesis 2a:

Neuroticism will predict overall candidate reactions, such that as scores on the trait of Neuroticism increase, total candidate reactions will be more negative.

Hypothesis 2b:

Neuroticism will predict reactions to the social factor of procedural justice, such that as scores on the trait of Neuroticism increase, candidate reactions to the social factor will be more negative.

Hypothesis 2c:

Neuroticism will predict perceived predictive validity of the selection process, such that as scores on the trait of Neuroticism increase, candidates will rate the perceived predictive validity of the selection process more negatively.

Hypothesis 2d:

Neuroticism will predict face validity of the selection process, such that as scores on the trait of Neuroticism increase, candidates will rate face validity of the selection process more negatively.

Openness to Experience

Costa and McCrae (1992) describe Openness to Experience as being curious and unconventional, with those who score lower on Openness to Experience being more traditional and pragmatic. In Barrick and Mount (1991) additional traits associated with Openness to Experience include being imaginative, broadminded, and having a high intellect. Openness to Experience was found to be positively related to perceptions of social fairness in Truxillo et al. (2006), and to face validity and perceived predictive validity in Oostrom et al. (2010). Further, Hausknecht et al. (2004) suggested that Openness to Experience was likely to be associated with reactions to newer selection processes in general. The following hypotheses have been developed:

Hypothesis 3a:

Openness to Experience will predict overall candidate reactions, such that as scores on the trait of Openness to Experience increase, total candidate reactions will be more positive.

Hypothesis 3b:

Openness to Experience will predict reactions to the social factor of procedural justice, such that as scores on the trait of Openness to Experience increase, candidate reactions to the social factor will be more positive.

Hypothesis 3c:

Openness to Experience will predict perceived predictive validity of the selection process, such that as scores on the trait of Openness to Experience increase, candidates will rate the perceived predictive validity of the selection process more positively.

Hypothesis 3d:

Openness to Experience will predict face validity of the selection process, such that as scores on the trait of Openness to Experience increase, candidates will rate face validity of the selection process more positively.

Conscientiousness

Conscientiousness is related to being disciplined and fastidious, with lower Conscientiousness associated with being more laid back and careless (Costa & McCrae, 1992). Other traits associated with Conscientiousness include being dependable, responsible, organised, and planful (Barrick & Mount, 1991). Only one study in the SLR found a significant relationship between Conscientiousness and candidate reactions: Truxillo et al. (2006) found a positive relationship with perceptions of social fairness. This relationship was not hypothesised by Truxillo et al. (2006), however theoretically this follows as those who are more disciplined and organised are more likely to engage with explanations around the process, taking time to process all the information and may therefore feel more positively about the social elements. Further, Hausknecht et al. (2004) did find some evidence for a positive relationship between Conscientiousness and candidate reactions in their metaanalysis. Drawing from these findings the following hypotheses have been developed:

Hypothesis 4a

Conscientiousness will predict overall candidate reactions, such that as scores on the trait of Conscientiousness increase, total candidate reactions will be more positive.

Hypothesis 4b

Conscientiousness will predict reactions to the social factor of procedural justice, such that as scores on the trait of Conscientiousness increase, candidate reactions to the social factor will be more positive.

Extraversion

Extraversion is defined by Costa and McCrae (1992) as being energetic and thrill seeking, compared to introversion which is associated with being more sober and solitary. Other traits typically associated with Extraversion are ambition, sociability, and gregariousness (Barrick & Mount, 1991). None of the relevant SLR studies found direct relationships between Extraversion and candidate reactions, however it was found to interact with other personality traits in Honkaniemi et al. (2013).

Interaction effects

Honkaniemi et al. (2013) investigated a number of personality types (represented by interactions between the different traits in the Big Five model) to investigate reactions to

cognitive ability tests. Honkaniemi et al. used a different framework to Gilliland's which measured only three dimensions of reactions: perceived predictive validity, face validity, and 'fairness'. Honkaniemi et al. found some differences in reactions to the general fairness of the tests between personality types; their 'Overcontrolled' type (i.e., high Neuroticism, low Extraversion, and low Agreeableness) was found to have the most negative reactions to the tests. The type with the most positive reactions were 'Bohemian' (i.e., low Extraversion and low Conscientiousness). It is therefore hypothesised that:

Hypothesis 5a

There will be an interaction effect of low Extraversion and low Conscientiousness on candidate reactions; total candidate reactions will be more positive for candidates with lower scores on Extraversion and Conscientiousness.

Hypothesis 5b

There will be an interaction effect of high Neuroticism, low Extraversion, and low Agreeableness on candidate reactions; total candidate reactions will be more negative for candidates with lower scores on Extraversion and Agreeableness and higher scores on Neuroticism.

These hypotheses are depicted in Figure 1 (see below).

Figure 1 Hypothesised relationship between personality factors and aspects of procedural justice



Note, GIO = Guaranteed interview offer, SEB = Socio-economic background. The predicted direction of the individual or interaction effects are shown in parentheses.

Method

Design

A cross-sectional field study was carried out to test the hypotheses as described in the introduction. Prior to the main study, a pilot of the survey was sent out to current development scheme participants for comment on the suitability and usability of the personality questionnaire and reaction questionnaire items, and the survey platform. Twenty-two participants completed the pilot, and no issues were raised.

Sample

Participating Organisation

The present study was carried out within the UK Civil Service, a public body made up of a number of Government Departments and Executive Agencies collectively employing approximately 480,000 members of staff (Institute for Government, 2022).

Participants

Study participants were applicants to two talent development schemes in the UK Civil Service: the Future Leaders Scheme (FLS) and the Senior Leaders Scheme (SLS). The FLS is targeted at middle managers looking to develop into more senior manager roles, while the SLS is targeted at senior managers looking to move to organisational leader positions (described as Directors within the UK Civil Service). The sample was an opportunity sample based on applications to the 2022 schemes. A total of 2,884 candidates applied to both schemes, and of these 2,603 (90.3%) were applications to FLS and 281 (9.7%) were applicants to SLS. All participants were employed as Civil Servants and met the inclusion criteria set out by the schemes (this is based on substantive job grade of participant at the time of application and sponsorship of the application from a line manager or other appropriate sponsor senior to the applicant).

A total of 681 participants (23.6% of the available population) took part in the study. Following the data clearing approach (detailed below), 625 participants we retained for analysis. 57.3% (n = 357) were female (41.6% male, n = 260; 1.3% other/missing, n = 8); 83.5% (n= 522) were white (13.8% ethnic minority, n= 86; 2.7% other/missing, n= 14), 84.6% (529) had no disability (11% disability, n= 69; 4.3% other/missing, n= 17); 59.4% (n= 371) from a professional socio-economic background (20.6 % working class, n= 129; 15.8% intermediate, n= 99; 4.2% other/missing, n= 29), and 71.7% (n= 448) were in the 30-50 age groups (11.5% under 30, n= 72; 14.6% over 50, n= 91; 2.2% other/missing, n= 14). Comparison of the demographic make-up of the sample with the total applicant population suggested that the sample data is representative of the population of interest.

Procedure

Selection for the schemes happens annually, with applications typically opening between February and April. The selection process comprises two stages, with Stage 1 typically completed by May and Stage 2 completed in the June - July period. The development schemes begin in September and run for approximately 18 months. The selection process is updated each year, however Stage 1 usually consists of online tests and a candidate statement (the full details of Stage 1 of the 2022 selection process are detailed in the Selection Methods section below) and Stage 2 is typically an interview. Following the completion of Stage 1, results are analysed by the selection team and candidates obtaining a pass mark (derived each year based on a minimum standard and the number of places available on the scheme) progress to Stage 2. Candidates do not receive their results immediately on completion of Stage 1, meaning that the reaction questionnaire was completed before candidates knew whether they had been successful or not in progressing to Stage 2. In general, about two thirds of FLS candidates and one third of SLS candidates do not progress to Stage 2. Once Stage 2 is completed, feedback on the online tests and candidate statement is shared with candidates. This includes scores for each exercise (on a 1-10 sten scale for the tests and a 1-5 rating scale for the candidate statements) and autogenerated feedback text for the online tests.

Prior to applying to the schemes, potential candidates are able to attend briefing sessions which provide them with an overview of the selection process and information on the number of places available on the schemes. Pass marks are not shared as these are not fixed each year (but rather developed based on the performance of the current years' applicants), however the expected selection ratios are shared with potential applicants. In addition to the briefings, information about the selection process is available on the Civil Service learning and development online platform. To apply to the schemes, candidates had to self-register with their email address via an online platform which hosts the selection activities. Once Stage 1 of the selection process was completed, the selection team provided the researcher with the candidates' contact emails and a unique candidate identifier. Candidates were contacted to take part in the survey via email which contained a link to an online survey. Stage 1 of the selection process for both FLS and SLS comprises four elements: a text based SJT, a leadership potential questionnaire (LPQ), a learning agility questionnaire (LAQ), and written candidate statement (CS). These exercises are completed in sequence, typically starting with the SJT, followed by the questionnaires, and finishing with written statement. Candidates usually have approximately 4 weeks in which to complete their application.

The SJT was developed specifically for the 2021 schemes and the content was refreshed for the 2022 schemes. The SJT is a fixed format test, meaning all candidates see the same set of questions. It is typically refreshed each year in order to avoid giving an unfair advantage to those who may be reapplying for the schemes. The test contains a number of scenarios and possible actions; candidates are asked to rate the effectiveness of each action. Their responses are compared to the responses of subject matter experts and scoring is based on the level of agreement with the subject matter experts. The test is set in a fictional government department and the scenarios were developed based on critical incident interviews with senior leaders within the UK Civil Service. Most SJTs present candidates with a number of unrelated scenarios against which to rate the actions. However, in the current SJT the scenarios are linked, and develop along a fictional timeline in order to assess candidates' ability to draw on existing information and apply new learning.

The LPQ and LAQ were also developed for the 2021 schemes and were used in the same format for 2022. The LAQ was based on a questionnaire developed for selection to the UK Civil Service graduate schemes ('Fast Stream') and was revised to be more suited to senior audiences. It measures the individuals' learning preferences and styles. The LPQ was developed to measure elements of the Model of Potential, a talent model used by the UK Civil Service to support various talent development activities. The questionnaire measures six scales important to successful leadership in the UK Civil Service: Drive; Achievement Striving; Commitment; Expression; Self-Management and Analytical Approach. Both questionnaires used a mixed rating and ranking response format. In the LPQ, candidates are presented with three statements at a time and asked to rate their level of agreement with each one on a 7-point rating scale. They are then asked to rank the statements by selecting which are most and least like them. The LAQ follows the same format, but candidates are presented with four statements at a time. In the candidate statement, candidates are asked to provide a 500-word response to a question designed to measure the aforementioned Model of Potential. The question is refreshed each year and asks candidates to provide evidence against the three elements of the Model of Potential; Performance, Values and Purpose. Precise wording of the current question is available on request.

At time of undertaking the study, predictive validity statistics for the SJT, LPQ and LAQ were not available. However, the content validity of the SJT was assessed during the development stage through a series of focus groups with subject matter experts and piloting with previous scheme participants. Divergent validity was demonstrated through intercorrelations of the test scores. The internal consistency reliability of the LPQ and LAQ was assessed during the development of questionnaires and was in line with accepted test development standards (EFPA, 2008). A key focus of the development of the current selection process was to ensure that no demographic groups were adversely impacted by the selection process. Data available for the 2021 schemes showed no adverse impact (using the

4/5th rule, e.g., Hauenstein et al., 2013), against underrepresented and protected groups and a high level of parity in pass rates across all groups was found. Candidate feedback on the schemes is also collected each year which has demonstrated a largely positive reception to the selection process. Further information on the reliability, validity and fairness of the selection tests is available on request.

Ethical approval

Ethical approval was granted for this study by Birkbeck, University of London in November 2021. The participating organisation was provided with an invitation letter and a meeting was held to answer any further questions about the process. The participant information sheet was sent to participants with the survey information and the online survey included a mandatory question regarding informed consent before participants could complete the study.

Measures

Two measures were examined in the survey, these were the candidate Reactions Questionnaire (RQ) and the Personality Questionnaire (PQ). A summary of the measures is provided below, the full list of items can be found in Appendix 1 and Appendix 2 respectively.

Candidate reactions questionnaire (RQ)

The RQ was developed based on Bauer et al.'s (2001) Selection Procedural Justice Scale (SPJS). This scale was developed to measure Gilliland's (1993) procedural justice rules in a selection context and measures a number of subscales: job-relatedness predictive (2 items), information known (3 items), chance to perform (4 items), reconsideration opportunity (5 items), feedback (3 items), consistency (3 items), openness (4 items), treatment (5 items), two-way communication (5 items), propriety of questions (3 items), and job-relatedness content (2 items). The SPJS has three higher order factors: the 'Social' factor, the 'Structure' factor, and a third factor containing the job-relatedness content subscale. The subscales of consistency, openness, treatment, two-way communication, and propriety of questions load on to the 'Social' factor, whilst the subscales of job-relatedness predictive, information known, chance to perform, reconsideration opportunity, and feedback load on to the 'Structure' factor.

As discussed in Chapter 2, the SPJS was adapted into the RQ for the current study. The RQ was administered in two sections: the first section investigates reactions to the process as whole and the second section investigates reactions to the individual elements of the process. The scales covered in each section are as follows, with an example question from each scale shown in brackets: section 1 of the RQ covered the subscales of job-relatedness predictive ('Doing well on this selection process means a person can do well on the schemes'), information known ('I understood in advance what the selection process would be like'), chance to perform ('I could really show my skills and abilities through this process'), feedback ('I had a clear understanding of when I would get my results for Stage 1 of the selection process'), two-way communication ('There was enough communication during the selection process'), job-relatedness content ('It would be clear to anyone that this process is related to leadership'), and a new single item scale looking at the contribution of each element to the overall outcomes ('contribution to process', 'I understood what contribution each of the activities would make to the selection process outcome'). Section two covered the scales of job-relatedness predictive ('Doing well on this activity means a person can do well on the schemes'), information known ('I understood in advance what the activity would be like'), chance to perform ('I could really show my skills and abilities through this activity'), feedback ('I had a clear understanding of when I would get my results for this activity'), proprietary of questions ('The content of the activities did not appear to be prejudiced'), jobrelatedness content ('It would be clear to anyone that this activity is related to leadership'), and contribution to process ('I understood what contribution this activity would make to the selection process outcome'). Participants were asked to give their level of agreement with each questionnaire item on a 1-5 rating scale ('Strongly Disagree' to 'Strongly Agree'). Responses to each subscale and factor were totalled such that higher scores indicated more positive reactions.

The convergent validity of the SPJS was demonstrated through its relationship with a global measure of procedural justice (Bauer et al, 2001). The divergent validity was also demonstrated through the weak relationships found between the SPJS and other unrelated measures (Bauer et al, 2001). With regard to reliability, Bauer et al. (2001) reports Cronbach's alphas ranging from .73 (Job-relatedness Content) to .92 (Treatment) during the development of the SPJS, this indicated that the questionnaire has good internal consistency reliability.

As the questionnaire was adapted for use in the current study, confirmatory factor analysis (CFA) was carried out to check whether the three-factor structure for the SPJS held with the reduced number of subscales used in the RQ. CFA was carried out on section 1 of the RQ, with the new 'contribution to process' scale expected to load on to the 'Structure' factor. In line with Pontes et al. (2021), the following fit indices were adopted: χ^2/df [1;4]; Comparative Fit Index (CFI); and Tucker-Lewis Fit Index (TLI) [0.90;0.95]; Root Mean Square Error of Approximation (RMSEA) [0.05;0.08]; and the Standardized Root Mean Square Residual (SRMR) [0.05;0.08]. The CFA was carried out using the lavaan package in R version 4.2.0 (Rosseel, 2012).

The results for the 3-factor model showed only moderate fit using Bauer et al.'s (2001) factor structure. Due to the exclusion of subscales, only one subscale remained relating to the Social factor ('two-way communication') and there were high correlations between the Social and Structure factors, and between the Structure factor and third factor. A two-factor model was then tested, mapping the job-relatedness content subscale to the Structure factor and the feedback and information known scales to the Social factor. This aligns with Gilliland's original model which proposed job-relatedness and opportunity to perform were part of the formal characteristics of the process, whereas feedback and selection information were rules to do with explanations of the process. The results for the two-factor model showed the following goodness of fit: χ2 < .001, df=145, CFI=0.961, TLI=0.954, RMSEA=0.054. SRMR=0.054. With the exception of the chi square statistic, the indices suggested that the current RQ fit a two-factor model. As noted in Cheung & Rensvold (2002), the χ^2 is less useful when sample sizes are large, therefore the other indicators were considered. Examination of the AIC and BIC statistics also showed a reduction for the two-factor model, therefore this was retained. Further details of the three-factor and two-factor models can be found in the Supplementary Materials.

The two-factor model was then tested for section 2 of the questionnaire, with CFA run for reactions to each exercise of the process; the 'propriety of questions' subscale replaces the 'two-way communication' subscale in this model. The two-factor model showed good fit for reactions to all four exercises. The internal consistency reliability of the RQ scales for the current sample was assessed using Cronbach's alpha. These ranged from .68 to .91, showing good reliability. Full information and results for the CFA and information on the means,

standard deviations, number of items and reliabilities of the RQ scales and subscales can be found in the Supplementary Materials.

Personality Questionnaire (PQ)

Personality was measured using scales from the International Personality Item Pool to measure the Big Five traits (Goldberg et al., 2006). The IPIP is an open-source resource of personality questionnaire items and scales which have been validated against other established questionnaires. This study used the 50-item questionnaire with measures of the Big Five domains and is validated against Costa and McCrae's (1992) NEO-PI-R (Goldberg, 1999, as cited in Goldberg et al., 2006). Ten items are used to measure each of the Big Five personality factors, five of which of positively worded and five of which are negatively worded. An example item from each factor is as follows: Neuroticism, '*Am not easily bothered by things*'; Extraversion, '*Feel comfortable around people*'; Openness to Experience, '*Enjoy hearing new ideas*'; Agreeableness, '*Have a good word for everyone;* Conscientiousness, '*Am always prepared*'. Participants were asked to give their level of agreement with how accurately each statement described them on a 1-5 rating scale ('Very Inaccurate' to 'Very Accurate'). Responses to the negatively worded items were reverse scored, then items responses were totalled such that higher scores related to higher level of the trait in question.

To assesses construct validity and internal consistency reliability, Goldberg et al. (2006) provide the following correlations between the IPIP factor scores and NEO-PI-R factor scores (Cronbach's alpha reliabilities shown in brackets): Neuroticism = .82 (r = .86); Extraversion = .77 (r = .86); Openness = .79 (r = .82); Agreeableness = .70 (r = .77); Conscientiousness = .79 (r = .81). These indicate that the IPIP are a valid and reliable measure of the Big Five factors of personality as defined by Costa and McCrae (1992).

The internal consistency reliability of the PQ scales for the current sample was assessed using Cronbach's alpha. These ranged from .65 to .81 showing good internal consistency reliability. The means, SDs, number of items and reliabilities for the PQ factor scores can be found in the Supplementary Materials.

