School of Management Studies

Do Rater Personality Traits Moderate the Relationship Between Intelligence and Rating Accuracy in Interviews?

By

Laeeqa Bassier
BSSLAE001

Supervisor: Professor Francois de Kock

A dissertation submitted in partial fulfilment of the requirements for the award of the Degree of Master of Commerce in Industrial and Organisational Psychology
Faculty of Commerce

University of Cape Town

2021

Compulsory Declaration:

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to and quotation in this dissertation from the work or works of other people has been attributed, cited and referenced.

Signature: L. Bassier Date: 15/03/2021

Signed by candidate

Acknowledgements

I would like to give special thanks to my supervisor, Francois de Kock, for his support in the writing and consolidation of this dissertation, as well as the primary researchers, Philippa Rauch, Kirsty Mostert, Naadir Soeker, Kavesh Vanmali and Melvynne Muturiki.

Thank you to my family, especially my mother and friends who provided me with unconditional love, support and inspiration throughout this challenging year.

Abstract

Research on judgment accuracy in human resource management shows that various rater characteristics predict accuracy, but emerging findings suggest that these individual differences may interact with one another (rather than being *direct effects*). The present study aimed to add to this area of research by determining how rater personality traits may moderate the relationship between GMA and rating accuracy. Secondary data collected in a prior study of police managers undergoing a seven-week managerial training course in South Africa (N =146) were analysed. The findings supported that selected rater-personality traits may moderate the relationship between intelligence and rating accuracy. For example, rater intelligence was a better predictor of accuracy when the judge was more agreeable. Intellectance and conscientiousness were found to have no significant moderating effect on the relationship between intelligence and rating accuracy. Only three out of the Big Five Personality traits were examined in this research study. Importantly, the study contributed to theory by expanding the Good Judge model (De Kock et al., 2020), analysing how individual differences in the ability and trait domains may potentially interact to influence accuracy. In addition to enhancing our understanding of how rater personality constructs may affect accuracy, the study discusses important implications for practices, such as rater training and selection.

Keywords: moderator effect, accuracy, intelligence, judgment, rating, personality traits, good judge model.

Table of Contents

Compulsory Declaration: **2**

Acknowledgements **3**

Abstract **4**

List of Figures **8**

List of Tables **9**

Chapter 1 .... **10**

1.1 Introduction 10

1.2 Aim of This Research 13

1.3 Research Question 13

Chapter 2: Literature Review **14**

2.1 The Importance of Rating Quality 14

2.2 Judgement Accuracy 14

2.3 Theoretical Approaches to Judgement Accuracy 15

2.3.1 Realistic Accuracy Model 15

2.3.2 Social Judgement Theory 16

2.4 Individual Differences of the Good Judge 17

2.5 Intelligence 17

2.5.1 Research Findings on Intelligence and Rating Accuracy 18

2.6 Personality 19

2.7 Personality and Judgement Accuracy 19

2.7.1 Research Findings 19

2.7.2 How Personality may Moderate Intelligence and Rating Accuracy 20

2.7.3 Does Personality Moderate the Effect of Intelligence on Accuracy? 21

2.8 Moderation 22

2.8.1 Agreeableness 23

2.8.2 Openness to Experience 25

2.8.3 Conscientiousness 26

2.9 Summary and Conclusion 27

Chapter 3: Method 28

3.1 Research Design 28

3.1.1 Research Design of Original Study by De Kock et al. (2015) 28

3.1.2 Materials 29

3.2 Participants 29

3.3 Measures 30

3.3.1 Accuracy Criterion Measure 30

3.3.2 Predictor Measures 31

3.4 Statistical Procedures 32

3.5 Data Analysis 33

3.5.1 Moderation Analysis 34

3.5.2 Moderated Multiple Regression (MMR) 34

3.5.3 Best Practices for MMR 35

3.6 Additional Analyses 36

3.6.1 Earlier Approaches to Estimating Interaction Effects 36

3.7 PROCESS 37

Chapter 4: Results 39

4.1 Data Preparation 39

4.2 Testing Assumptions: Multiple Regression Analysis 39

4.3 Skewness and Kurtosis 42

4.4 Descriptive Statistics 43

4.5 Correlation Analysis 45

4.5.1 Tests of Hypotheses: Moderator Effects 46

4.5.2 Simple Slope Analysis: Agreeableness 47

4.6 Summary of Moderation Results 52

4.7 Further Analyses 52

Chapter 5: Discussion 53

5.1 Introduction 53

5.2 Main Findings 53

5.2.1 Agreeableness 54

5.2.2 Openness to Experience 55

5.2.3 Conscientiousness 56

5.3 Moderated Multiple Regression and Statistical Power 58

5.4 Limitations 59

5.5 Recommendations for Future Research 60

5.6 Implications for Theory and Future Research 61

5.7 Implications for Practice 62

5.8 Conclusion 62

References 64

Appendix A: A 20-item short version of the International Personality Item Pool Big Five Factor Markers 75