Control variables

The findings from the SLR (Chapter 3) suggested that there are other variables that may impact on candidate reactions. Some studies have controlled for the relationship of biographical factors (typically collected as demographic data such as age, gender and ethnicity) to candidate reactions, other studies have investigated factors such as outcomes of the process and test performance as control variables. Specifically, four studies in the SLR controlled for age (Brenner et al., 2016; Bye & Sandal, 2016; Honkaniemi et al., 2013; Oostrom et al., 2010;) and four controlled for gender (Brenner et al., 2016; Honkaniemi et al., 2013; Oostrom et al., 2010; Truxillo et al., 2006). Actual test performance was controlled for in Merkulova et al. (2014) and Truxillo et al. (2006) and four studies controlled for test taking experience (Brenner et al., 2016; Oostrom et al., 2010; Van Vianen et al., 2004; Truxillo et al., 2006). To understand the unique contribution personality makes to the variance in candidate reaction, the factors of age, gender, ethnicity, and test score will be considered as control variables. While test taking experience is not measured, data is available on whether candidates had applied to the FLS scheme previously, which can be used as a proxy for test taking experience as a further control. While not investigated in any of the previous studies, socio-economic status will be considered as a control as previous research has found this to be related to other perceptions, such as organisational attractiveness (Nadler et al., 2017). Finally, some candidates were eligible for a guaranteed interview and could therefore chose to bypass elements of the selection process if they wished. Guaranteed interview offer (GIO) status was considered as the final control variable as this choice in whether to complete the exercises or not may have caused participants to react differently than those who were obligated to complete them.

Data cleaning

681 participants started the online survey. However, 12 (1.8%) did not accept the informed consent declaration, therefore were removed from the study. One participant (.15%) entered text to state they had not been able to access the participant information sheet but were happy to take part; they were removed from the study as they had not been able to give fully informed consent. Two participants (.3%) entered ID numbers that did not match the biodata file and so were removed from further analysis. Duplicate ID numbers were then checked and none were found. To check for level of engagement with the surveys, the survey completion times and response patterns were reviewed. Five participants (.7%) were missing one or more survey response; as all questions required a response this may have suggested errors in the

completion, therefore they were removed from further analysis. Candidates who complete the selection process under the 'Guaranteed Interview Offer' (GIO) can choose to bypass any part, or all parts, of the selection process. Candidates who were GIO and were missing scores for one or more of the elements of the process were removed from the analysis as they would not have been able to provide legitimate feedback on all elements of the selection process. Twenty-one participants (3.1%) were missing the SJT scores and were GIO, so were removed from the analysis due to missing data. Similarly, 9 GIO candidates (1.3%) were missing scores for the candidate statements and were removed. This data cleaning process resulted in 625 participants being retained for the analysis.

Analysis of control variables

To identify the impact of the socio-demographic data and test scores on the RQ results, tests of difference were carried out. Independent samples t tests were carried out where there were two main categories of interest and Analysis of Variance (ANOVA) was carried out for multiple categories.

The independent t tests for Gender showed significant differences at the p < .05 level for seven of the subscale scores. Investigation of effect sizes using Cohen's *d* (Gignac and Szodorai, 2016) showed that these were all small (ranging from .18 to .31). The independent t tests for Ethnicity showed significant differences at the p < .05 level for eight of the subscales. Investigation of effect sizes showed that these were also all small (ranging from .23 to .33). There were no significant differences in RQ scores between the different age groups or on the basis of socio-economic background. For those who had applied for the schemes before, there were significant differences (at the p < .05 level) for four of the five Information Known subscales, which could be expected. There was no significant difference in reactions between those who applied for FLS and those who applied to SLS. There were some significant differences RQ scores for the GIO candidates compared to non-GIO, with the GIO candidates scoring significantly lower (at the p < .05 level) on 21 of the 40 subscales. Effect sizes for the difference ranged from .33 to .63.

The impact of socio-demographic data and test scores on the PQ results were also investigated. In terms of personality differences, there were significant gender differences (at the p < .05 level) for three of the five factors, with females scoring significantly higher on Agreeableness, Conscientiousness and Neuroticism than males. The effect sizes were small, ranging from .25 to .31. For ethnicity, Ethnic Minority participants scored significantly higher on Extraversion and lower on Neuroticism than White participants, effect sizes again were small (.26 and .37 respectively). In terms of SEB and age, there were significant differences for Extraversion (SEB) and Agreeableness and Neuroticism (age, all at the p < .05level).

As noted, GIO candidates can choose to undertake the selection process, or to bypass some or all elements. No significant differences in personality scores were found for the GIO participants. Significantly more were from an Ethnic Minority background and had a disability, this is to be expected as some of the criteria for GIO is whether an individual has a disability and whether they have completed another development scheme which is aimed at ethnic minority employees.

The final control to explore was the impact of actual test performance. Small but significant correlations were found between a number of the RQ scales and scores on the different elements of the selection process. Based on these findings, the following variables were controlled for in the regression: Gender, Age, Ethnicity, SEB, GIO status and exercise scores. Full details of the control variable analysis are available on request.

Statistical analysis

Prior to testing the hypotheses, preliminary analysis was carried out as detailed above to confirm the structure of the RQ, check the internal consistency reliability of the RQ and the PQ, identify control variables for use in the regression analysis, and to clean the data to remove any invalid cases. These steps were necessary in order to identify any issues with the data and ensure that the assumptions for the specified analysis had been met.

The assumption of normality of measures (PQ and RQ) was also examined. The Kolmogorov-Smirnov test showed significant deviation from normality for the majority of variables, however the large sample size means it is likely that the test is over-powered (Field, 2009). The skewness and kurtosis were then evaluated. The normal distribution assumption was fulfilled as there were no absolute values of skewness >3 and kurtosis >8 (Kline, 2011, as cited in Pontes et al., 2022). The assumption of linearity was tested through

examination of the Q-Q plots, which indicated no clear violations in the assumption. The full analysis table testing for normality can be found in the Supplementary Materials.

The assumptions for multiple linear regression were checked prior to running the regression analysis to determine the suitability of the data. There are various 'rules of thumb' for sample sizes, for example Field (2009) suggests 50 + 8k (with *k* being the number of predictors) for testing model fit and 104 + k for testing the contribution of individual variables. Where previous data exists, power analysis can be carried out to compute minimum sample sizes. While previous research in this area is limited, previous studies have found effect sizes in the smalll to medium range (e.g., Oostrom et al., 2010, Truxillo et al., 2006). Power analysis using G*Power (version 3.1.9., Faul et al., 2007) with a small to medium effect size ($f^2 = .85$), alpha = .05, six control variables (i.e., Gender, Age, Ethnicity, SEB, GIO status and exercise score) and one predictor (the personality factor) gives a required sample size of 265 with a power of .95. The current sample exceeds these values. Analysis of the correlations between predictor variables revealed that none of the predictor variables were highly correlated, and examination of the variation inflation factors (VIF) and tolerance were within acceptable limits (Field, 2009), meaning that the assumptions for linear regression were met.

One of the control variables was exercise performance. In order to use this as a control in reactions to the overall process, a composite score was created based on the four exercise scores. This was achieved by converting all scores to z scores to provide equivalent scaling, and then computing a mean exercise score per participant. The alternative would have been to include each exercise score in the regression as controls, however this would have increased the number of predictors in the model, making it potentially less robust for the sample size.

Descriptive analysis was carried out to describe the demographic characteristics of the sample. To test the main hypotheses 1-4 in the study, correlational analysis was carried out first to calculate Pearson's correlation coefficient. Where significant correlations were found (at the p < .05 level), regression analysis was conducted to identify whether the personality factors were significant predictors of reactions once other factors had been controlled for. As hypothesis 5 was concerned with interaction effects, only regression analysis was conducted. A hierarchical regression was used, with the control variables entered in the first step followed by the personality variable(s) of interest. This type of regression analysis allows for the unique contribution of the step 2 variable(s) to be calculated so is the most suitable for the

current study, where a goal of the research is to identify how much variance in reactions is explained by the personality factors once other factors are controlled for. Hypotheses 1-5were first investigated for reactions to the selection process overall (measured in section 1 of the RQ), followed by reaction to the individual exercises (measured in section 2 of the RQ). The analysis was carried out using IBM SPSS Statistics for Windows, Version 26 (IBM Corp, 2019).

Results

Correlations between PQ and RQ variables

Table 8 presents the means, standard deviations and correlations to test the direct relationships between the PQ and the RQ variables. The hypothesised relationships were tested for reactions to the total selection process and for reactions to each of the four exercises.

As shown in the Table 8, the hypothesised relationships for total process were largely supported. Agreeableness was significantly positively related to total reactions, (r(623) = .177, p < .001) the social factor (r(623) = .127, p < .05), and job-relatedness content (r(623) = .103, p = .01), thus hypotheses 1a, 1b and 1d were supported. The relationship with job-related predictive was not significant, meaning hypothesis 1c was not supported. Neuroticism was significantly negatively related to total reactions (r(623) = -.179, p < .001), the social factor (r(623) = -.150, p < .001), and job-relatedness content (r(623) = -.104, p < .05), meaning hypotheses 2a, 2b and 2d were supported. The relationship with job-relatedness predictive was not significant, meaning hypothesis 2c was not supported. No significant correlations were found for the hypothesised relationships were between Openness to Experience and candidate reactions, meaning hypotheses 3a to 3d were not supported. Conscientiousness was significantly positively related to total reactions (r(623) = .212, p < .001), and the social factor (r(623) = .144, p < .001), meaning that hypotheses 4a and 4b were supported

Exercises

Tables 9-12 present the correlations for reactions to the four exercises. The hypotheses regarding the relationship between personality variables and candidate reactions were tested for each of the four exercises. Agreeableness was significantly positively related to total reactions for the CS, LAQ, LPQ and SJT, (r(623) = .156, p < .001; r(623) = .214, p < .001; r(623) = .186, p < .001; r(623) = .162, p < .001 respectively), meaning hypothesis 1a was supported for all exercises. Agreeableness was significantly positively related to the social factor for the CS, LAQ, LPQ and SJT (r(623) = .137, p < .001; r(623) = .176, p < .001; r(623) = .165, p < .001; r(623) = .121, p < .05 respectively), meaning that hypothesis 1b was supported for all exercises. Agreeableness was significantly positively related to job-relatedness predictive for the CS, LAQ, LPQ and SJT (r(623) = .101, p = .01; r(623) = .117, p < .05; r(623) = .094, p = .02; r(623) = .086, p = .03 respectively), meaning hypothesis 1c was supported for all exercises. Agreeableness was significantly positively related to the job-related content for the CS, LAQ, LPQ and SJT (r(625) = .124, p < .05; r(623) = .138, p < .001; r(623) = .115, p < .05; r(623) = .037, r(623) = .138, p < .001; r(623) = .115, p < .05; r(623) = .037, r(623) = .138, p < .001; r(623) = .115, p < .05; r(623) = .127, p < .057 respectively), meaning hypothesis 1c was supported for all exercises. Agreeableness was significantly positively related to the job-related content for the CS, LAQ, LPQ and SJT (r(625) = .124, p < .057; r(623) = .138, p < .0017; r(623) = .115, p < .057; r(623) = .127, p < .057 respectively), meaning hypothesis 1d was supported for all exercises.

Neuroticism was significantly negatively related to total reactions for the CS, LAQ, LPQ and SJT (r(623) = -.120, p < .05; r(623) = -.155, p < .001; r(623) = -.148, p < .001; r(623) = -.166, p < .001 respectively), meaning hypothesis 2a was supported for all exercises. Neuroticism was significantly negatively related to the social factor for the CS, LAQ, LPQ and SJT (r(623) = -.107, p = .02; r(623) = -.134, p < .001; r(623) = -.150, p < .05; r(623) = -.146, p < .001 respectively), meaning hypothesis 2b was supported for all exercises. Neuroticism was significantly negatively related to job-relatedness predictive for the CS, LAQ and LPQ (r(623) = -.100, p = .01; r(623) = -.099, p = .01; r(623) = -.107, p = .01 respectively). The correlation for job-related predictive of the SJT was not significant, meaning hypothesis 2c was partially supported for the exercises. Neuroticism was negatively related to the job-related content for the CS, LAQ, LPQ and SJT (r(623) = -.101, p = .01; r(623) = -.101; p = .01; r(623) = -.101; p = .01; r(623) = -.101, p = .01; r(623) = -.101; p

Openness to Experience was positively related to job-relatedness predictive of the CS (r = .082, p = .04) meaning hypothesis 3d was partially supported for the exercises. No other significant correlations were found, meaning hypotheses 3a - 3c were not supported for the exercises.

Conscientiousness was positively related to total reactions for the CS, LAQ, LPQ and SJT (r(623) = .170, p < .001; r(623) = .219, p < .001; r(623) = .217, p < .001; r(623) = .188, p < .001 respectively), meaning hypothesis 4a was supported for all exercises. Conscientiousness was positively related to the social factor for the CS, LAQ, LPQ and SJT (r(623) = .136, p < .001; r(623) = .168, p < .001; r(623) = .184, p < .001; r(623) = .169, p < .001 respectively), meaning hypothesis 4b was supported for all exercises.

Bonferroni correction

To explore the likelihood of Type 1 errors occurring with multiple comparisons, a Bonferroni correction was applied to the PQ and RQ correlations. This was carried out using the method as described in a number of papers (Armstrong, 2014; Nakagawa, 2004) where the proposed alpha (.05) is divided by the number of comparisons made. This leads to a corrected alpha of <.001 for the current study. Correlations reaching this significance level are denoted in tables 8 - 12 below.

Applying this correction, Agreeableness was significantly positively related to total reactions, (r(623) = .177, p < .001) but not the social factor, job-relatedness predictive or job-relatedness content, thus hypotheses 1a was supported but 1b, 1c and 1d were unsupported. Neuroticism was significantly negatively related to total reactions (r(623) = -.179, p < .001), the social factor (r(623) = -.150, p < .001), but not job-relatedness predictive or job-relatedness content meaning hypotheses 2a and 2b were support but 2c and 2d were not supported. No significant correlations were found for the hypothesised relationships were between Openness to Experience and candidate reactions, meaning hypotheses 3a to 3d were not supported. Conscientiousness was significantly positively related to total reactions (r(623)=.212, p < .001), and the social factor (r(623)=.144, p < .001), meaning that hypotheses 4a and 4b were supported

Exercises

The Bonferroni correction was also applied to the relationships between personality variables and candidate reactions for each of the four exercises. Applying this correction, Agreeableness was significantly positively related to total reactions for the CS, LAQ, LPQ and SJT, (r(623) = .156, p < .001; r(623) = .214, p < .001; r(623) = .186, p < .001; r(623) = .162, p < .001 respectively), meaning hypothesis 1a was supported for all exercises. Agreeableness was significantly positively related to the social factor for the CS, LAQ, and LPQ (r(623) = .137, p < .001; r(623) = .176, p < .001; r(623) = .165, p < .001 respectively) but not the SJT, meaning that hypothesis 1b was supported for all exercises except the SJT.. Agreeableness was not significantly positively related to job-relatedness predictive for the exercises, meaning hypothesis 1c was not supported for any exercise. Agreeableness was significantly positively related content of the LAQ (r(623) = .138, p < .001) meaning hypothesis 1d was partially supported for the exercises.

Neuroticism was significantly negatively related to total reactions for the LAQ, LPQ and SJT; r(623) = -.155, p < .001; r(623) = -.148, p < .001; r(623) = -.166, p < .001 respectively), but not the CS meaning hypothesis 2a was supported for all exercises except the CS. Neuroticism was significantly negatively related to the social factor for the LAQ and the SJT (r(623) = -.134, p < .001; r(623) = -.146, p < .001 respectively), but not the CS or the LPQ respectively meaning hypothesis 2b was supported for some exercises. Neuroticism was not significantly negatively related to job-relatedness predictive for the exercises, meaning hypothesis 2c was not supported for the exercises. Neuroticism was not significantly negatively related content for the exercises meaning hypothesis 2d was not supported for all exercises.

Openness to Experience was not significantly related any of the candidate reactions, meaning hypothesis 3a - 3d were not supported for the exercises.

Conscientiousness was positively related to total reactions for the CS, LAQ, LPQ and SJT (r(623) = .170, p < .001; r(623) = .219, p < .001; r(623) = .217, p < .001; r(623) = .188, p < .001 respectively), meaning hypothesis 4a was supported for all exercises. Conscientiousness was positively related to the social factor for the CS, LAQ, LPQ and SJT (r(623) = .136, p < .001; r(623) = .186, p < .001; r(623) = .184, p < .001; r(623) = .169, p < .001 respectively), meaning hypothesis 4b was supported for all exercises.

While the Bonferroni correction aims to reduce the risk of a Type 1 error, in doing so the risk of Type 2 errors can be increased (e.g., Armstrong, 2014, Rothman, 1990). Further, the Bonferroni correction has been criticised by some as unnecessary, for example by Perneger (1998), who suggest "The

integration of prior beliefs with evidence is best achieved by Bayesian methods, not be Bonferroni adjustment" (pp1237-1238) and by Nakagawa, who concludes in their review that effect sizes ought to be reported rather than applying Bonferroni procedures. Finally, applying such corrections means that the power to detect small effect will be compromised (Glickman et al., 2014). In response to these criticisms, the subsequent regression analysis and discussion of findings are carried out on the basis of uncorrected values.

Descriptive statistics and correlations between PQ and RQ variables for Total Process

Variable	Mean	SD	Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Total Reactions	65.56	9.85	625	1													
2 Structure Factor	27.32	5.89	625	$.807^{\dagger}$	1												
3 Social Factor	38.24	6.17	625	$.826^{\dagger}$.334†	1											
4 Job-relatedness Predictive	6.25	1.44	625	$.526^{\dagger}$	$.692^{\dagger}$	$.179^{\dagger}$	1										
5 Information Known	11.59	2.34	625	.632†	.269†	.752 [†]	.133†	1									
6 Chance to Perform	10.56	3.33	625	$.688^{\dagger}$.897†	.243†	.490†	.201†	1								
7 Feedback	11.63	2.60	625	$.601^{\dagger}$.191†	$.776^{\dagger}$.112**	.331†	.134†	1							
8 Two-way Communication	15.02	2.78	625	.741†	.336†	$.862^{\dagger}$	$.182^{\dagger}$	$.518^{\dagger}$	$.245^{\dagger}$.511†	1						
9 Job-relatedness Content	7.40	1.73	625	.644†	.749†	.313†	.419†	.230†	.495†	.198†	.316†	1					
10 Contribution to Process	3.09	1.07	625	$.548^{\dagger}$.566†	.334†	.254†	.306†	.362†	.165†	.331†	.399†	1				
11 Agreeableness	43.49	3.54	625	.177 [†]	.163†	.127**	.070	.099*	.164†	.059	.143†	.103**	.124**	1			
12 Conscientiousness	42.21	4.49	625	.212†	.203†	.144†	.124**	.113**	.169†	.108**	.124**	.169†	.150†	.358†	1		
13 Extraversion	38.22	5.51	625	.130†	.169†	.046	.173†	.083*	.130†	061	.090*	.125**	.091*	.214†	.339†	1	
14 Neuroticism	21.65	5.57	625	179†	142†	150†	048	118**	139†	061	175†	104**	115**	317†	319†	325†	1
15 Openness	39.77	4.39	625	013	015	007	.033	.005	001	050	.027	034	066	.152†	.091*	.282†	027

Note: PQ = Personality Questionnaire, RQ = Reaction Questionnaire. [†]Correlation significant at p < .001 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed). * Correlation is significant at the 0.05 level (2-tailed).