List of Figures

Figure 1: Conceptual Diagram of the Moderator Effect.. 22
Figure 2: Simple Moderation Model 34
Figure 3: Normal Probability Plot (P-P) of the Regression Standardised Residuals of the Three Regression Models 41
Figure 4: Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Agreeableness... 48
Figure 5: Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Openness to Experience 50
Figure 6: Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Conscientiousness 51

List of Tables

Table 1: Descriptive Statistics and Correlations for Study Variables 44
Table 2: Moderator Test Results: Agreeableness as a Moderator of Effect between General Mental Ability and Accuracy 47
Table 3: Moderator Test Results: Openness to Experience as a Moderator of Effect between General Mental Ability and Accuracy 49
Table 4: Moderator Test Results: Conscientiousness as a Moderator of Effect between General Mental Ability and Accuracy.. 51

Chapter 1

The following chapter provides a brief overview of individual differences in rating quality and how they influence judgement accuracy and intelligence. The chapter also discusses the main aims of this study, followed by the research question.

1.1 Introduction

Personality judgements are considered a vital aspect of everyday life; they are consequential for the individual who makes them (Funder, 2012). In human resource management (HRM), ratings play a pervasive role (Guion & Highhouse, 2011). Importantly, several organisations rely on ratings to make crucial decisions regarding selection, promotion and performance management (De Kock et al., 2020). The attributes of the "good judge" have compelled researchers for several decades (Funder, 2012). Establishing a more precise perspective of the characteristics which affect accuracy may assist organisations with recruitment and training processes, which are to hire the best possible employees.

The focus of individual differences research has been focused more towards investigating particular abilities associated with being a good judge (Letzring, 2008; Powell, 2008). Even though most reviews of interview literature suggest that good judges have higher general intelligence, researchers continue to search for other attributes that may help identify accurate judges (Letzring, 2008). The accurate judge does not appear to score higher or lower than others on any specific personality trait (De Kock et al., 2020). Findings from De Kock et al. (2020) discovered none of the Big Five Personality traits to be consistent predictors of judgement accuracy in HRM. Studies that have reported significant effects have illustrated small effects (Cohen, 1988), except for openness to experience and Agreeableness (Christiansen et al., 2005), therefore suggesting that traits may not be crucial to shaping rating quality. Furthermore, De Kock et al.'s (2020) review proposes that there might be more complex interactions between personality and judgment accuracy, such as moderating effects, which may account for these inconsistent trends.

One study discovered a gender distinction in the accuracy of personality judgments, and their findings report that interpersonal orientation was positively associated to judges' accuracy in rating the personality traits of videotaped targets (Vogt & Colvin, 2003). In contrast, Lippa and Dietz (2000) found no gender differences in making accurate judgements, even though intelligence was associated with more accurate personality judgements. These inconsistent findings potentially indicate that personality indirectly influences judgement accuracy.

It is anticipated that accurate judges generate evaluations in HRM, which demonstrate sufficient rating quality. Rating quality can be referred to as the extent to which an individual's ratings are considered accurate, not only as a measure of other individuals' attributes, but also as predictors of significant outcomes (e.g. employment performance criteria) (De Kock et al., 2020). However, researchers are presently experiencing inconsistencies relating to the profile of the accurate judge in HRM (Jones & Born, 2008).

Most research on individual difference attributes is known to be tested in relative isolation and limited research is available on how these variables may interact when predicting rating accuracy within organisational settings. Moderation arises when the association between an independent and dependent variable is conditional upon a third variable (Li et al., 2019), commonly referred to as a moderator variable (Aiken et al., 1991; Dawson, 2014). Furthermore, in the absence of a moderator variable, the independent variable (X) is anticipated to predict the dependent variable (Y) (Li et al., 2019). However, the interaction term (Z) implies that the strength of the XY association will differ depending on the level of Z. Certainly, the interaction term (Z) may theoretically enhance, weaken or reverse the effect of the independent variable and dependent variable (Li et al., 2019). Evaluating interactions assists researchers with the capacity to enhance their perspectives of socioeconomic relationships by discovering the circumstances which apply to such relationships (Andersson et al., 2020).