Descriptive statistics and	correlations betwe	en PQ and RQ	variables for	Candidate Statement

Variable	Mean	SD	Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Total Reactions	67.11	9.19	625	1													
2 Structure Factor	40.15	6.34	625	.899†	1												
3 Social Factor	26.96	4.25	625	$.840^{\dagger}$.518†	1											
4 Job-relatedness Predictive	6.65	1.55	625	.635†	$.769^{\dagger}$	$.288^{\dagger}$	1										
5 Information Known	12.00	2.20	625	.595†	.340†	.741†	.199†	1									
6 Chance to Perform	13.69	3.47	625	.799†	.923†	$.419^{\dagger}$	$.622^{\dagger}$	$.288^{\dagger}$	1								
7 Feedback	11.58	2.55	625	$.616^{\dagger}$.342†	$.780^{\dagger}$.172†	$.262^{\dagger}$.275†	1							
8 Two-way Communication	12.03	1.54	625	$.697^{\dagger}$	$.530^{\dagger}$	$.706^{\dagger}$.313†	$.400^{\dagger}$.414†	.353†	1						
9 Job-relatedness Content	7.78	1.38	625	.698†	.699†	$.498^{\dagger}$.398†	.292†	$.506^{\dagger}$.344†	.533†	1					
10 Contribution to Process	3.38	1.05	625	$.490^{\dagger}$.444†	$.408^{\dagger}$.210 [†]	.257†	.230†	$.284^{\dagger}$.411†	.301†	1				
11 Agreeableness	43.49	3.54	625	.156†	.135†	.137†	.101*	.081*	.100*	.082*	$.170^{\dagger}$.124**	.106**	1			
12 Conscientiousness	42.21	4.49	625	$.170^{\dagger}$.159†	.136†	$.160^{\dagger}$.051	.094*	.137†	.116**	.157†	.127**	.358†	1		
13 Extraversion	38.22	5.51	625	.129**	$.147^{\dagger}$.069	$.205^{\dagger}$.054	.089*	001	.136†	.117**	.068	.214†	.339†	1	
14 Neuroticism	21.65	5.57	625	120**	118**	090*	100*	036	062	079*	092*	101*	166 [†]	317†	319†	325†	1
15 Openness	39.77	4.39	625	.025	.035	.006	.082*	.036	.021	045	.039	.036	042	.152†	.091*	.282†	027

Note: PQ = Personality Questionnaire, RQ = Reaction Questionnaire, SD = standard deviation. [†]Correlation significant at*p*<.001 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed). *Correlation is significant at the 0.05 level (2-tailed).

Descriptive statistics and correlations between PQ and RQ variables for LAQ

Variable	Mean	SD	Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Total Reactions	63.26	9.57	625	1													
2 Structure Factor	37.11	6.69	625	.895†	1												
3 Social Factor	26.16	4.59	625	$.828^{\dagger}$.492†	1											
4 Job-relatedness Predictive	6.47	1.50	625	.667†	.787†	.314†	1										
5 Information Known	11.39	2.30	625	.624†	$.350^{\dagger}$.779†	$.220^{\dagger}$	1									
6 Chance to Perform	11.56	3.49	625	.789 [†]	.922†	.381†	$.678^{\dagger}$	$.275^{\dagger}$	1								
7 Feedback	11.50	2.59	625	$.567^{\dagger}$.259†	.782†	.155**	.345†	$.185^{\dagger}$	1							
8 Two-way Communication	11.79	1.65	625	.699 [†]	.565†	$.656^{\dagger}$	$.380^{\dagger}$	$.376^{\dagger}$	$.458^{\dagger}$	$.270^{\dagger}$	1						
9 Job-relatedness Content	7.29	1.66	625	.673 [†]	.713†	.418†	$.410^{\dagger}$	$.276^{\dagger}$	$.489^{\dagger}$.210 [†]	$.526^{\dagger}$	1					
10 Contribution to Process	3.27	1.07	625	$.602^{\dagger}$.538†	$.500^{\dagger}$.283†	$.384^{\dagger}$.339†	.344†	$.409^{\dagger}$.381 [†]	1				
11 Agreeableness	43.49	3.54	625	.214†	$.192^{\dagger}$	$.176^{\dagger}$.117**	$.147^{\dagger}$	$.184^{\dagger}$.087*	$.181^{\dagger}$.138 [†]	.129**	1			
12 Conscientiousness	42.21	4.49	625	.219†	.193†	.186†	$.147^{\dagger}$.164†	.153†	.138†	.107**	$.172^{\dagger}$.137†	$.358^{\dagger}$	1		
13 Extraversion	38.22	5.51	625	$.162^{\dagger}$.183†	.086*	$.142^{\dagger}$.102*	.159†	003	.116**	.156 [†]	.095*	$.214^{\dagger}$.339†	1	
14 Neuroticism	21.65	5.57	625	155†	135†	134†	099*	140†	091*	086*	067	109**	171†	317†	319†	325†	1
15 Openness	39.77	4.39	625	002	.011	018	.054	004	.033	051	.034	041	055	.152†	.091*	$.282^{\dagger}$	027

Note: PQ = Personality Questionnaire, RQ = Reaction Questionnaire, LAQ = Learning Agility Questionnaire, SD = standard deviation. [†]Correlation significant at p < .001 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed). **. Correlation is significant at the 0.01 level (2-tailed).
Table 11

Descriptive statistics and correlations between PQ and RQ variables for LPQ

Variable	Mean	SD	N	1	2	3	4	5	6	7	8	0	10	11	12	13	14
1 Total Reactions	63.51	9.75	625	1	2	5		5	0	7	0)	10	11	12	15	14
2 Structure Factor	37.32	6.82	625	.902†	1												
3 Social Factor	26.20	4.60	625	$.840^{\dagger}$.524 [†]	1											
4 Job-relatedness Predictive	6.44	1.55	625	.692†	.823†	.331†	1										
5 Information Known	11.44	2.32	625	.624†	.362†	.777†	$.206^{\dagger}$	1									
6 Chance to Perform	11.49	3.54	625	.799†	.926†	.414†	.711†	.296†	1								
7 Feedback	11.49	2.59	625	.574†	.288†	.772†	.179 [†]	.340†	.210 [†]	1							
8 Two-way Communication	11.72	1.73	625	.697†	.576†	.652†	.399†	.363†	.467†	.247†	1						
9 Job-relatedness Content	7.66	1.54	625	.693†	.705†	.481†	.466†	.304†	.485†	.256†	.579†	1					
10 Contribution to Process	3.27	1.08	625	.596†	.539†	$.500^{\dagger}$.313†	.384†	.342†	.348†	.387†	.375†	1				
11 Agreeableness	43.49	3.54	625	.186†	.161†	.165†	.094*	.137†	.154†	.087*	.157†	.115**	.124**	1			
12 Conscientiousness	42.21	4.49	625	.217†	.194†	$.184^{\dagger}$	$.176^{\dagger}$.153†	$.160^{\dagger}$.137†	.115**	.142†	$.140^{\dagger}$.358†	1		
13 Extraversion	38.22	5.51	625	.156†	$.187^{\dagger}$.072	.192†	.070	.155†	003	.117**	.135†	.102*	$.214^{\dagger}$.339†	1	
14 Neuroticism	21.65	5.57	625	148†	134†	124**	107**	121**	087*	086*	062	111**	174†	317†	319†	325†	1
15 Openness	39.77	4.39	625	006	.007	020	.045	003	.003	051	.022	.010	049	.152†	.091*	.282†	027

Note: PQ = Personality Questionnaire, RQ = Reaction Questionnaire, LPQ = Leadership Potential Questionnaire, SD = standard deviation. [†]Correlation significant at p < .001 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).

Table 12

Descriptive statistics and correlations between PQ and RQ variables for SJT

Variable	Mean	SD	Ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14
1 Total Reactions	63.18	10.21	625	1													
2 Structure Factor	37.17	7.43	625	.913†	1												
3 Social Factor	26.01	4.60	625	.832†	.534†	1											
4 Job-relatedness Predictive	6.41	1.63	625	.710 [†]	$.828^{\dagger}$.346†	1										
5 Information Known	11.21	2.52	625	.574†	.328†	$.744^{\dagger}$	$.202^{\dagger}$	1									
6 Chance to Perform	11.51	3.81	625	.825†	.937†	.437†	.721†	$.274^{\dagger}$	1								
7 Feedback	11.50	2.57	625	.531†	.265†	.742 [†]	.151†	.247†	$.190^{\dagger}$	1							
8 Two-way Communication	11.69	1.78	625	.725†	.632†	.643†	.453†	$.288^{\dagger}$	$.547^{\dagger}$.259†	1						
9 Job-relatedness Content	7.56	1.64	625	.717†	.738†	.483†	.519†	.257†	.545†	.259†	.599†	1					
10 Contribution to Process	3.30	1.07	625	.589†	.519 [†]	.515†	$.288^{\dagger}$	$.370^{\dagger}$.358†	$.350^{\dagger}$.397†	.349†	1				
11 Agreeableness	43.49	3.54	625	.162†	.158†	.121**	.086*	.065	.151†	.081*	.126**	.127**	.125**	1			
12 Conscientiousness	42.21	4.49	625	.188†	.163†	$.169^{\dagger}$.128**	.110**	.134†	.134†	.117**	.143†	.126**	.358†	1		
13 Extraversion	38.22	5.51	625	.111**	.108**	.083*	.067	.084*	.085*	004	.116**	.121**	.086*	.214†	.339†	1	
14 Neuroticism	21.65	5.57	625	166 [†]	147†	146 [†]	072	137†	131**	086*	086*	103**	182†	317 [†]	319†	325†	1
15 Openness	39.77	4.39	625	.005	.021	018	.052	014	.025	051	.042	.007	046	.152†	.091*	$.282^{\dagger}$	027

Note: PQ = Personality Questionnaire, RQ = Reaction Questionnaire, SJT = Situational Judgement Test, SD = standard deviation. [†]Correlation significant at*p*<.001 level (2-tailed).**Correlation is significant at the 0.01 level (2-tailed).

Hierarchical regression analyses

Candidate reactions to the total process

Tables 13 shows the results for the regressions to test the hypothesised relationships for the individual personality factors and reactions to the total process (where significant correlations were found). Model 1 tested hypothesis 1a, that Agreeableness will be a significant predictor of candidate reactions, in that as levels of Agreeableness increase candidates will react more positively to the process. The regression showed that Agreeableness explained a significant proportion of the variance (1.9%) beyond the control variables ($\beta = .14$, t = 3.4, p < .001), therefore hypothesis 1a was supported. Model 2 tested hypothesis 1b, which stated that Agreeableness will be a significant predictor of the social factor of the process, namely that as levels of Agreeableness increase candidate reactions to the social factor of the process will be more positive. Model 2 testing hypothesis 1b showed that Agreeableness explained a significant proportion of the variance of the social factor (.8%) beyond the control variables $(\beta = .09, t = 2.16, p = .03)$, therefore hypothesis 1b was supported. Model 3 tested hypothesis 1d, which stated that Agreeableness will a significant predictor of the face validity of the process, namely that as Agreeableness increased the reaction to the face validity of the process will be more positive. Model 3 showed that Agreeableness did not explain a significant proportion of the variance of face validity beyond the control variables ($\beta = .08$, t=1.96, p=.05), therefore hypothesis 1d was not supported.

Model 4 tested hypothesis 2a, that Neuroticism will be a significant predictor of candidate reactions, in that as levels of Neuroticism increase candidates will react more negatively to the process. The regression showed that Neuroticism explained a significant proportion of the variance (1.8%) beyond the control variables ($\beta = -.14$, t = -3.28, p < .001), therefore hypothesis 2a was supported. Models 5 and 6 tested hypotheses 2b and 2d, which state that Neuroticism will a significant predictor of specific elements of candidates' reactions, namely that as levels of Neuroticism increase candidate reactions to the social factor of the process and the face validity of the process will be more negative. Model 5 testing hypothesis 2b showed that Neuroticism explained a significant proportion of the variance in the social factor (1.7%) of reactions beyond the control variables ($\beta = -.14$, t = -3.19, p < .001), therefore hypothesis 2b was supported. Model 6 testing hypothesis 2d showed that Neuroticism did not explain a significant proportion of the variance of face validity beyond the control variables ($\beta = -.06$, t = -1.48, p = .14), therefore hypothesis 2d was not supported.

Models 7 and 8 tested hypotheses 4a and 4b, that Conscientiousness would be a significant predictor of overall candidate reactions and the social factor of candidate reactions. Model 7 testing hypothesis 4a showed that Conscientiousness explained a significant proportion of the variance in total candidate reactions (3.7%) beyond the control variables ($\beta = .20$, t = 4.77, p < .001), therefore hypothesis 4a was supported. Model 8 testing hypothesis 4b showed that Conscientiousness explained a significant proportion of the variance in the social factor of reactions (1.4%) beyond the control variables ($\beta = .12$, t = 2.90, p < .05) therefore hypothesis 4b was supported.

Table 13

Hierarchical regression testing associations between Agreeableness, Neuroticism and Conscientiousness and Candidate Reactions to Total Process

Variables and Model Stats	Model 1:	Total Read	ctions	Model	2: Social F	Factor	Model	3: Job-related Content	ness	Model 4	: Total Re	eactions	Model	5: Social	Factor	Mo related	del 6: Job iness Con	- tent	Mo	odel 7: To Reactions	otal s	Mode	l 8: Socia	al Factor
Predictors	В	SE	β	В	SE	β	В	SE	β	В	SE	β	В	SE	β	В	SE	β	В	SE	β	В	SE	β
Controls																								
Gender	-0.07	0.81	0	-0.16	0.52	01	0.05	0.14	.01	-0.68	0.81	03	-0.47	0.51	04	-0.01	0.14	0	0.17	0.81	.01	-0.07	0.52	01
Age	0.33	0.24	.06	0.09	0.16	.03	0.03	0.04	.03	0.26	0.25	.05	0.03	0.16	.01	0.02	0.04	.02	0.31	0.24	.05	0.09	0.15	.02
Ethnicity	0.41	1.16	.01	-1.12	0.74	06	0.19	0.21	.04	-0.24	1.17	01	-1.47	0.74	08*	0.13	0.21	.03	0.15	1.15	.01	-1.23	0.74	07
SEB	-0.11	0.50	01	-0.31	0.32	04	0.01	0.09	.01	-0.16	0.5	01	-0.32	0.31	04	0.01	0.09	0	-0.02	0.49	0	-0.27	0.32	04
GIO	3.69	1.76	.09*	2.13	1.12	.08	0.33	0.31	.04	4.06	1.76	.09*	2.29	1.12	.08*	0.37	0.31	.05	4.51	1.75	.10*	2.45	1.12	.09*
Exercise Performance Independent variables	1.62	0.74	.09*	1.05	0.47	.09*	0.14	0.13	.04	1.71	0.74	.10*	1.01	0.47	.09*	0.15	0.13	.05	1.34	0.74	.08	0.94	0.47	.09*
Agreeableness	0.4	0.12	$.14^{\dagger}$	0.16	0.07	.09*	0.04	0.02	.08	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Neuroticism	-	-	-	-	-	-	-	-	-	-0.25	0.08	14^{+}	-0.15	0.05	14†	-0.02	0.01	06	-	-	-	-	-	-
Conscientiousness	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	0.44	0.09	$.20^{\dagger}$	0.17	0.06	.12**
R squared (R ²)	.0	47 (4.7%)		.(035 (3.5%)			.014 (1.4%)		.(46 (4.6%)	.0)44 (4.4%)	.0	12 (1.2%)		.()65 (6.5%	5)		.041 (4.19	%)
Adjusted $R^2 (\Delta R^2)$.0	35 (3.5%)			023 (2.3%)			.002 (0.2%)		.()34 (3.4%)	.0)32 (3.2%)	.0	01 (.1%)		.()54 (5.4%	5)		.029 (2.9	%)
R ² Change	.0	19 (1.9%)			.008 (.8%)			.007 (0.7%)		.(018 (1.8%)	.0	017 (1.7%)	.00	04 (.04%)		.()37 (3.7%	5)		.014 (1.4	%)
Significance of model	F (7,	572) = 4.04	4 [†]	F (7,	572) = .2.9)3**	F(7,572) = 1.198		F (7	,572) = 3.	92†	F(7,	,572) = 3.	.74†	F(7,	572) = 0.9	96	F(7	,527) = 5	$.68^{\dagger}$	F(7,	,572) = 3	3.48**

Note: SEB = Socio-economic background, GIO = Guaranteed Interview Offer, B = unstandardized regression coefficient, SE = standard error, β = standardized regression coefficient.

*significant at p < .05 level, **significant at p < .01 level, †significant at p < .001 level.

Table 14 shows the results for the predicted interaction effects between personality factors and candidate reactions to the total process. In line with Iacobucci et al (2016), the personality factors entered into the regression were centred around the mean to avoid multicollinearity. Model 9 testing hypothesis 5a showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = .00, *t* = -0.11, *p* = .91), therefore hypothesis 5a was not supported. Model 10 testing hypothesis 5b showed that the interaction effect of Extraversion, Neuroticism, Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = .06, *t* = 1.22, *p* = .02), therefore hypothesis 5b was not supported.

Table 14

X 11 1X 110/		110 7 (1.D. C	Madal 10. Tatal Daratiana						
Variables and Model Stats	Moc	iel 9: Tota	I Reactions	Model	10: Total I	Reactions				
Predictors	В	SE	В	В	SE	β				
Controls										
Gender	0.19	0.81	.01	-0.34	0.82	02				
Age	0.3	0.24	.05	0.23	0.25	.04				
Ethnicity	0.07	1.15	0	-0.03	1.17	0				
SEB	0.03	0.5	0	-0.05	0.5	0				
GIO	4.58	1.75	.11**	3.91	1.76	.09*				
Exercise Performance	1.26	0.75	.07	1.33	0.75	.08				
Independent variables										
Conscientiousness	0.41	0.1	.19†	-	-	-				
Extraversion	0.08	0.08	.04	-	-	-				
Extraversion*Conscientiousness	0	0.02	0	-	-	-				
Extraversion	-	-	-	0.11	0.08	.06				
Neuroticism	-	-	-	-0.18	0.08	10*				
Agreeableness	-	-	-	0.35	0.13	.13**				
Extraversion*Neuroticism*Agreeableness	-	-	-	0	0	.06				
R squared (R ²)		.067 (6.	.7%)	.062 (6.2%)						
Adjusted $R^2 (\Delta R^2)$.052 (5.	.2%)	.045 (4.5%)						
R ² Change		.039 (3.	.9%)	.034 (3.4%)						
Significance of model]	F (9,570)	$= 4.52^{\dagger}$	$F(10,569) = 3.74^{\dagger}$						

Hierarchical regression model testing the associations between the interaction Personality Factors and Candidate Reactions to the Total Process

Note: SEB = Socio-economic background, GIO = Guaranteed Interview Offer, B = unstandardized regression coefficient, SE = standard error, β = standardized regression coefficient. *significant at *p* <.05 level, **significant at *p* <.01 level, *significant at *p* <.00 level.

Candidate reactions to the individual exercises

Additional models tested the hypothesised relationships between personality factors and candidate reactions for each of the exercises. The summary of results for this analysis are shown below; full tabular outputs of the regression testing can be found in the Supplementary Materials.

Model 12 tested hypothesis 1a, the association between Agreeableness and overall candidate reactions by exercises. Model 12a showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (2.4%) beyond the control variables ($\beta = .17, t = 3.92, p < .001$) for the CS. Model 12b showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (3.5%) beyond the control variables ($\beta = .19, t = 4.65, p < .001$) for the LAQ. Model 12c showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (2.6%) beyond the control variables ($\beta = .19, t = 4.65, p < .001$) for the LAQ. Model 12c showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (2.6%) beyond the control variables ($\beta = .16, t = 3.96, p < .001$) for the LPQ. Model 12d showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (2.4%) beyond the control variables ($\beta = .16, t = 3.96, p < .001$) for the LPQ. Model 12d showed that Agreeableness explained a significant proportion of the variance in total candidate reactions (2.4%) beyond the control variables ($\beta = .16, t = 3.96, p < .001$) for the SJT. Therefore, hypothesis 1a was supported for all exercises.

Model 13 tested hypothesis 1b, the association between Agreeableness and the social factor of candidate reactions by exercises. Model 13a showed that Agreeableness explained a significant proportion of the variance in the social factor of candidate reactions (2.0%) beyond the control variables ($\beta = .15$, t = 3.42, p < .05) for the CS. Model 13b showed that Agreeableness explained a significant proportion of the variance in social factor of candidate reactions (1.9%) beyond the control variables ($\beta = .14$, t = 3.39, p < .05) for the LAQ. Model 13c showed that Agreeableness explained a significant proportion of the variance in the social factor of candidate reactions (1.9%) beyond the control variables ($\beta = .14$, t = 3.39, p < .05) for the LAQ. Model 13c showed that Agreeableness explained a significant proportion of the variance in the social factor of candidate reactions (1.9%) beyond the control variables ($\beta = .14$, t = 3.41, p < .05) for the LPQ. Model 13d showed that Agreeableness explained a significant proportion of the variance in the social factor of candidate reactions (1.3%) beyond the control variables ($\beta = .11$, t = 2.74, p < .05) for the SJT. Therefore, hypothesis 1b was supported for all exercises.

Model 14 tested hypothesis 1c, the association between Agreeableness and perceived predictive validity by exercise. Model 14a showed that Agreeableness explained a significant proportion of the variance in perceived predictive validity (.9%) beyond the control variables ($\beta = .09, t = 2.20, p = .03$) for the CS. Model 14b showed that Agreeableness explained a

significant proportion of the variance in perceived predictive validity of candidate reactions (.9%) beyond the control variables ($\beta = .10, t = 2.28, p = .02$) for the LAQ. Model 14c showed that Agreeableness did not explain a significant proportion of the variance in perceived predictive validity beyond the control variables ($\beta = .06, t = 2.53, p = .13$) for the LPQ. Model 14d showed that Agreeableness did not explain a significant proportion of the variance in perceived predictive validity beyond the control variables ($\beta = .06, t = 2.53, p = .13$) for the LPQ. Model 14d showed that Agreeableness did not explain a significant proportion of the variance in perceived predictive validity beyond the control variables ($\beta = .07, t = 1.74, p = .08$) for the SJT. Therefore, hypothesis 1c was partially supported for the exercises.

Model 15 tested hypothesis 1d, the association between Agreeableness and face validity by exercise. Model 15a showed that Agreeableness explained a significant proportion of the variance in face validity (1.3%) beyond the control variables (β = .12, *t* = 2.68, *p* = .01) for the CS. Model 15b showed that Agreeableness explained a significant proportion of the variance in face validity of candidate reactions (1.5%) beyond the control variables (β = 0.13, *t* = 2.96, *p* < .05) for the LAQ. Model 15c showed that Agreeableness explained a significant proportion of the variance in face validity (1.0%) beyond the control variables (β = .10, *t* = 2.48, *p* = .01) for the LPQ. Model 14d showed that Agreeableness explained a significant proportion of the variance in face validity (1.3%) beyond the control variables (β = .11, *t* = 2.70, *p* = .01) for the SJT. Therefore, hypothesis 1d was supported for all exercises.

Model 16a tested hypothesis 2a, the association between Neuroticism and total candidate reactions by exercise. Model 16a showed that Neuroticism explained a significant proportion of the variance in total candidate reactions (.9%) beyond the control variables ($\beta = -.10$, t = -2.30, p = .02) for the CS. Model 16b showed that Neuroticism explained a significant proportion of the variance in total candidate reactions (1.2%) beyond the control variables ($\beta = -.12$, t = -2.72, p = .01) for the LAQ. Model 16c showed that Neuroticism explained a significant proportion of the variance in total candidate reactions (1.1%) beyond the control variables ($\beta = -.11$, t = -2.59, p = .01) for the LPQ. Model 16d showed that Neuroticism explained a significant proportion of the variance in total candidate reactions (1.1%) beyond the control variables ($\beta = -.13$, t = -3.17, p < .05) for the SJT. Therefore, hypothesis 2a was supported for all exercises.

Model 17 tested hypothesis 2b, the association between Neuroticism and the social factor of candidate reactions by exercise. Model 17a showed that Neuroticism explained a significant proportion of the variance in the social factor of candidate reactions (.7%) beyond the control

variables ($\beta = -.09$, t = -1.98, p = .05) for the CS. Model 17b showed that Neuroticism explained a significant proportion of the variance in social factor of candidate reactions (1.0%) beyond the control variables ($\beta = -.11$, t = -2.52, p = .01) for the LAQ. Model 17c showed that Neuroticism explained a significant proportion of the variance in the social factor of candidate reactions (1.0%) beyond the control variables ($\beta = -.11$, t = -2.51, p = .01) for the LPQ. Model 17d showed that Neuroticism explained a significant proportion of the variance in the social factor of candidate reactions (1.5%) beyond the control variables ($\beta = -.13$, t = -3.01, p < .05) for the SJT. Therefore, hypothesis 2b was supported for all exercises.

Model 18 tested hypothesis 2c, the association between Neuroticism and perceived predictive validity for the CS, LAQ, LPQ. Model 18a showed that Neuroticism explained a significant proportion of the variance in perceived predictive validity (.8%) beyond the control variables ($\beta = -.09, t = -2.09, p = .04$) for the CS. Model 18b showed that Neuroticism did not explain a significant proportion of the variance in perceived predictive validity of candidate reactions beyond the control variables ($\beta = -.06, t = -1.47, p = .14$) for the LAQ. Model 18c showed that Neuroticism did not explain a significant proportion of the variance proportion of the variance in perceived predictive validity of the LAQ. Model 18c showed that Neuroticism did not explain a significant proportion of the variance in perceived predictive validity beyond the control variables ($\beta = -.08, t = -1.87, p = .06$) for the LPQ. Therefore, hypothesis 2c was partially supported for the exercises.

Model 19 tested hypothesis 2d, the association between Neuroticism and face validity by exercise. Model 19a showed that Neuroticism did not explain a significant proportion of the variance in face validity beyond the control variables ($\beta = -.08$, t = -1.81, p = .07) for the CS. Model 19b showed that Neuroticism explained a significant proportion of the variance in face validity of candidate reactions (.7%) beyond the control variables ($\beta = -.09$, t = -2.08, p = .04) for the LAQ. Model 19c showed that Neuroticism did not explain a significant proportion of the variance in face validity beyond the control variables ($\beta = -.07$, t = -1.62, p = .11) for the LPQ. Model 19d showed that Neuroticism did not explain a significant proportion of the variance in face validity (1.3%) beyond the control variables ($\beta = -.07$, t = -.07, t = -1.65, p = .10) for the SJT. Therefore, hypothesis 2d was partially supported for all exercises.

Model 20 tested the hypothesised relationships for Openness to Experience and face validity for the CS. Model 20 showed that Openness to Experience did not explain a significant proportion of the variance in face validity beyond the control variables ($\beta = .04$, t = 0.94, p = .35) for the CS. Hypothesis 3d was therefore not supported for the exercises.

Model 21 tested hypothesis 4a, the association between Conscientiousness and overall candidate reactions by exercise. Model 21a showed that Conscientiousness explained a significant proportion of the variance in total candidate reactions (3.5%) beyond the control variables ($\beta = .19$, t = 4.35, p < .001) for the CS. Model 21b showed that Conscientiousness explained a significant proportion of the variance in total candidate reactions (4.6%) beyond the control variables ($\beta = .23$, t = 5.36, p < .001) for the LAQ. Model 21c showed that Conscientiousness explained a significant proportion of the variance proportion of the variance in total candidate reactions (4.8%) beyond the control variables ($\beta = .23$, t = 5.36, p < .001) for the variance in total candidate reactions (4.8%) beyond the control variables ($\beta = .23$, t = 5.43, p < .001) for the LPQ. Model 21d showed that Conscientiousness explained a significant proportion of the variance in total candidate reactions (4.5%) beyond the control variables ($\beta = .22$, t = 5.25, p < .001) for the SJT. Therefore, hypothesis 4a was supported for all exercises.

Model 22 tested hypothesis 4b, the association between Conscientiousness and the social factor of candidate reactions by exercise. Model 22a showed that Conscientiousness explained a significant proportion of the variance in the social factor of candidate reactions (2.6%) beyond the control variables ($\beta = .16$, t = 3.90, p < .001) for the CS. Model 22b showed that Conscientiousness explained a significant proportion of the variance in social factor of candidate reactions (3.0%) beyond the control variables ($\beta = .18$, t = 4.27, p < .001) for the LAQ. Model 22c showed that Conscientiousness explained a significant proportion of the variance proportion of the variance in the social factor of candidate reactions (3.5%) beyond the control variables ($\beta = .20$, t = 4.63, p < .001) for the LPQ. Model 22d showed that Conscientiousness explained a significant proportion of the variance in the social factor of candidate reactions (3.6%) beyond the control variables ($\beta = .19$, t = 4.66, p < .001) for the SJT. Therefore, hypothesis 4b was supported for all exercises.

Model 23 tested the interaction effect of Conscientiousness and Extraversion on overall candidate reactions by exercise. Model 23a showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = .03, *t* = 0.59, *p* = .55) for the CS. Model 23b showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = .03, *t* = 0.59, *p* = .55) for the CS. Model 23b showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = -.02, *t* = -0.50, *p*= .62) for the LAQ. Model 23c showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = -.02, *t* = -0.50, *p*= .62) for the LAQ. Model 23c showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in the variance in the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversion did not explain a significant proportion of the variance in Conscientiousness and Extraversio

total candidate reactions beyond the control variables or the individual personality factors (β = .00, *t* = -.01, *p* = .99) for the LPQ. Model 23c showed that the interaction effect of Conscientiousness and Extraversion did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors (β = -.01, *t* = -0.97, *p* = .33) for the SJT. Therefore, hypothesis 5a was not supported for any of the exercises.

Model 24 tested the interaction effect of Extraversion, Neuroticism, Agreeableness on overall candidate reactions by exercise. Model 24a showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .01, t = 0.13, p = .89$) for the CS. Model 24b showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .01, t = 0.20, p = .84$) for the LAQ. Model 24c showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .01, t = -0.20, p = .84$) for the LAQ. Model 24c showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .01, t = 0.15, p = .88$) for the LPQ. Model 24c showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .01, t = 0.15, p = .88$) for the LPQ. Model 24c showed that the interaction effect of Extraversion, Neuroticism and Agreeableness did not explain a significant proportion of the variance in total candidate reactions beyond the control variables or the individual personality factors ($\beta = .04, t = 0.87, p = .39$) for the SJT. Therefore, hypothesis 5b was not supported for any of the exercises.

Discussion

Discussion of findings

The aim of this study was to examine the impact of personality on candidate reactions to an online selection process for talent development schemes. The specific focus was the reactions of senior, internal candidates to technology-enabled selection methods. The results indicated that personality factors do have an impact on candidate reactions, however there are variations in the impact personality has across the five factors and some differences between

reactions to individual exercises. The results showed that Agreeableness, Conscientiousness and Neuroticism had an impact on candidate reactions, but not systematically.

Hypothesis 1 - 3 tested the relationships between Agreeableness, Neuroticism and Openness to Experience and overall candidate reactions, reactions to the social factor of candidate reactions, perceived predictive validity and face validity. In line with previous studies (Truxillo et al., 2006; Oostrom et al., 2010; Bernerth et al., 2006) Hypotheses 1a - 1d were largely supported, as significant positive correlations were found for all hypothesised relationships, with the exception of the perceived predictive validity of the total process.

The regression analysis showed that a significant proportion of the variance in overall reactions and the social factor of reactions was explained by Agreeableness once sociodemographic characteristics and test performance were controlled for. This was true for both the total process and the individual exercises. Agreeableness also explained a significant proportion of the variance in face validity of the exercises once other factors were controlled for and explained a significant proportion of the variance in the perceived predictive validity of the CS and the LAQ (but not the LPQ or the SJT). This differs from Oostrom et al., (2010) who did find positive relationships between Agreeableness and perceived predictive validity, in this case of a cognitive ability test and an SJT. This suggests that the impact of Agreeableness on perceptions of the predictive validity of the selection process is dependent on the exercise content. When comparing reactions to the elements of the selection process, the CS had the most favourable reactions in terms of perceived predictive validity and face validity, and reactions to the process overall were generally lower than for the individual exercises. This suggests that there may be a threshold at which Agreeableness begins to have an impact on how predictive of future success candidates perceive elements of a selection process to be.

Hypotheses 2a - 2d were largely supported, in line with previous studies (Truxillo et al., 2006; Oostrom et al., 2010). Neuroticism was found to be significantly negatively related to total candidate reactions, the social factor of reactions and the face validity of the process and all four exercises. The relationship with the perceived predictive validity of the process was not significant, nor was the relationship with the perceived predictive validity of the SJT. In the regression analysis, Neuroticism explained a significant proportion of the variance in total reactions and the social factor of reactions when controlling for the socio-demographic

characteristics and test performance. This held for the total process and for the exercises. Neuroticism explained a significant proportion of the variance in the perceived predictive validity of the CS and in the face validity of the LAQ. As with Agreeableness, this suggest that the impact of Neuroticism will vary depending on generally how valid the selection process content is seen to be.

There was very limited support for hypotheses 3a - 3d, which investigated the relationship between Openness to Experience and candidate reactions. Truxillo et al. (2006) and Oostrom et al. (2010) found significant relationships between Openness to Experience and various aspects of candidate reactions, and Hausknecht et al. (2004) predicted that there would be a relationship, particularly with newer selection methods. The only significant relationship found was with the perceived predictive validity of the CS, however once the sociodemographic characteristics and exercise performance were controlled for this was no longer significant. It is possible that Openness to Experience did not play a role in this study as the participants had some familiarity with this type of selection process and did not see it as particularly novel. As middle and senior managers in the organisation it is highly likely that they will have familiarity with the types of selection exercise employed by in the current study, either as candidates themselves going through a selection process for their current role or as managers using such exercises to recruit employees into their teams. Oostrom et al.'s (2010) study made use of students and Truxillo et al., (2006) studied applicants to the police; while details of their familiarity with the selection processes are not available is it reasonable to assume that they will have had less exposure to different selection methods than the participants in the current study. Further research may help establish whether the impact of Openness to Experience differs depending on factors such as position in the organisation or years of work experience.

There was clear support for hypotheses 4a and 4b, which investigated the relationship between Conscientiousness and candidate reactions. In line with Hausknecht et al. (2004) and Truxillo et al. (2006), there was a significant positive relationship between Conscientiousness and total candidate reactions and the social factor of candidate reactions for both the overall process and the individual exercises. Conscientiousness explained a significant proportion of the variance in reactions once socio-demographic characteristics and performance was controlled for. It is notable that several previous studies investigating the relationships between personality and candidate reactions have omitted Conscientiousness from their studies (e.g., Bernerth et al., 2006; Van Vianen et al., 2004), however in this study Conscientiousness had the strongest relationships with candidate reactions. Future researchers are therefore encouraged to include the full five-factor model of personality in their research so that potentially significant findings are not missed. Conscientiousness is typically related to being planned, organised, reliable and dutiful. With regard to the current selection process, the organisation employs various communication strategies about the selection process prior to the applications opening, with various briefing sessions offered and information made available via internal online communication channels. Further, information is shared with participants as they work through the selection process regarding where they are in the process and what the next exercise entails. Each exercise (with the exception of the candidate statement) has several instruction pages and example questions prior to starting the tests. It may be that those who are higher on conscientious plan their application, taking time to engage with all the information channels and therefore perceive the process to be more fair. Those who are lower on Conscientiousness may be less planful in their approach and find the information overwhelming, which may lead them to disengage and perceive the process as less fair.

The final hypotheses (5a and 5b) investigated the interaction effects of Extraversion with other personality factors on candidate reactions. These interaction effects were largely unsupported in the current study. Honkaniemi et al. (2013) found a significant interaction effect of Conscientiousness and Extraversion and a significant interaction effect of Neuroticism, Extraversion and Agreeableness, however these were not found in the current study. This lack of support for Honkaniemi et al.'s (2013) findings may have been partly due to the differences in methods, as Honkaniemi et al. used cluster analysis to create profile types whereas the current study used simple interactions of personality factor scores.

Strengths and limitations

A key strength of the current study is the use of real-world applicants. It has been noted by previous researchers (e.g., Hiemstra et al., 2019) that a typical methodology for applicant reactions research is to use short descriptions of hypothetical situations, however, it is unlikely that this will be a valid proxy for the real high-stakes experience of going through a selection process. Further, as far as is known, this is the first study investigating reactions of more senior internal candidates.

The use of Gilliland's (1993) model of candidate reactions (measured by the adapted version of the SPJS) provides a comprehensive view of reactions to all aspects of the selection process. While the current study focused on testing the hypothesised relationships, the reporting of data for all elements of Gilliland's model means that future research can build on this to focus on specific aspects of reactions. For example, the new 'contribution to process' scale provides information to practitioners regarding how well candidates understand the combining of exercises into a single outcome.

This study was not without limitations. First, the small effect sizes and small amount of variance explained by the personality factors could be raised as having limited practical significance. However, recent research has suggested that Cohen's effect size boundaries (typically considered as small being r = .10, medium being .30 and large being .50) as unsuitable for social science research (Funder & Ozer; 2019). Bosco et al. (2015) proposed more appropriate benchmarks for use in applied psychology research. They suggest that for research involving the prediction of behaviours, medium effect sizes are between .1 and .25 and medium effects for research involving attitudes are between .2 and .4. Further, Gignac and Szodorai (2016) suggest, based on a meta-analysis of 708 studies, benchmarks of .10, .20 and .30 for individual differences research. The findings for the current study would therefore fall within the small-medium range for uncorrected corelations. Funder and Ozer (2019) suggest that a sense of magnitude can be gained by comparing these effect sizes with intuitively understood relationships outside of the psychology field, giving an example of the correlation between typical pain relivers and effective alleviation of headaches as .14. They also note the importance of small effects cumulating over time, suggesting that in personality research stable traits that have small effects on what an individual does can add up very quickly

While a deliberate choice in the design of the study, it is also somewhat of a limitation that participants did not know the outcomes of their performance when responding to the RQ. Previous studies which have included feedback have found that the selection decision has a significant impact on reactions (Wiechmann & Ryan, 2003). It would be interesting to investigate whether this impact differed depending on personality traits.