The findings of Christiansen et al. (2005) demonstrate that conscientiousness and agreeableness traits moderated the relationship between intelligence and acquaintance accuracy (De Kock et al., 2020). In addition, openness to experience was found to be positively associated with measures of General Intelligence (De Kock et al., 2020), an attribute known to foster higher accuracy levels. Furthermore, these findings are consistent with past research (Hollenbeck et al., 1988; Wright et al., 1995), demonstrating that employees' performance on specific tasks or activities is more efficient when people are driven to perform successfully on the job and present greater levels of relevant abilities. Additionally, Wright et al.'s (1995) study supports the interactive findings of Hollenbeck et al. (1988), who propose that personality and mental ability interactively predict job performance.

Moreover, if the moderator effect exists, this means that there are not only direct effects which influence judgement accuracy, as demonstrated by the Good Judge Model (De Kock et al., 2020), but that there are also indirect effects, such as individual characteristics interacting with each another which may influence judgement accuracy. Importantly, this could be relevant to rating practices, as the success of rater selection as well as rater training could be dependent on an appropriate framework of the specific interviewer characteristics which drive accuracy. Additionally, the moderation effect could be pertinent to organisations as they could consider implementing rater training interventions which look how different personality traits moderate GMA on rating accuracy, not only the direct effects of judgement accuracy, and develop the relevant knowledge structures which could be crucial in curriculum design and evaluation (Christiansen et al., 2005). Individuals may develop the ability to make better decisions once they are capable of accurately evaluating the personality of people, as well as adopting the relevant knowledge concerning the process of accurate judgement, which could assist HRM practitioners to make more appropriate judgements (Letzring, 2008). Importantly, speculating about the good judge hypothesises that a good judge may be knowledgeable about how personality correlates to behaviour, as well as demonstrates high cognitive abilities and general intelligence (Funder, 1999).

One approach explaining the process of accurate judgement is the Realistic Accuracy Model (RAM), which stems from Brunswik's Lens Model (Brunswick, 1956; Funder, 1995). According to RAM (Funder, 1995), an individual is required to fulfil four stages successfully in order for accuracy to be attained. For example, the individual has to perform an action pertinent to the personality characteristic being judged, the relevant cue is required to be accessible to as well as detected by the judge, and lastly the judge has to use the cue efficiently to formulate a judgement (Letzring, 2008). From a cognitive point of view, obtaining an accurate outcome regarding an employee's personality is exceptionally abstruse. As a minimum requirement, accuracy demands a good memory to recall past behaviour (Christiansen et al., 2005). Because inferential trait frameworks are required to be developed and examined, the capacity to cope with abstruse notions is also pertinent. As a result, it is anticipated that the association between intelligence and accuracy may potentially depend on one's personality, where personality functions as an indirect rather than direct effect in rating accuracy.

1.2 Aim of This Research

The limited amount of research on the moderating effects of personality on rating accuracy and intelligence requires further development. Investigating how individual differences interact with one another may be relevant, as most research tends to focus on the direct effects. The moderation effect may therefore help us to understand these inconsistencies of the relationship between accuracy and personality and that these mixed findings may indicate the presence of a moderator or moderators. The present study extended the research of Christiansen et al. (2005) by demonstrating how personality may moderate GMA on rating accuracy.

Moderation analysis is utilised when a researcher is particularly fascinated in investigating whether or not the size of a variable's effect on an outcome variable of interest is dependent on a third variable (Hayes, 2012). Investigating this interaction effect also added to the Good Judge Model (De Kock et al., 2020) by demonstrating how individual characteristics act as moderators and explained this complex interaction. Importantly, another fundamental aim of this study was to investigate whether the findings of the research study by Christiansen et al. (2005) were spurious or robust, if replicated in a new study. By studying replication, researchers build a stronger science (Farrar et al., 2020).

1.3 Research Question

To address the aims of this research, the research question is: Do rater-personality traits moderate the effect of other individual difference constructs, such as GMA, on rating quality? More specifically, do personality traits such as agreeableness, conscientiousness and openness to experience moderate the relationship between intelligence and rating accuracy in interviews?

Chapter 2: Literature Review

The following chapter introduces the theoretical background and empirical research around rater-personality traits and how they influence rating quality. Most research on individual difference attributes tends to be examined in relative isolation; limited research has acknowledged how these variables interact in the prediction of rating accuracy within organisational settings. This chapter begins by describing rating accuracy and its theoretical underpinnings. This is followed by a section on theoretical approaches to judgement accuracy and intelligence. The individual attributes of the good judge are then described in detail, along with the theoretical framework for moderating effects. Lastly, the three personality dimensions which influence rating accuracy and intelligence are discussed.