Finally, as with any study based in a single organisation, there are limitations with the generalisability of the results to other settings. While the selection process of interest is unique to the Civil Service, it does reflect the content of selection processes currently used by other large-scale organisations, meaning lessons learned for the current study are likely to be applicable to a range of other organisations and selection settings.

Future research

While a number of controls were considered in the study, these were limited to sociodemographic factors and exercise performance. Other studies have included attitudinal measures such as test taking motivation. While it could be assumed that the high level of competition in current selection processes is likely to mean all participants would be highly motivated, this was not formally tested in the current design. Future research could include additional variables in the model to identify whether these moderate or mediate the effects of personality factors.

The current study used a short five-factor measure of personality. In order to further explore the effect of personality on candidate reactions, future studies could use a more granular measure of personality to identify sub-scale differences or use an alternative to the five-factor model. The current study investigated one stage of a two-stage selection process, with the second stage being an interview. It would be possible for future research to consider monitoring the impact of personality before and after each stage.

Implications for theory and practice

Implications for Theory

This study makes a significant contribution to the evidence base of candidate reactions research. Despite several calls over the preceding years (e.g., Ryan & Ployhart, 2000, Hausknecht et al., 2004), research into the impact of personality factors on candidate reactions remains limited. Where research does exist, there is limited replicability due to the range of selection methods available. By investigating reactions to four different exercises as well as the process as a whole, and by considering the whole procedural justice framework, this study has significantly expanded the knowledge base in application reactions research. This study also demonstrated the continued relevance of Gilliland's (1993) procedural justice framework. Despite its development being prior to the introduction of technology-enabled

selection processes, it has been shown to be applicable in current selection settings. However, as noted in more recent research (e.g., Wood's et al, 2020), there is scope to build on the existing framework to account for more recent developments in technology and to develop existing questionnaires (such as the SPJS; Bauer et al., 2001) to measure newer frameworks. The current study excluded elements around the administration of the selection process which are unlikely to apply in a large number of newer selection processes (e.g., the item '*Test administrators made no distinction in how they treated applicant*' will not apply to any selection process where candidates do not interact directly with administrators); less relevant scales and/or items could be removed and new scales incorporated which assess aspects of reactions more relevant to current methods. For example, in 1993 Ball et al. suggested that understanding 'procedural pain' such as the risk of discomfort or humiliation in the process could be part of candidate reactions; this could be developed to consider the element of 'creepiness' that has been described as a concern with some newer technologies (e.g., Langer et al., 2018; Tene & Polonetsky, 2013).

Implications for Practice

Competition for talented employees is high, and organisations must ensure they keep candidates engaged and feeling positive during their selection processes, or risk losing them to competitors. This can become a particular issue when recruiting for senior level or leadership positions, where applicant pools are often already limited, and organisations may struggle to attract diverse candidate groups. The difference in reactions to the different elements of the selection process will be of direct interest to practitioners. Despite its preponderance in use in selection, candidate statements are not frequently included in inventories of selection methods, therefore the finding that this may be more positively received than other selection methods, regardless of personality, is something practitioners may wish to consider in their selection process.

The current study found differences in reactions based on the social factors of the process which is concerned with communication between the organisation and candidates. Giving focus to the amount and content of messaging about the selection process and potentially trialling messaging with different candidate groups could be a relatively straightforward way of improving reactions without the need for practitioners to make significant changes to the content and structure of the process. One risk of focussing on improving reactions to a selection process may be that reliability and validity is compromised. For example, Anderson et al.'s meta-analysis (2010) showed that cognitively ability tests are less preferred by candidates than resumes, despite having much stronger predictive validity (e.g., Jackson et al., 2018; Schmidt, & Hunter, 1998). Practitioners should therefore ensure that they are applying best practice in selecting or developing tools that are valid and reliable as well as considering how well they will be received by candidates.

As noted, the relationship between personality factors differed by exercises, suggesting that variance in reactions may be a factor of both personality factors and exercises content. Often, organisations use a 'hurdle' approach to selection (Mendoza et al., 2004), where candidates must pass each exercise in turn before progressing to the next. This means that candidates may be put off by exercises early in the selection process before they reach later types of exercises which they view more favourably. The process used in the current study applied multiple exercises in a single stage, meaning there was potential for exercises that are less favourable to those lower on Conscientiousness and Agreeableness and higher on Neuroticism to be 'offset' by elements that were more favourable. Practitioners could therefore benefit from including multiple exercises in a single stage early in the selection process.

Conclusion

Personality factors play a significant role in reactions to selection processes, having an impact on reactions beyond differences that may be due to socio-demographic factors. Practitioners would benefit from considering how their selection process will be perceived by potential candidates with traits that are desirable to the job role in question, and should consider using selection methods which comprise different exercises in a single early stage in order to maximise their chances of attracting the broadest range of potential talent.

Chapter 5 - Discussion General

This thesis examined the impact of individual differences on candidate reactions to selection processes. There have been numerous calls for research in this area over the previous decades: in their 2000 review Ryan and Ployhart noted the lack of research on the impact of individual differences in candidate reactions research; in 2004 Hausknecht et al. suggested further research into personality differences was needed; in the same year Chan and Schmitt highlighted the need for research into applicant characteristics and also reactions to newer technologies in selection; and in 2017 McCarthy et al.'s meta-analysis supported the ongoing relevance of research into candidate reactions, noting that the effect of personality is an emerging consideration. Yet, despite these calls the research examining the impact of individual differences on candidate reactions to selection processes has remained scant. The studies presented in this thesis aimed to address this gap.

In Study 1, a systematic literature review was carried out to understand what is currently known and unknown regarding the role individual differences play in candidate reactions. Study 2 aimed to address some of the gaps identified in the literature review by exploring the impact of individual differences on candidate reactions to a selection process for talent development schemes using Gilliland's (1993) model of procedural justice as the basis of measuring candidate reactions.

A synthesis of findings from both studies is presented here, followed by a discussion of the practical and theoretical implications of these findings, limitations of the research, and suggestions for further research.

Synthesis of findings from Study 1 and Study 2

Study 1 identified 18 papers which met the inclusion criteria. The papers were organised by the types of individual differences investigated, which covered three broad categories of attitudinal, biographical, and personality factors. In some papers the individual differences were mediators, moderator, or controls rather than predictors. Study 2 focused on personality

factors as the mechanism, with biographical factors considered as control variables. No attitudinal factors were considered in Study 2 due to practical constraints.

The following section synthesises the findings from Study 1 and Study 2 to discuss the role different aspects of individual difference (Attitudinal, Biographical and Personality) have on individual differences.

Summary of Findings

Table 15 below presents the high-level findings from study 1 and study 2 against the CIMO framework:

Table 15

Summary of findings from Study 1 and Study 2 against the CIMO framework

Framework	SLR (Study 1)	Empirical Study (Study 2)
Study aim	To understand existing research regarding the impact that individual differences such as personality and biographical factors have on	To investigate the impact of personality on candidate reactions to an online selection process
	candidates' reactions to selection procedures.	for talent development schemes.
Context	Median sample of 218, 45% female across studies, even split of field and lab studies.	Sample 625, 57.3% female, field study.
Intervention	Cognitive ability test, personality tests, interviews, ACs, SJTs, resumes, written tests	4 exercise process comprising candidate statement, two questionnaires and an SJT
Mechanisms	Biographical factors (gender, age, ethnicity, disability, educational attainment, job role, GMA). Personality (Big Five). Attitudes (employment commitment, self-efficacy, motivational control, test taking motivation, core-self evaluations, trait affectivity, well-being, and locus of control)	Personality (Big Five), gender, age, ethnicity, SEB, disability, test performance, previous test experience.
Outcomes	Measured by the TAS, SPJS and other researcher-developed questionnaires. The findings suggest that there is promising evidence for the impact of attitudinal, biographical and personality factors on candidate reactions, however papers are of mixed quality and there is little replicability of study methodologies.	Measured by the SPJS (adapted version). The findings suggest that the personality factors of Conscientious, Agreeableness and Neuroticism play a small but significant role in candidate reactions, and the impact of these factors varies depending on the selection method used.

Note: CIMO = Context, intervention, Mechanism and Outcomes, SLR = Systematic Literature Review, ACs = Assessment Centres, SJTs = Situational Judgement Tests, GMA = General Mental Ability, SEB = Socio-economic background, TAS = Test Attitude Survey, SPJS = Selection Procedural Justice Scale

The role of attitudinal, biographical and personality factors in candidate reactions

The following sections bring together the findings from Study 1 and Study 2 to explore the role of attitudinal, biographical and personality factors in candidate reactions.

Attitudinal factors

The SLR findings on attitudinal factors were the most consistent, with self-efficacy and motivation found to be positively related to candidate reactions. Other attributes related to self-efficacy such as locus of control, affectivity and core-self evaluations were also found to have positive relationships. The replicability of findings was somewhat limited with a number of selection methods investigated, and there was some evidence for different patterns of results depending on the selection method used (Oostrom et al., 2010). The findings suggest that attitudinal factors do have an impact on reactions to selection methods, with those who have higher levels of attributes such as self-efficacy, motivation, positive affectivity and core self-evaluations more likely to react positively to selection processes. However, the impact of these attributes will likely vary depending on the content of the selection process. Attitudinal factors were not investigated in Study 2 due to practical constraints of the study; further research could therefore include an exploration of attitudinal factors.

Biographical factors

Ethnicity (sometimes investigated as 'race') and gender were the two most commonly studied areas regarding biographical characteristics. The findings for ethnicity differences were inconsistent, with two studies in the SLR finding ethnic minority candidates had more negative perceptions (one of which only found differences when individuals were highly identified with their ethnic group), two finding that perceptions were more positive for some ethnic minority groups and one study finding no differences. Study 2 found some small but significant differences in reactions, with ethnic minority participants having more positive reactions to some of the subscales measuring perceptions of having a chance to perform and

the perceived predictive validity of some elements of the selection process, but lower reactions to subscales regarding how much information was known about the process.

Some effects of gender were found, with one study (Bernerth, 2005) finding that perceptions interacted with outcomes leading to rejected females having lower procedural and distributive justice ratings than males. Other studies did not find any significant impact of gender on reactions. In Study 2, some gender differences were found in a small number of the candidate reactions scales, however these were small in magnitude. Where differences were found, they tended to be regarding the levels of understanding about the contribution each activity made to the process and regarding the information known about each activity, with males rating these factors more positively than females.

Age, educational level and employment status were investigated in the SLR (Macan et al., 1994; Visser & Schaap, 2017) and no significant differences were found. In line with this, Study 2 did not find any significant differences in candidate reactions on the basis of age or socio-economic background.

Taken collectively, these findings suggest that there is some evidence for the impact of biographical characteristics on candidate reactions, specifically the impact of gender and ethnicity. Of note is the finding that differences were limited to specific subscales, and that these tended to be regarding the communications around the process (i.e., how much candidates understood about the activities and how they would contribute to the selection outcomes). This warrants further research to investigate whether this finding can be replicated in other samples and with other selection processes.

Personality factors

The studies investigating personality differences found a number of relationships with candidate reactions, both direct and indirect. To understand the findings across Study 1 and Study 2, it is beneficial to reconsider the hypothesised relationships, the theoretical basis for these hypotheses, and the proposed explanations of the findings. The range of selection methods considered can make synthesis of study findings challenging; to help identify common themes across the studies the selection methods were grouped as either 'high-touch'

or 'low-touch' (e.g., Hart, 2008; Schwencke, 2022). High-touch selection methods refer to methods where there is significant input from both assessors and candidates, usually with a direct interaction between the two. The methods classed as high-touch in the SLR include interviews (including group interviews and asynchronous interviews) and assessment centres, those classed as low-touch are cognitive ability tests, personality tests, SJTs and written tests. The selection processes used in Study 2 are considered low-touch.

Agreeableness

Individuals higher on the trait of Agreeableness are typically described as being more trusting, co-operative, forgiving and tolerant (Barrick & Mount, 1991; Costa & McCrae, 1992). Merkulova et al. (2014) hypothesised that Agreeableness would be positively related to perceptions of an assessment centre due to the tendency of those high on Agreeableness to be compassionate and sympathetic. They suggest that these positive feelings may be applied to the assessors of the assessment centre. Merkulova et al. also suggested that less agreeable individuals may experience more strain and perceive less control, which could impact on their fairness perceptions. These hypotheses were supported, with their study finding that Agreeableness was positively related to face validity, measurement quality, controllability and quality of administration. Regression analysis also showed that Agreeableness was a significant predictor of measurement quality and controllability. Merkulova et al.'s (2014) findings were supported by Bye and Sandal (2016); they found that Agreeableness was significantly related to the structural fairness factor of Gilliland's model when investigating group interviews. This suggests that individuals who are higher on Agreeableness will react more positively to a number of procedural justice factors in high-touch selection processes, likely due to their tendency to be more positive in general in outlook and particularly feel positively towards others with whom they interact during the selection process.

Hiemstra et al. (2019) investigated reactions to asynchronous interviews in two studies, the first with paid participants and the second with real job applicants. They did not find any significant correlations with face validity, predictive validity or opportunity to perform for Agreeableness in their first study, and in their second study a negative relationship was found with face validity. While still considered a high-touch selection method due to the level of input required from candidates and assessors, it may be that the lack of direct interaction with

assessors meant Agreeableness had a limited role in reactions. The results of the second study also needs to be treated with caution due to the very small sample size.

With regard to the relationship between Agreeableness and low-touch selection methods, Truxillo et al. (2009) suggested that agreeable individuals would rate social fairness more highly when investigating reactions to a written test. They suggested that those high on Agreeableness were more likely to hold positive beliefs about the organisation and feel that they had been treated fairly. A positive correlation was found for the social factor but not the structure factor. Bernerth et al. (2006) hypothesised that Agreeableness would be related to perceived fairness of a personality test, suggesting that those low on Agreeableness will have higher levels of emotional arousal which will influence their perceptions of the selection process. It was found that Agreeableness had a positive relationship with both procedural justice and distributive justice, and that this was found regardless of selection decision. Oostrom et al. (2010) investigated justice perceptions of SJTs and cognitive ability tests, hypothesising that Agreeableness would be related to fairness perceptions based on the tendency for those low in Agreeableness to react more negatively to selection instruments in general. They found a positive relationship with Agreeableness for the face validity and perceived predictive validity of their cognitive ability test but not the SJT. In Study 2 there was support for the hypothesised relationships between Agreeableness and candidate reactions; significant positive correlations were found for all hypothesised relationships with the exception of the perceived predictive validity of the total process.

To summarise, these studies suggest that there is a link between Agreeableness and low-touch selection methods. The social factor of Gilliland's model relates to the level of information shared, including the opportunity for feedback and two-way communication between the organisation and the participant. The findings suggest that being adaptable, co-operative and positive means individuals will be more likely to perceive that they have been provided with enough information, that they will receive an acceptable level of feedback and have had sufficient communication with the organisation. Interestingly, Oostrom et al. (2010) found some differences in reactions to the different selection methods used. Their study variables of face validity and predictive validity relate more to the structure factor of Gilliland's model; their findings may suggest that the SJT was more oblique, and that candidates had less of an understanding of what this was testing, meaning that Agreeableness had less of an impact.

These findings provide promising evidence that levels of trait Agreeableness have an impact on candidate reactions, and in particular impact on the social factor of reactions in both hightouch and low-touch selection processes.

Extraversion

Extraversion describes individuals who are energetic, sociable, gregarious and ambitious (Barrick & Mount, 1991; Costa & McCrae, 1992). Merkulova et al. (2014) hypothesised that Extraversion would be related to applicant perceptions of an assessment centre, suggesting that tasks requiring social interaction are more likely to be seen favourably by individuals with this trait. However, their findings did not support this hypothesis as they did not find any significant direct relationships between Extraversion and candidate reaction in the scales they measured. Hiemstra et al. (2019) hypothesised that Extraversion would relate to perceptions of fairness in asynchronous interviews, as previous research has shown that extraverts tend to present more online, and the socially expressive, attention seeking element of Extraversion would mean they respond more positively to the ability to present themselves in an asynchronous interview. They went on to suggest that extraverts' social and expressive nature means they will be more appreciative of selection instruments with more social cues. This finding was supported in their second study with a real applicant sample. In a similar vein, Brenner et al. (2016) suggested that Extraversion would be related to positive reactions to asynchronous interviews as there are more opportunities for impression management and self -presentation. However, Brenner et al. did not find any relationships with attitudes towards video interviews (investigating the reaction scales of appropriateness, chance to perform, fairness, and informativeness). These findings suggest limited support for the notion that the attention seeking, outgoing nature of extraverts will mean that they react more positively to high-touch selection instruments. It may be that the selection situation means that those higher on Extraversion are not able to achieve the level of social interaction they desire with others and therefore Extraversion plays less of a role than it would in other social situations.

With regard to low-touch selection methods, Truxillo et al. (2009) suggested that individuals higher on Extraversion will be more positive about the social fairness of a written test as they are more likely to be optimistic and perceive that they were treated fairly. No correlations were found with either the social or structure factors. Oostrom et al. (2010) did not make any hypothesis regarding Extraversion and they did not find any significant relationships with face or predictive validity for an SJT or a cognitive ability test. Based on these findings, no

hypotheses were tested focusing on the direct relationship of Extraversion and candidate reactions in Study 2 (interaction effects were tested, detailed in the Interaction Effects section). However, it is possible to examine the correlational relationship *post hoc*. In line with previous findings, there was not a significant relationship between Extraversion and the social factor of reactions, however small but significant relationships were found with the total reactions to the process, the structure factor, and a number of the subscales. Significant relationships were also found with a number of the exercise-specific reactions scales. In their studies looking at interaction effects (discussed in greater detail in the Interaction effects section); Bye and Sandal (2016) found that Extraversion did relate to fairness perceptions, but only at specific levels of Neuroticism and Agreeableness. This suggests that Extraversion may play an indirect role in candidate reactions and is perhaps more affected by the specific selection process content.

Conscientiousness

Conscientiousness is associated with being disciplined, organised, planful and dependable (Barrick & Mount, 1991; Costa & McCrae, 1992). While Bye and Sandal (2016) did not hypothesise a relationship between Conscientiousness and candidate reactions, they did investigate it for exploratory purposes. A significant correlation was found with social fairness, however there was no further discussion of this finding and Conscientiousness was not a significant predictor in their regression equation. Merkulova et al. (2014) hypothesised that Conscientiousness would be positively related to reactions to an assessment centre. Their study was carried out with Armed Forces applicants who were similar to internal candidates as they had previous experience with the organisation. Merkulova et al. suggested that Conscientiousness would play a role as individuals high on this trait could be assumed to have invested more time and effort into the organisation and therefore react more strongly to perceived injustice. Merkulova et al. (2014) found that Conscientiousness was positively correlated with face validity and perceptions of measurement quality. Hiemstra et al. (2019) found that Conscientiousness was related to opportunity to perform but not to face or predictive validity in their first study, but in their second with real applicants no relationships were found. Brenner et al. (2016) hypothesised that higher Conscientiousness would be associated with more positive reactions to asynchronous interviews due to individuals higher on this trait seeing the use of new technology as a positive factor in supporting their ability to achieve on the job. They found Conscientiousness was positively correlated with attitudes towards video interviews (appropriateness, chance to perform, fairness, and informativeness). While not hypothesised, Truxillo et al. (2009) found a positive correlation between Conscientiousness and the social factor (but not the structure factor) of candidate reactions. Oostrom et al. (2010) did not make any hypothesis regarding Conscientiousness and they did not find any significant relationships with face or predictive validity for an SJT or a cognitive ability test. The findings for the social factor and Conscientiousness were aligned to both Bye and Sandal (2016) and Truxillo et al. (2009) suggesting that Conscientiousness has an impact on reactions to the social aspect of procedural justice across both high- and low-touch selection settings. Conscientiousness was not investigated in Bernerth et al.'s (2006) study so it is not possible to compare to their results for low-touch methods, however the findings in Study 2 offer further support. In Study 2, Conscientiousness was found to be a significant predictor of total candidate reactions and the social factor of reaction for the process overall and for the individual exercises.

There is promising evidence for the role Conscientiousness plays in candidate reactions, with studies showing there is an impact in both high- and low-touch selection settings across a range of selection methods. Theoretical discussion around the reasons for the impact of Conscientiousness are limited, however it is suggested that the dutiful and organised nature of those high in Conscientiousness drives how they interact with the selection process. These traits could mean that high Conscientiousness individuals ensure they have time to engage fully with the selection process such as by reading all communications thoroughly and asking questions about the process. Where organisations have provided detailed information about the process, individuals high on Conscientiousness may be more likely to engage with this and therefore see the process as fairer. Future research could test this theory either through an experimental design or by comparing selection processes with varying levels of candidate information provided.

Neuroticism

Neuroticism has been associated with a higher propensity to experience negative emotions such as anxiety, anger, embarrassment, and insecurity (Barrick & Mount, 1991; Costa & McCrae, 1992). Merkulova et al. (2014) hypothesised that Emotional Stability (as the opposite of Neuroticism) would be positively related to assessment centre perceptions, theorising that those who are emotionally calm and stable will experience less negative psychological effects (such as feeling out of control and under strain in the selection process)

and therefore have more positive perceptions. Merkulova et al. found that Emotional Stability was positively correlated with measures of controllability, absence of strain and quality of administration, but not face validity. In their regression, Emotional Stability was a significant predictor of controllability. Hiemstra et al. (2019) found Neuroticism was a significant negative predictor of opportunity to perform but not of face or predictive validity in their first study, however no significant relationships were found in their study with real applicants. Brenner et al. (2016) suggest that individuals higher on Neuroticism would react more negatively to asynchronous interviews due to the stress of the unfamiliar environment and potential privacy concerns over how their data was being used, however they found no relationship with attitudes towards video interviews (measuring perceived appropriateness, chance to perform, fairness, and informativeness).

When investigating reactions to low-touch selection methods, Truxillo et al. (2009) suggested that Neuroticism would be negatively related to the social factors of the written test and found support for this in their study. Truxillo et al. suggest that the likelihood of individuals high on Neuroticism to experience events more negatively than others and cope less well with stress may be heightened in the selection situation. Bernerth et al. (2006) suggest that Neuroticism will be more negatively related to the selection process due to the likelihood of individuals higher on the trait being more fearful of novel situations, more insecure, and experiencing feelings of helplessness. Bernerth et al. found a negative association of Neuroticism with distributive justice but not procedural justice. Oostrom et al. (2010) hypothesised that Emotional Stability would be related to fairness perceptions based on the tendency for those low in Emotional Stability to project their negative emotions on to the perceived job relatedness of the selection instrument. They found a positive relationship of Emotional Stability and face validity and predictive validity for the cognitive ability test but not the SJT. Study 2 supported the findings that higher Neuroticism is related to more negative candidate reactions. The study found significant negative relationships between Neuroticism and many of the procedural justice scales for reactions to the process overall and the individual exercises.

These findings suggest that Neuroticism does play a role in candidate reactions, however this is not uniform across all selection methods or all aspects of reactions. In particular Neuroticism appears to have a stronger relationship to perceptions of the social factors of the selection process than the structural factors. Further research could explore this trait in more

detail to identify where Neuroticism has the biggest impact and how this could be mitigated for candidates higher on this trait.

Openness to experience

Openness to Experience describes individuals who are imaginative, unconventional, curious, and intellectual (Barrick & Mount, 1991; Costa & McCrae, 1992). Bye and Sandal (2016) suggest that Openness to Experience will predict fairness perceptions as the selection process will often involve intellectually stimulating and novel tasks which those high on this trait are more likely to find favourable, and Merkulova et al. (2014) expected that the trait would be related to favourable assessment centre perceptions as the imaginative and curious elements of the trait were likely to help people cope with different testing situations. However, these hypotheses were not supported in either study. Bye and Sandal suggested that the effect many be more likely to be observed when the selection methods are more cognitively loaded, such as selection processes using cognitive ability test. Hiemstra et al. (2019) found Openness to Experience had a significant relationship with opportunity to perform but not with face validity or perceived predictive validity in their first study, however this was not supported in their second study. Brenner et al. (2016) proposed that individuals high on Openness to Experience would have more positive perceptions of asynchronous interviews as they will be more adaptable to new technologies. Brenner et al. did not find a relationship between Openness to Experience and attitudes towards the interviews (measured using scales of perceived appropriateness, chance to perform, fairness, and informativeness), however an interaction effect between Openness to Experience and perceived usefulness of the asynchronous interview was found.

Truxillo et al. (2009) suggested that those higher on Openness to Experience would be more positive about the social fairness of a written test as it may affect how they approach novel testing situations. They found a positive correlation with the social factor but not the structure factor. Van Vianen et al. (2004) focussed specifically on Openness to Experience when investigating reactions to cognitive ability tests, personality tests and SJTs. This study did not find any significant direct relationships with perceptions of fairness, however an indirect effect through belief in tests was found. Bernerth et al. (2006) suggested that those high on Openness to Experience will rate the fairness of a personality test more favourably as they are more likely to view the process as a new and novel experience, while those low on Openness to Experience are more likely to fear the unknown and therefore react more negatively to the

selection process. They found a positive association of Openness to Experience with procedural justice. Oostrom et al. (2010) investigated justice perceptions of SJTs and cognitive ability tests, hypothesising that Openness to Experience would be related to fairness perceptions based on the tendency for those low in Openness to Experience to be more resistant to modern computer-based selection instruments. They found significant relationships for the face validity and the predictive validity of both the SJT and the cognitive ability test. In the regression, Openness to Experience explained the variance in the predictive and face validity of the cognitive test beyond the control variables, and explained variance in the face validity of the SJT. Oostrom et al. suggest that individuals who are more amenable to new experiences react more positively to computer-based selection instruments. In Study 2, the only significant relationship between Openness to Experience and candidate reactions was the perceived predictive validity of the candidate statement. This may have been due to the majority of participants being familiar with the types of tests used, therefore they did not seem particularly 'novel', thus limiting the impact of levels of Openness to Experience.

Across the studies, the support for the relationship between Openness to Experience and candidate reactions is quite weak, particularly with high-touch selection methods. Further research into the impact of this trait on other technology-enabled selection methods (such as virtual assessment centres and gamified assessments) would be beneficial to better understand the relationship.

Interaction effects

In their study investigating reactions to selection interviews, Bye and Sandal (2016) focused on the interactions between Neuroticism and Extraversion (described as the 'affective plane') and the interaction between Extraversion and Agreeableness (the 'interpersonal plane'). They suggested that higher levels of Emotional Stability will mean individuals are more relaxed and comfortable with a group interview setting regardless of their levels of Extraversion, and that Agreeableness and Extraversion interact to shape levels of trust and comfort in the interactions between the interviewer and other candidates These were of particular interest when investigating reactions to group interviews, where there is a need to interact with others as well as being evaluated by the interviewers. Bye and Sandal found that those higher on Neuroticism and lower on Extraversion perceived the interview as less fair than those higher on Neuroticism but high on Extraversion (in terms of both social and structural fairness). Those who were higher on Agreeableness and higher on Extraversion had favourable reactions to the social justice of the interview (but not structural). Hiemstra et al. (2019) also tested the interaction of Extraversion and Neuroticism on reactions to asynchronous interviews, however this interaction effect was not supported in their study.

Honkaniemi et al., (2013) examined differences in personality types on applicant reactions to cognitive ability tests, with the types characterised by different combinations of the Big Five traits. They identified four types: Resilient (characterised by low Neuroticism, high Conscientiousness, higher Extraversion and Agreeableness), Overcontrolled (characterised by high Neuroticism, low Extraversion and Agreeableness), Undercontrolled (characterised by high Neuroticism, higher Extroversion, low Conscientiousness) and Bohemian (characterised by low Extraversion and low Conscientiousness). Honkaniemi did not find an effect of personality type on face validity or predictive validity, however in terms of general fairness perceptions, Overcontrolled types had the least favourable perceptions and Bohemian had the highest. Interaction effects were unsupported in Study 2; this suggests that interaction effects are particularly sensitive to the selection method used, potentially with them playing a greater role when the selection methods required a high degree of interpersonal interactions.

Practical & Theoretical Implications

Theoretical implications

This thesis has increased understanding of the impact of individual differences on candidate reactions in a number of ways. By taking a holistic view across a spectrum of individual differences (attitudinal, biographical and the full Big Five model of personality) a more complete understanding of how these factors affect candidate reactions is achieved. A comprehensive range of selection methods has been examined within this thesis, ranging from more traditional methods such as face-to-face interviews and cognitive ability tests, to newer methods such as SJTs and asynchronous video interviews. Study 1 highlighted the lack of replicability of studies in the area of candidate reactions and the relatively limited number of studies which have focused on individual differences. Study 2 sought to address some of the gaps in the literature by focussing on personality differences and investigating reactions of senior, internal candidates. By investigating reactions to a number of different selection

exercises within a single process, Study 2 also expanded the evidence base for the favourability of different selection methods within the same sample. This builds on the work of previous research (e.g., Anderson et al., 2010; Anderson and Witvliet, 2008; Hausknecht et al., 2004) to understand the favourability of different selection methods in a real-world setting.

In addition to expanding knowledge regarding the impact of individual differences on candidate reactions, this thesis has also begun to address the calls for research into newer, technology-enabled selection processes. As described in Woods et al. (2020) much of what is understood about selection processes has been drawn from research into more traditional selection methods and research has not kept pace with technological developments in the field. Several previous studies have investigated reactions to selection instruments; however, these have not included common tools such as written applications or SJTs. The inclusion of newer and less well-researched tools broadens researchers' understanding of the current selection landscape.

Practical implications

Many organisations are aiming to increase the diversity of their employees (Gartner, 2022) and diversity in senior leaders can have direct positive impacts on organisational performance (Hunt et al., 2020). It is important to understand whether aspects of the selection process are causing more negative perceptions with target groups as this could reduce the diversity of applicant pools. While the impact of gender and ethnicity on reactions is likely to be small, small effects can accumulate over time and organisations could see a measurable impact if there are group differences in reactions to their selection processes.

The pattern of results described above suggests there are some differences between how people respond to the social factors of a process (such the interactions with the organisation and the level and quality of communications about the process) and the structural factors (regarding the content of the process). Organisations may find that issues regarding the social factors are easier and cheaper to address than changing the content of their selection processes. For example, Truxillo et al. (2009) found that providing information had a positive effect on reactions to the selection process, and Study 2 found that there may be gender and ethnicity differences in reactions to the perceptions regarding the level of information known

about a process. Testing messaging with a diverse group of individuals to ensure messages about the process are understood and are helpful could be a relatively straightforward way for practitioners to improve reactions to their selection processes.

The findings suggest that the impact of personality factors will vary depending on the selection process used. This variation means that the most inclusive selection processes should include a range of activities, preferably with each stage of the process employing more than one selection method (rather than a series of 'hurdles' where each activity needs to be passed in turn before moving to the next, Mendoza et al., 2004). This means that aspects of the selection process that are off-putting to individuals with certain traits or attributes can be counterbalanced with elements that they find more favourable. This could result in fewer applicants self-selecting out of the process and reduce the risk of negative feedback being shared.

The current economic climate is uncertain and competition for talented employees is high (De Smet et al., 2022). In such uncertain markets, organisations must do everything they can to keep candidates engaged during what can be lengthy selection processes; this is particularly key when it comes to retaining talent already with the organisation (Giumetti & Sinar, 2012). As has been previously explored, the implications of poorly received selection processes can be far reaching; negative perceptions can impact on turnover intentions, organisational commitment, job satisfaction (Ambrose & Cropanzano, 2003), reduce employee engagement (Ford et al., 2009), and result in the loss of talent from the organisation (Giumetti & Sinar, 2012). Konradt et al. (2017) identified some of the longer-term consequences of selection processes, with fairness perceptions being related to acceptance of the job offer, and later job performance. Understanding more about how applicants may react to different selection methods can help with the design of a selection process that will keep candidates engaged throughout and reduce the risk of it being poorly received.

Research implications

While other models have been developed, Gilliland's (1993) procedural justice model has been the driver for many studies of application reactions (McCarthy et al., 2017). This thesis highlights the continued relevance of this model as a framework for understanding candidate reactions. Further research could build on the work of Bauer et al. (2001) in developing a measure of Gilliland's model (the SPJS) to build a tool that is grounded in reactions to technology-enabled selection methods.

One related area of applicant reactions has been to investigate the impact of providing more or less information about the selection process to candidates. Gilliland's model would suggest that more information is desirable, however there can be downsides to providing such information. Langer et al. (2018) investigated the impact of providing procedural information to explain and justify the use of novel technologies in selection. They found that while there was a positive relationship between providing information and applicant reactions, and an indirect positive affect on organisation attractiveness through applicant reactions, there was also a direct negative relationship between information known and organisational attractiveness. The authors suggest that applicants might be intimidated by the technological aspects of the process or may have had enough information to question the process but not enough to explain it fully. Further research could explore how messaging is used in technology-enabled selection processes.

As described in the preceding sections, the role of personality in candidate reactions to selection processes is complex and future research could focus on the role of each of the Big Five factors of personality individually as well as expanding on research across the whole model. For example, the role of Extraversion on newer selection methods such as virtual assessment centres could be investigated, or specific elements of Extraversion could be explored (e.g., 'Activity' or 'Warmth' from the NEO; Costa & McCrae, 1992) as there may be differences in the role individual facets play which are not seen at the overall factor level. Interestingly, Conscientiousness appears to play more of a role in reactions to the social factors of candidate reactions, therefore it would be beneficial for further research to test the hypothesis that this is due to candidates spending more time engaging with the communications around the selection process. Neuroticism also appears to play more of a role in reactions to the social factors. Given that it is theoretically related to a higher propensity to experience more negative emotions during the selection process, it is important to understand this better to reduce the negative impact on candidates. Further research could seek to understand in which selection settings Neuroticism plays the greatest role to better understand how this impact could be mitigated.

The impact of Openness to Experience was somewhat surprisingly weak, given the clear theoretical basis for this relationship. It may be that none of the selection methods used in the studies identified were seen as particularly 'novel' to their candidates. Research into tools that are likely to become more common but may be less familiar to candidates (such as gamified assessments and virtual reality assessment centres) would be beneficial to better test the hypotheses regarding the role Openness to Experience may play in candidate reactions. There was more support for the findings regarding Agreeableness and candidate reactions, however there was still some variability on the impact on reactions to different selection process. These findings provide promising evidence that levels of trait Agreeableness have an impact on candidate reactions, particularly on the social factor of reactions in both high- and low-touch selection processes. Further research could aim to replicate these findings with different samples and selection methods.

Only two of the studies (Merkulova et al., 2014 and Study 2) investigated internal candidates as their study population, and these did not compare internal candidates' reactions to external candidates. Other research has found differences in the reactions of internal and external candidates, for example Sylva and Mol (2009) investigated online application processes for both internal and external applicants. They found that while perceptions were generally positive, internal applicants had less positive reactions than external candidates. The authors suggest that internal candidates may expect a more personalised and informal process, given that they are known to the organisation. Further research could investigate differences between reactions to the communication elements of the selection process, as it may be that internal candidates require a more tailored communication strategy. Lastly, none of the studies in the SLR considered disability as a factor that could impact candidate reactions, however in Study 2 candidates applying under the guaranteed interview offer (the majority of whom had a disability) were found to react more negatively to the selection processes in candidates with a disability.

Limitations

The studies in this thesis all employed quantitative methods which are particularly sensitive to sample sizes. Where studies in the SLR provided justification for their sample size, they tended to use an expected medium effect size to estimate minimum samples required. This is likely due to a lack of meta-analysis in the area to systematically identify expected effect sizes. Many of the studies identified in the SLR found small or small-to-medium effect sizes, with small effect sizes being non-significant in their given samples. It is likely that many of the studies were underpowered, therefore future research should aim to investigate much larger samples. Further, the application of quantitative methodologies only means that the studies tended to be reductionist in nature and there was no opportunity to explore additional factors driving reactions beyond those specified in the study designs. The inclusion of some mixed methods research could help to both describe and explain the observed outcomes (Sousa, 2010).

Selection processes are likely to be unique to each organisation, meaning that the generalisability of findings across this thesis could be limited. As noted, there is very little replication of studies, therefore caution must be taken in generalising results to different selection situations and different samples. However, by presenting a synthesis of findings from both Study 1 and Study 2, this thesis provides some promising evidence on which further research can build.

Some limitations of the individual studies have been discussed in earlier chapters. For example, with regard to the SLR, the date restriction could have resulted in relevant papers published prior to 1993 being missed. However, given the changes in the selection landscape over the past 30 years it is unlikely that papers published even further back would have significant practical implications for current researchers and practitioners. In addition, the search terms and selection of databases could also have resulted in relevant studies being excluded from the review. The careful selection of databases and search terms as described in Chapter 3 means that it is unlikely that a large number of papers were missed and restrictions in the search strategy were mitigated through additional hand searching.
Study 2 took place in a single organisation which, as noted, could impact on the generalisability of findings. However, the applicants themselves were heterogeneous, as they came from a broad range of Civil Service departments and had a wide variety of professional backgrounds. It could also be noted that the applicants were applying for a talent scheme, rather than for a new job role which could limit how much the findings will generalise to a job selection situation. The talent development schemes are highly prestigious within the UK Civil Service and form a key career development step for scheme participants. It is therefore highly likely that the participants in Study 2 were as motivated to succeed in the talent scheme selection process as they would be for any other work-related selection process.

Conclusions

With the continuing War for Talent and the Great Resignation, candidate experiences are ever more important. This thesis responds to calls for research into candidate reactions by providing a synthesis of existing literature; and by investigating the reactions of a candidate group previously under-researched in the field. This literature review is the first known systematic review to bring together findings regarding the impact of a range of individual differences on candidate reactions. The field study provided a unique opportunity to investigate reactions of senior applicants internal to the organisation undertaking a multimethod, technology-enabled selection process.

There is some promising evidence to suggest that personality, attitudes, and biographical factors such as gender and ethnicity do play a role in candidate reactions, however further research is needed to better understand in which selection situations these factors have the most impact. Practitioners would therefore benefit from ensuring their selection processes comprise a range of activities and employ multiple methods in each stage to give the best chance of receiving a positive reception from their candidates.

References

- Alharahsheh, H. H., & Pius, A. (2020). A review of key paradigms: Positivism VS interpretivism. Global Academic *Journal of Humanities and Social Sciences*, 2(3), 39-43. 10.36348/gajhss.2020.v02i03.001
- Armstrong, R. A. (2014). When to use the Bonferroni correction. Ophthalmic and Physiological Optics, 34(5), 502-508. https://doi.org/10.1111/opo.12131
- Ambrose, M. L., & Cropanzano, R. (2003). A longitudinal analysis of organizational fairness: An examination of reactions to tenure and promotion decisions. *Journal of Applied Psychology*, 88(2), 266-275. 10.1037/0021-9010.88.2.266
- Anderson, N., & Goltsi, V. (2006). Negative psychological effects of selection methods: Construct formulation and an empirical investigation into an assessment center. *International Journal* of Selection and Assessment, 14(3), 236-255
- Anderson, N., Salgado, J. F., & Hülsheger, U. R. (2010). Applicant reactions in selection: Comprehensive meta-analysis into reaction generalization versus situational specificity. *International Journal of Selection and Assessment, 18*(3), 291-304. 10.1111/j.1468-2389.2010.00512.x
- Anderson, N., & Witvliet, C. (2008). Fairness reactions to personnel selection methods: An international comparison between the Netherlands, the United States, France, Spain, Portugal, and Singapore. *International Journal of Selection and Assessment, 16*(1), 1-13.
- Arvey, R. D., Strickland, W., Drauden, G., & Martin, C. (1990). Motivational components of test taking. *Personnel Psychology*, 43(4), 695-716
- Ball, G. A., Trevino, L. K., & Sims, H. P. (1993). Justice and organizational punishment: Attitudinal outcomes of disciplinary events. Social Justice Research, 6(1), 39-67.

- Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press.
- Bangerter, A., König, C. J., Blatti, S., & Salvisberg, A. (2009). How widespread is graphology in personnel selection practice? A case study of a job market myth. *International Journal of Selection and Assessment*, 17(2), 219-230. 10.1111/j.1468-2389.2009.00464.x
- Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, 44(1), 1-26. 10.1111/j.1744-6570.1991.tb00688.x
- Bauer, T. N., Truxillo, D. M., Sanchez, R. J., Craig, J. M., Ferrara, P., & Campion, M. A. (2001).
 Applicant reactions to selection: Development of the selection procedural justice scale (SPJS). *Personnel Psychology*, 54(2), 387-419. 10.1111/j.1744-6570.2001.tb00097.x
- Bernerth, J. B. (2005). Perceptions of justice in employment selection decisions: The role of applicant gender. *International Journal of Selection and Assessment*, 13(3), 206-212. 10.1111/j.1468-2389.2005.00316.x
- Bernerth, J. B., Feild, H. S., Giles, W. F., & Cole, M. S. (2006). Perceived fairness in employee selection: The role of applicant personality. *Journal of Business and Psychology*, 20(4), 545-563. 10.1007/s10869-005-9004-3
- Bertolino, M., & Steiner, D. D. (2007). Fairness reactions to selection methods: An Italian study. International Journal of Selection and Assessment, 15(2), 197-205. 10.1111/j.1468-2389.2007.00381.x
- 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), 12-20. 10.25035/pad.2016.002
- Blunt. (2022, August 5). *How to write a supporting statement for a job application*. The Job Studio https://thejobstudio.co.uk/application-advice/how-to-write-a-supporting-statement-for-a-job-application/

- Bosco, F. A., Aguinis, H., Singh, K., Field, J. G., & Pierce, C. A. (2015). Correlational effect size benchmarks. *Journal of Applied Psychology*, 100(2), 431-449. 10.1037/a0038047
- BPS. (2021). Code of Ethics and Conduct. BPS https://www.bps.org.uk/sites/www.bps.org.uk/files/Policy/Policy%20-%20Files/BPS%20Code%20of%20Ethics%20and%20Conduct.pdf
- Branine, M. (2008). Graduate recruitment and selection in the UK: A study of the recent changes in methods and expectations. *Career Development International*. 13(6), 497-513. 10.1108/13620430810901660
- 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*, 1-11. 10.3389/fpsyg.2016.00863
- Briner, R. B., & Denyer, D. (2012). Systematic review and evidence synthesis as a practice and scholarship tool. In D. M. Rousseau (Ed.), *The Oxford Handbook of Evidence-Based Management*, (pp 112-129). Oxford University Press
- Briner, R. B., & Rousseau, D. M. (2011). Evidence-based I–O psychology: Not there yet but now a little nearer? *Industrial and Organizational Psychology*, 4(1), 76-82. 10.1111/j.1754-9434.2010.01301.x
- Butucescu, A., Iliescu, D., & Opariuc, C. D. (2019). The generalizability of reactions to assessment: An application of the Selection Procedural Justice Scale (SPJS) in academic settings. *International Journal of Selection and Assessment*, 27(4), 406-410. 10.1111/ijsa.12262
- Bye, H. H., & Sandal, G. M. (2016). Applicant personality and procedural justice perceptions of group selection interviews. *Journal of Business and Psychology*, *31(4)*, 569-582. 10.1007/s10869-015-9430-9
- Chambers, E. G., Foulon, M., Handfield-Jones, H., Hankin, S. M., & Michaels III, E. G. (1998). The war for talent. *The McKinsey Quarterly*, (3), 44-57.

- Chan, D., & Schmitt, N. (2004). An agenda for future research on applicant reactions to selection procedures: A construct-oriented approach. *International Journal of Selection and Assessment*, 12(1-2), 9-23. 10.1111/j.0965-075X.2004.00260.x
- Chan, D., Schmitt, N., DeShon, R. P., Clause, C. S., & Delbridge, K. (1997). Reactions to cognitive ability tests: the relationships between race, test performance, face validity perceptions, and test-taking motivation. *Journal of Applied Psychology*, 82(2), 300. 10.1037/0021-9010.82.2.300
- Chan, D., Schmitt, N., Sacco, J. M., & DeShon, R. P. (1998). Understanding pretest and posttest reactions to cognitive ability and personality tests. *Journal of Applied Psychology*, 83(3), 471-485. 10.1037/0021-9010.83.3.471
- Chapman, D. S., Uggerslev, K. L., & Webster, J. (2003). Applicant reactions to face-to-face and technology-mediated interviews: A field investigation. *Journal of Applied Psychology*, 88(5), 944–953. https://doi.org/10.1037/0021-9010.88.5.944
- Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling*, *9*(2), 233-255.
- Church, A. H. (2001). Is there a method to our madness? The impact of data collection methodology on organizational survey results. *Personnel Psychology*, 54(4), 937-969. 10.1111/j.1744-6570.2001.tb00238.x
- CIPD. (2021). *Recruitment: An Introduction*. CIPD. https://www.cipd.co.uk/knowledge/fundamentals/people/recruitment/factsheet#gref
- Cohen, A. (2021, May 10). *How to Quit Your Job in the Great Post-Pandemic Resignation Boom*. Bloomberg. <u>https://www.bloomberg.com/news/articles/2021-05-10/quit-your-job-how-to-resign-after-covid-pandemic</u>

- Collins, C. J., & Han, J. (2004). Exploring applicant pool quantity and quality: The effects of early recruitment practice strategies, corporate advertising, and firm reputation. *Personnel psychology*, 57(3), 685-717. 10.1111/j.1744-6570.2004.00004.x
- Costa, P., & McCrae, R. (1992). NEO PI-R Professional Manual. *Psychological Assessment Resources, 396*.
- Costa, P. T., Jr., & McCrae, R. R. (1985). *The NEO Personality Inventory Manual*. Odessa, FL: Psychological Assessment Resources
- Costa, P. T., Jr., & McCrae, R. R. (2008). The Revised NEO Personality Inventory (NEO-PI-R). In G. J. Boyle, G. Matthews, & D. H. Saklofske (Eds.), *The SAGE Handbook of Personality Theory and Assessment, Vol. 2. Personality Measurement and Testing* (pp. 179–198). Sage Publications, Inc.
- Creswell, J. W. (2014). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches. Sage.
- Dahl, A. (2017, March 8). Virtual Reality Coming to Your Assessment Centre Soon? University of Bath. https://blogs.bath.ac.uk/careers/2017/03/08/virtual-reality-coming-to-yourassessment-centre-soon/
- Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics* and Methodology, 8(3), 513-539. 10.1093/jssam/smz008
- Denyer, D., & Tranfield, D. (2009). Producing a systematic review. In D. A. Buchanan & A.
 Bryman (Eds.), *The Sage handbook of organizational research methods* (pp. 671–689).
 Sage Publications Ltd.
- Denyer, D., Tranfield, D., & van Aken, J. E. (2008). Developing Design Propositions through Research Synthesis. Organization Studies, 29(3), 393–413. <u>https://doi.org/10.1177/0170840607088020</u>

- Derous, E., Bord, M., & De Witte, K. (2004). How applicants want and expect to be treated:
 Applicants' selection treatment beliefs and the development of the Social Process
 Questionnaire on Selection. *International Journal of Selection and Assessment, 12 (1/2), 99-119.*
- De Smet, A., Dowline, B., Mugayar-Baldocchi, M., & Schaninger, B. (2022). Gone for Now, or Gone for Good? How to Play the New Talent Game and Win Back Workers. McKinsey Quarterly, 1-9.
- Donaldson-Feilder, E., Lewis, R., & Yarker, J. (2019). What outcomes have mindfulness and meditation interventions for managers and leaders achieved? A systematic review.
 European Journal of Work and Organizational Psychology, 28(1), 11-29.
 10.1080/1359432X.2018.1542379
- EFPA. (2008). EFPA Review Model for the Description and Evaluation of Psychological and Educational Tests. EFPA.
- Elkins, T. J., & Phillips, J. S. (2000). Job context, selection decision outcome, and the perceived fairness of selection tests: biodata as an illustrative case. *Journal of Applied Psychology*, 85(3), 479-484. 10.1037/0021-9010.85.3.479
- Eploy. (2021). 5th Annual UK Candidate Attraction Report 2020-2021. Eploy. https://www.eploy.co.uk/resources/whitepaper/uk-candidate-attraction-report-2021-22/
- Evertz, L., & Süß, S. (2017). The importance of individual differences for applicant attraction: a literature review and avenues for future research. *Management Review Quarterly*, 67(3), 141-174. 10.1007/s11301-017-0126-2
- Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39, 175-191. 10.3758/BF03193146
- Field, A. (2009). Discovering Statistics using SPSS: Book plus code for E version of text (Vol. 896).London, UK: SAGE Publications Limited.

- Fletcher, C. (1991). Candidates' reactions to assessment centres and their outcomes: A longitudinal study. *Journal of Occupational Psychology*, *62* (2),117-127. 10.1111/j.2044-8325.1991.tb00547.x
- Ford, D. K., Truxillo, D. M., & Bauer, T. N. (2009). Rejected but still there: Shifting the focus in applicant reactions to the promotional context. *International Journal of Selection and Assessment*, 17(4), 402-416. 10.1111/j.1468-2389.2009.00482.x
- Fosnacht, K., Sarraf, S., Howe, E., & Peck, L. K. (2017). How important are high response rates for college surveys? *The Review of Higher Education*, 40(2), 245-265. 10.1353/rhe.2017.0003
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. Advances in Methods and Practices in Psychological Science, 2(2), 156-168.
- Gartner. (2022, April 27). CEOs Turn a Sharp Eye to Workforce Issues and Sustainability in 2022-23. Gartner https://www.gartner.com/en/articles/ceos-turn-a-sharp-eye-to-workforce-issuesand-sustainability-in-2022-23
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74-78. 10.1016/j.paid.2016.06.069
- Gilliland, S. W. (1993). The perceived fairness of selection systems: An organizational justice perspective. Academy of Management Review, 18(4), 694-734. 10.5465/AMR.1993.9402210155
- Gilliland, S. W. (1994). Effects of procedural and distributive justice on reactions to a selection system. *Journal of Applied Psychology*, *79*(5), 691-701. 10.1037/0021-9010.79.5.691
- Giumetti, G. W., & Sinar, E. F. (2012). Don't you know me well enough yet? Comparing reactions of internal and external candidates to employment testing. *International Journal of Selection and Assessment, 20*(2), 139-148. 10.1111/j.1468-2389.2012.00587.x

- Glickman, M. E., Rao, S. R., & Schultz, M. R. (2014). False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies. *Journal of clinical epidemiology*, 67(8), 850-857. https://doi.org/10.1016/j.jclinepi.2014.03.012
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, *4*(1), 26-42.
- Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of publicdomain personality measures. *Journal of Research in Personality*, 40, 84-96.
- Goldman, B. M. (2001). Toward an understanding of employment discrimination claiming: An integration of organizational justice and social information processing theories. *Personnel Psychology*, 54(2), 361-386. 10.1037/1040-3590.4.1.26
- Hart, K. A. (2008). High-tech, high-touch recruitment: an oxymoron?. *Nursing Economics*, 26(4), 270-279.
- Hauenstein, N. M., Holmes, J. T., & Tison, E. B. (2013). Detecting adverse impact: the Four-fifths rule versus significance testing. Public Personnel Management, 42(3), 403-420. https://doi.org/10.1177/0091026013495762
- Hausknecht, J. P., Day, D. V., & Thomas, S. C. (2004). Applicant reactions to selection procedures: An updated model and meta-analysis. *Personnel Psychology*, 57(3), 639-683. 10.1111/j.1744-6570.2004.00003.x
- Hiemstra, A. M., Derous, E., Serlie, A. W., & Born, M. P. (2012). Fairness perceptions of video resumes among ethnically diverse applicants. *International Journal of Selection and Assessment*, 20(4), 423-433. 10.1111/ijsa.12005
- Hiemstra, A. M., Oostrom, J. K., Derous, E., Serlie, A. W., & Born, M. P. (2019). Applicant perceptions of initial job candidate screening with asynchronous job interviews: Does personality matter?. *Journal of Personnel Psychology*, 18(3), 138-147. 10.1027/1866-5888/a000230

- Hong, Q. N., & Pluye, P. (2019). A conceptual framework for critical appraisal in systematic mixed studies reviews. *Journal of Mixed Methods Research*, 13(4), 446-460.
- Honkaniemi, L., Feldt, T., Metsäpelto, R. L., & Tolvanen, A. (2013). Personality types and applicant reactions in real-life selection. *International Journal of Selection and Assessment*, 21(1), 32-45. 10.1111/ijsa.12015
- Hunt, V., Prince, S., Dixon-Fyle, S., & Dolan, K. (2020). Diversity Wins. McKinsey & Company.
- Iacobucci, D., Schneider, M. J., Popovich, D. L., & Bakamitsos, G. A. (2016). Mean centering helps alleviate "micro" but not "macro" multicollinearity. *Behavior research methods*, 48(4), 1308-1317. <u>https://doi.org/10.3758/s13428-015-0624-x</u>

IBM Corp. (2019). IBM SPSS Statistics for Windows, Version 26.0. Armonk, NY: IBM Corp.

- Institute for Government. (2022, September 14). *Civil Service Staff Numbers*. The Institute for Government https://www.instituteforgovernment.org.uk/explainers/civil-service-staff-numbers
- Jackson, D. J., Dewberry, C., Gallagher, J., & Close, L. (2018). A comparative study of practitioner perceptions of selection methods in the United Kingdom. *Journal of Occupational and Organizational Psychology*, 91(1), 33-56. 10.1111/joop.12187
- Johnson, P., & Cassell, C. (2001). Epistemology and work psychology: New agendas. *Journal of Occupational and Organizational Psychology*, 74(2), 125-143. 10.1348/096317901167280
- Johnson, W. (2020). Individual Differences. Edinburgh University https://www.handbooks.psy.ed.ac.uk/psychology-2/%09psychology-2a/lecturetopics/individual-differences
- Jones, T. L., Baxter, M. A. J., & Khanduja, V. (2013). A quick guide to survey research. The Annals of The Royal College of Surgeons of England, 95(1), 5-7. 10.1308/003588413X13511609956372

- Judge, T. A., & Bono, J. E. (2001). Relationship of core self-evaluations traits—self-esteem, generalized self-efficacy, locus of control, and emotional stability—with job satisfaction and job performance: A meta-analysis. *Journal of Applied Psychology*, 86(1), 80-92. 10.1037/0021-9010.86.1.80
- Kaplan, S., Bradley, J. C., Luchman, J. N., & Haynes, D. (2009). On the role of positive and negative affectivity in job performance: a meta-analytic investigation. *Journal of Applied Psychology*, 94(1), 162-176. 10.1037/a0013115
- Konradt, U., Garbers, Y., Böge, M., Erdogan, B., & Bauer, T. N. (2017). Antecedents and consequences of fairness perceptions in personnel selection: A 3-year longitudinal study. *Group & Organization Management*, 42(1), 113-146.
- Konradt, U., Oldeweme, M., Krys, S., & Otte, K. P. (2020). A meta-analysis of change in applicants' perceptions of fairness. *International Journal of Selection and Assessment*, 28(4), 365-382. 10.1111/ijsa.12305
- Konradt, U., Warszta, T., & Ellwart, T. (2013). Fairness perceptions in web-based selection: Impact on applicants' pursuit intentions, recommendation intentions, and intentions to reapply. *International Journal of Selection and Assessment*, 21(2), 155-169.
 10.1111/ijsa.12026
- LaHuis, D. M. (2005). Individual differences in applicant reactions: A job-search perspective. International Journal of Selection and Assessment, 13(2), 150-159. 10.1111/j.0965-075X.2005.00309.x
- 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. 10.1016/j.chb.2017.11.036
- Leucht, S., Kissling, W., & Davis, J. M. (2009). How to read and understand and use systematic reviews and meta-analyses. *Acta Psychiatrica Scandinavica*, *119*(6), 443-450. 10.1111/j.1600-0447.2009.01388.x

- Liberati, A., Altman, D. G., Tetzlaff, J., Mulrow, C., Gøtzsche, P. C., Ioannidis, J. P., Clarke, M., Devereaux, P. J., Kleijnen, J., & Moher, D. (2009). The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *PLoS Medicine*, 6(7), 1-28. 10.1371/journal.pmed.1000100
- Lievens, F., & Slaughter, J. E. (2016). Employer image and employer branding: What we know and what we need to know. Annual Review of Organizational Psychology and Organizational Behavior, 3, 407-440. 10.1146/annurev-orgpsych-041015-062501
- Macan, T. H., Avedon, M. J., Paese, M., & Smith, D. E. (1994). The effects of applicants' reactions to cognitive ability tests and an assessment center. *Personnel Psychology*, 47(4), 715-738. 10.1111/j.1744-6570.1994.tb01573.x
- McCarthy, J. M., Bauer, T. N., Truxillo, D. M., Anderson, N. R., Costa, A. C., & Ahmed, S. M. (2017). Applicant perspectives during selection: A review addressing "So what?,""What's new?," and "Where to next?". *Journal of Management*, 43(6), 1693-1725. 10.1177/0149206316681846
- McCarthy, J., Hrabluik, C., & Jelley, R. B. (2009). Progression through the ranks: Assessing employee reactions to high-stakes employment testing. *Personnel Psychology*, 62(4), 793-832. https://doi.org/10.1111/j.1744-6570.2009.01158.x
- McHugh, M. L. (2012). Interrater reliability: the kappa statistic. *Biochemia Medica*, 22(3), 276-282. 10.11613/bm.2012.031
- Mendoza, J. L., Bard, D. E., Mumford, M. D., & Ang, S. C. (2004). Criterion-related validity in multiple-hurdle designs: Estimation and bias. *Organizational Research Methods*, 7(4), 418-441. 10.1177/1094428104268752
- Merkulova, N., Melchers, K. G., Kleinmann, M., Annen, H., & Tresch, T. S. (2014). Effects of individual differences on applicant perceptions of an operational assessment center. *International Journal of Selection and Assessment*, 22(4), 355-370. 10.1111/ijsa.12083

- Miles, S. J., & McCamey, R. (2018). The candidate experience: Is it damaging your employer brand?. *Business Horizons*, 61(5), 755-764. 10.1016/j.bushor.2018.05.007
- Nadler, J., Gann-Bociek, M., & Skaggs, B. (2017). Interview support on perceptions of organizational attractiveness: The effects of applicant gender and socio-economic status. *Management Research Review*, 40(7), 783-799. 10.1108/MRR-06-2016-0148
- Nakagawa, S. (2004). A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral ecology*, 15(6), 1044-1045. https://doi.org/10.1093/beheco/arh107
- Nikolaou, I. (2021). What is the Role of Technology in Recruitment and Selection?. *The Spanish Journal of Psychology*, 24(2), 1-6. 10.1017/SJP.2021.6
- Nikolaou, I., & Foti, K. (2018). Personnel selection and personality. *The SAGE Handbook of Personality and Individual Differences, 3*, 458-474.
- Nikolaou, I., & Judge, T. A. (2007). Fairness Reactions to Personnel Selection Techniques in Greece: The role of core self-evaluations. *International Journal of Selection and Assessment*, 15(2), 206-219. 10.1111/j.1468-2389.2007.00382.x
- ONS. (2022, May 17). Vacancies and Iobs in the UK: May 2022. ONS https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/employmentandemplo yeetypes/bulletins/jobsandvacanciesintheuk/may2022
- Oostrom, J. K., Born, M. P., Serlie, A. W., & Van Der Molen, H. T. (2010). Effects of individual differences on the perceived job relatedness of a cognitive ability test and a multimedia situational judgment test. *International Journal of Selection and Assessment, 18*(4), 394-406. 10.1111/j.1468-2389.2010.00521.x
- Oostrom, J. K., Bos-Broekema, L., Serlie, A. W., Born, M. P., & van der Molen, H. T. (2012). A field study of pretest and posttest reactions to a paper-and-pencil and a computerized inbasket exercise. *Human Performance*, 25(2), 95-113. 10.1080/08959285.2012.658928

- Park, Y. S., Konge, L., & Artino, A. R. (2020). The positivism paradigm of research. Academic Medicine, 95(5), 690-694. <u>http://dx.doi.org/10.1097/ACM.000000000003093</u>
- Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *British Medical Journal,* 316(7139), 1236-1238. https://doi.org/10.1136/bmj.316.7139.1236
- Petty-Saphon, K., Walker, K. A., Patterson, F., Ashworth, V., & Edwards, H. (2017). Situational judgment tests reliably measure professional attributes important for clinical practice. *Advances in Medical Education and practice*, 8, 21-23.
- Ployhart, R. E., Ziegert, J. C., & McFarland, L. A. (2003). Understanding racial differences on cognitive ability tests in selection contexts: An integration of stereotype threat and applicant reactions research. *Human Performance*, 16(3), 231-259. 10.1207/S15327043HUP1603_4
- Pontes, H. M., Schivinski, B., Kannen, C., & Montag, C. (2022). The interplay between time spent gaming and disordered gaming: A large-scale world-wide study. *Social Science & Medicine*, 296, 1-9. 10.1016/j.socscimed.2022.114721
- Pontes, H. M., Schivinski, B., Sindermann, C., Li, M., Becker, B., Zhou, M., & Montag, C. (2021). Measurement and conceptualization of Gaming Disorder according to the World Health Organization framework: The development of the Gaming Disorder Test. *International Journal of Mental Health and Addiction*, *19*(2), 508-528. 10.1007/s11469-019-00088-z
- Potočnik, K., Anderson, N. R., Born, M., Kleinmann, M., & Nikolaou, I. (2021). Paving the way for research in recruitment and selection: recent developments, challenges and future opportunities. *European Journal of Work and Organizational Psychology*, 30(2), 159-174. 10.1080/1359432X.2021.1904898
- Puttick, R., & Ludlow, J. (2013). Nesta Standards of Evidence: An Approach That Balances the Need for Evidence with Innovation. Nesta
- Reeder, M. C., Powers, C. L., Ryan, A. M., & Gibby, R. E. (2012). The role of person characteristics in perceptions of the validity of cognitive ability testing. *International Journal of Selection and Assessment*, 20(1), 53-64. 10.1111/j.1468-2389.2012.00579.x

- Revilla, M., & Ochoa, C. (2017). Ideal and maximum length for a web survey. *International Journal of Market Research*, 59(5), 557-565. 10.2501/IJMR-2017-039
- Rojon, C., McDowall, A., & Saunders, M. N. (2011). On the experience of conducting a systematic review in industrial, work, and organizational psychology. *Journal of Personnel Psychology*, 10(3), 133-138. 10.1027/1866-5888/a000041
- Rojon, C., Okupe, A., & McDowall, A. (2021). Utilization and development of systematic reviews in management research: What do we know and where do we go from here? *International Journal of Management Reviews*, 23(2), 191-223. 10.1111/ijmr.12245
- Rosseel, Y. (2012). lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software, 48*(2), 1-36. 10.18637/jss.v048.i02
- Rousseau, D. M., Manning, J., & Denyer, D. (2008). Evidence in management and organizational science: assembling the field's full weight of scientific knowledge through syntheses. *Academy of Management Annals*, 2(1), 475-515. 10.1080/19416520802211651
- Ryan, A. M., & Huth, M. (2008). Not much more than platitudes? A critical look at the utility of applicant reactions research. *Human Resource Management Review*, 18(3), 119-132. https://doi.org/10.1016/j.hrmr.2008.07.004
- Ryan, A. M., McFarland, L., & Baron, H. (1999). An international look at selection practices: Nation and culture as explanations for variability in practice. *Personnel Psychology*, 52(2), 359-392. 10.1111/j.1744-6570.1999.tb00165.x
- Ryan, A. M., & Ployhart, R. E. (2000). Applicants' perceptions of selection procedures and decisions: A critical review and agenda for the future. *Journal of management*, 26(3), 565-606. 10.1177/014920630002600308
- Ryan, A. M., & Ployhart, R. E. (2014). A century of selection. Annual Review of Psychology, 65, 693-717. 10.1146/annurev-psych-010213-115134

- Rynes, S. L., & Connerley, M. L. (1993). Applicant reactions to alternative selection procedures. *Journal of Business and Psychology*, 7(3), 261-277
- Sackett, P. R., Lievens, F., Van Iddekinge, C. H., & Kuncel, N. R. (2017). Individual differences and their measurement: A review of 100 years of research. *Journal of Applied Psychology*, *102(3)*, 254-273. 10.1037/apl0000151
- Saks, A. M., & McCarthy, J. M. (2006). Effects of discriminatory interview questions and gender on applicant reactions. *Journal of Business and Psychology*, 21(2), 175-191. 10.1007/s10869-006-9024-7
- Scarborough, J. (2020, March 16). The Candidate Experience Gap: How employers are getting it wrong 95% of the time. TMP Worldwide. http://insights.tmpw.co.uk/post/102g1f6/thecandidate-experience-gap-how-employers-are-getting-it-wrong-95-of-the-time
- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. Psychological bulletin, 124(2), 262.
- Schinkel, S., van Dierendonck, D., van Vianen, A., & Ryan, A. M. (2011). Applicant reactions to rejection: Feedback, fairness, and attributional style effects. *Journal of Personnel Psychology*, 10(4), 146-156.
- Schuler, H. (1993). Social validity of selection situations: A concept and some empirical results. In:
 H. Schuler, J. L. Farr, & M. Smith (Eds), *Personnel Selection and Assessment: Individual* and Organizational Perspectives. Hillsdale, NJ: Lawrence Erlbaum Associates. https://doi.org/10.4324/9780203772706
- Schwencke, B. (2022, April 22). *The Four-Point Model of Candidate Experience*. Test Partnership. https://www.testpartnership.com/blog/four-point-model-candidate-experience.html
- Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching*, 5(9), 9-16. 10.5539/elt.v5n9p9

- Scott, D. (2005). Critical realism and empirical research methods in education. *Journal of Philosophy of Education*, *39*(4), 633-646. 10.1111/j.1467-9752.2005.00460.x
- Searle, R. H., & Al-Sharif, R. (2018). Recruitment and selection. In D. G. Collings, G. T. Wood, & L. T. Szamosi, (Eds.), *Human Resource Management* (pp. 215-237). Routledge.
- Sedgwick, P. (2014). Cross sectional studies: advantages and disadvantages. *Bmj*, 348, 1-2.10.1136/bmj.g2276
- Sharf, J. C., & Jones, D. P. (2000). Employment risk management. In J. Kehoe (Ed.), Managing Selection in Changing Organizations: Human Resource Strategies (pp. 271–318). San Francisco, CA:Jossey-Bass Inc.
- Smith, M. (2022, May 3). Job openings and number of people quitting hit record highs—it's 'the pit of despair for employers'. CNBC. https://www.cnbc.com/2022/05/03/job-openings-andthe-number-of-people-quitting-hit-record-highs.html
- Snape, D., Meads, C., Bagnall, A., Tregaskis, O., & Mansfield, L. (2017). What Works Wellbeing: A guide to our evidence review methods. Project Report. What Works Centre for Wellbeing, London, UK.
- Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. *Journal of Business Research, 104,* 333-339. 10.1016/j.jbusres.2019.07.039
- Sousa, F. J. (2010). Chapter 9: Metatheories in research: positivism, postmodernism, and critical realism. In Woodside, A. G. (Ed.), Organizational Culture, Business-to-Business Relationships, and Interfirm Networks. Emerald Group Publishing Limited, Bingley, pp.455-503. https://doi.org/10.1108/S1069-0964(2010)0000016012
- Steiner, D. D., & Gilliland, S. W. (1996). Fairness reactions to personnel selection techniques in France and the United States. *Journal of Applied Psychology*, 81(2), 134-141. 10.1037/0021-9010.81.2.134

- Suen, H. Y., Hung, K. E., & Lin, C. L. (2019). TensorFlow-based automatic personality recognition used in asynchronous video interviews. *IEEE Access*, 7, 61018-61023. 10.1109/ACCESS.2019.2902863
- Sylva, H., & Mol, S. T. (2009). E-Recruitment: A study into applicant perceptions of an online application system. *International Journal of Selection and Assessment*, 17(3), 311-323. 10.1111/j.1468-2389.2009.00473.x
- Tay, C., Ang, S., & Van Dyne, L. (2006). Personality, biographical characteristics, and job interview success: A longitudinal study of the mediating effects of interviewing selfefficacy and the moderating effects of internal locus of causality. *Journal of Applied Psychology*, 91(2), 446-454. 10.1037/0021-9010.91.2.446
- Tene, O., & Polonetsky, J. (2013). A theory of creepy: technology, privacy and shifting social norms. Yale Journal of Law and Technology, 16(59), 1-32
- Traynor, M., Neill, C. M., & Roulston, A. (2022). Personal statement versus psychological test as admission to the nursing degree: an evaluation. *BMC nursing*, 21(1), 1-11. 10.1186/s12912-022-00919-w
- Truxillo, D. M., Bauer, T. N., Campion, M. A., & Paronto, M. E. (2006). A field study of the role of big five personality in applicant perceptions of selection fairness, self, and the hiring organization. *International Journal of Selection and Assessment*, 14(3), 269-277. 10.1111/j.1468-2389.2006.00351.x
- Truxillo, D. M., Bodner, T. E., Bertolino, M., Bauer, T. N., & Yonce, C. A. (2009). Effects of explanations on applicant reactions: A meta-analytic review. *International Journal of Selection and Assessment*, 17(4), 346-361. 10.1111/j.1468-2389.2009.00478.x
- van Esch, P., & Black, J. S. (2019). Factors that influence new generation candidates to engage with and complete digital, AI-enabled recruiting. *Business Horizons*, 62(6), 729-739. 10.1016/j.bushor.2019.07.004

- van Esch, P., Black, J. S., & Arli, D. (2021). Job candidates' reactions to AI-enabled job application processes. *AI and Ethics*, 1(2), 119-130. 10.1016/j.bushor.2019.07.004
- van Teijlingen, E., & Hundley, V. (2001). The importance of pilot studies. *Social Research Update,* 35.
- Van Vianen, A. E., Taris, R., Scholten, E., & Schinkel, S. (2004). Perceived fairness in personnel selection: Determinants and outcomes in different stages of the assessment procedure. *International Journal of Selection and Assessment, 12*(1-2), 149-159. 10.1111/j.0965-075X.2004.00270.x
- Visser, R., & Schaap, P. (2017). Job applicants' attitudes towards cognitive ability and personality testing. SA Journal of Human Resource Management, 15(1), 1-11. 10.4102/sajhrm.v15i0.877
- Viswesvaran, C., & Ones, D. S. (2004). Importance of perceived personnel selection system fairness determinants: Relations with demographic, personality, and job characteristics. *International Journal of Selection and Assessment*, 12(1-2), 172-186. 10.1111/j.0965-075X.2004.00272.x
- Viswesvaran, C., & Ones, D. S. (2018). Non-test methods and techniques used in employee selection. In D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), *The SAGE Handbook of Industrial, Work & Organizational Psychology: Personnel Psychology and Employee Performance* (pp. 451–473). Sage Reference.
- Wheeler, A. R., Shanine, K. K., Leon, M. R., & Whitman, M. V. (2014). Student-recruited samples in organizational research: A review, analysis, and guidelines for future research. *Journal of Occupational and Organizational Psychology*, 87(1), 1-26.
- Wiechmann, D., & Ryan, A. M. (2003). Reactions to computerized testing in selection contexts. International Journal of Selection and Assessment, 11(2-3), 215-229. 10.1111/1468-2389.00245

- Woods, S. A., Ahmed, S., Nikolaou, I., Costa, A. C., & Anderson, N. R. (2020). Personnel selection in the digital age: a review of validity and applicant reactions, and future research challenges. *European Journal of Work and Organizational Psychology*, 29(1), 64-77. 10.1080/1359432X.2019.1681401
- Wu, M. J., Zhao, K., & Fils-Aime, F. (2022). Response rates of online surveys in published research: A meta-analysis. *Computers in Human Behavior Reports*, 7, (2-11). https://doi.org/10.1016/j.chbr.2022.100206

Appendix 1 – Reactions Questionnaire Items

Section 1

Q3.1. Doing well on this selection process means a person can do well on the schemes.

Q3.2. A person who scored well in this selection process will be a good future leader in the Civil Service.

Q3.3. I understood in advance what the selection process would be like.

Q3.4. I knew what to expect during the selection process.

Q3.5. I had ample information about what the format of the selection process would be.

Q3.6. I could really show my skills and abilities through this process.

Q3.7. This process allowed me to show what my leadership skills are.

Q3.8. This selection process gives applicants the opportunity to show what they can really do.

Q3.9. I was able to show what I can do during this selection process.

Q3.10. I had a clear understanding of when I would get my results for Stage 1 of the selection process.

Q3.11. I knew when I would receive feedback about my results for Stage 1 of the selection process.

Q3.12. I am satisfied with the amount of time it will take to get feedback on my results to Stage 1 of the selection process.

Q3.13. There was enough communication during the selection process.

Q3.14. I was able to ask questions about the process.

Q3.15. I am satisfied with the communication that occurred during the selection process.

Q3.16. I would have felt comfortable asking questions about the process if I had any.

Q3.17. I understood what contribution each of the activities would make to the selection process outcome.

Q3.18. It would be clear to anyone that this process is related to leadership.

Q3.19. The content of the selection process is clearly related to leadership skills.

Section 2

Q4. Doing well on this activity means a person can do well on the schemes.

Q5. A person who scored well on this activity will be a good future leader in the Civil Service.

Q6. I understood in advance what the activity would be like.

Q7. I knew what to expect during the activity.

Q8. I had ample information about what the format of the activity would be.

Q9. I could really show my skills and abilities through this activity

Q10. This activity allowed me to show what my leadership skills are.

Q11. This activity gives applicants the opportunity to show what they can really do.

Q12. I was able to show what I can do during this activity.

Q13. I had a clear understanding of when I would get my results for this activity.

Q14. I knew when I would receive feedback about my results for this activity.

Q15. I am satisfied with the amount of time it will take to get feedback on my results for this activity.

Q16. I understood what contribution this activity would make to the selection process outcome.

Q17. The content of the activities did not appear to be prejudiced.

Q18. The activities themselves did not seem too personal or private.

Q19. The content of the activities seemed appropriate.

Q20. It would be clear to anyone that this activity is related to leadership.

Q21. The content of the activity was clearly related to leadership skills.

Appendix 2 – Personality Questionnaire Items

Goldberg's IPIP representation of Costa and McCrae's (1992) NEO-PI-R Domains (Goldberg, 1999, as cited in Goldberg et al., 2006):

NEUROTICISM

10-item scale + keyed: Often feel blue. Dislike myself. Am often down in the dumps. Have frequent mood swings. Panic easily.

keyed: Rarely get irritated.
Seldom feel blue.
Feel comfortable with myself.
Am not easily bothered by things.
Am very pleased with myself.

EXTRAVERSION

10-item scale

- + keyed: Feel comfortable around people.
 - Make friends easily. Am skilled in handling social situations. Am the life of the party. Know how to captivate people.

- keyed : Have little to say.

Keep in the background. Would describe my experiences as somewhat dull. Don't like to draw attention to myself. Don't talk a lot.

OPENNESS TO EXPERIENCE

10-item scale

+ keyed: Believe in the importance of art.

Have a vivid imagination.

Tend to vote for liberal political candidates.

Carry the conversation to a higher level.

Enjoy hearing new ideas.

- keyed: Am not interested in abstract ideas.

Do not like art. Avoid philosophical discussions. Do not enjoy going to art museums. Tend to vote for conservative political candidates.

AGREEABLENESS

10-item scale

+ keyed: Have a good word for everyone.
Believe that others have good intentions.
Respect others.
Accept people as they are.
Make people feel at ease.

- keyed : Have a sharp tongue.

Cut others to pieces. Suspect hidden motives in others. Get back at others. Insult people.

CONSCIENTIOUSNESS

10-item scale

+ keyed: Am always prepared.
Pay attention to details.
Get chores done right away.
Carry out my plans.
Make plans and stick to them.

- keyed : Waste my time.

Find it difficult to get down to work. Do just enough work to get by. Don't see things through. Shirk my duties.