2.1 The Importance of Rating Quality

Individuals may have the ability to make more efficient decisions in the workplace if they have the competency to judge employees or candidates more accurately (Letzring, 2008). Rating quality can be referred to as the extent to which an individual's ratings are considered accurate, not only as a measure of other individual's attributes, but also as predictors of significant outcomes (i.e. employment performance criteria) (De Kock et al., 2020). A fundamental element within the procedure of accurate judgment involves the individual making the judgement (Letzring, 2008). Speculations concerning the good judge propose that it is someone who is well-informed about the ways in which personality pertains to behaviour, demonstrates advanced levels of intellectual competencies and is motivated to be accurate, along with various other attributes (Funder, 1995). Research also indicates that good judges are individuals who are capable of attaining fairly large amounts of relevant information concerning their targets (Letzring, 2008; Powell; 2008). As a result, accurate judges are expected to provide assessments in HRM that illustrate sufficient rating quality due to the competencies associated with them.

2.2 Judgement Accuracy

In several research studies involving judgment accuracy, the focal point encompasses the question: "is the judgement correct?" (Funder, 2012, p. 177). Thus, accuracy could be defined as the true principle against which judgement is differentiated

(Funder, 1995). Importantly, all domains of science require assessments of validity and reliability. Funder (2012) proposes three criteria which could be implemented when evaluating accuracy. The first is known as *self-other agreement*, and various studies have evaluated accuracy based on the degree to which the rater's judgement is in line with that of the target's own judgement of personality (Funder, 2012). The second criterion is *consensus,* also known as *other-other agreement*, which is defined as the level of agreement between the personality judgements manifested by two or more individuals regarding another person (Funder et al., 2004). Moreover, it is crucial to be mindful of the fact that consensus is not inevitably related to what the individual is actually like; a group of people may agree about the personality of a particular individual, which could result in inaccurate stereotypes, affecting the way that individual is being judged.

Lastly, the third criterion for accurate judgement is referred to as *behaviour prediction* (Funder, 1996). A judgement of personality is considered accurate to the extent that it can anticipate independent evaluations of the behaviour of the individual being judged (Funder & West, 1993). Although the three criteria mentioned present an idea of ways in which researchers could evaluate accuracy, it is also necessary to comprehend the procedure which enables accurate judgement. Therefore, Funder (2012) recommends an accuracy model to fulfil this objective.

2.3 Theoretical Approaches to Judgement Accuracy

Various theories, including the RAM (Funder, 1995) and Social Judgment Theory (SJT) (Cooksey, 1996), emerge in rating accuracy literature to explain the phenomenon. These theoretical frameworks provide an empirical grounding for understanding the foundations of judgment accuracy.

2.3.1 Realistic Accuracy Model

The RAM, established by Funder (1995), integrates an individual's personality trait with the perceiver's correct judgment of that trait (Funder, 2012). RAM also provides a general description for moderators of accuracy and clarifies the ways in which these variables interact (Funder, 1995). Importantly, for this association to be developed for accurate judgement to be attained, four processes need to take place: relevance, availability, detection and utilisation (Funder, 1995). The Good Judge Model (De Kock et al., 2020) links rater individual differences to fundamental judge processes, namely cue detection and cue utilisation, thought to cause accuracy (De Kock et al., 2020).

Judges may affect the detection phase by demonstrating more attentive characteristics and acknowledging their surroundings instead of only being concerned with their inner emotions and thoughts (Adams, 1927; Letzring, 2008). For example, making more eye contact and showing full interest in what the individual has to say are considered behaviours that could possibly illustrate that the judge is concentrating on the target. Moreover, in situations like this, targets are more likely to reveal more information about themselves when the judge appears to be interested in them and paying attention to what they are saying (Letzring, 2008). Judges may influence the utilisation phase by having the ability to effectively integrate and comprehend cues (Letzring, 2008). Additionally, judges who present high cognitive abilities, such as general and social intelligence, are predicted to be highly efficient during this phase of RAM, as they have the potential to recall and effectively manipulate cues (Funder, 1999). Moreover, certain researchers have found promising results relating to the concept that intelligence and mental complexity are associated with judgement accuracy (Christiansen et al., 2005; Vernon, 1933).

2.3.2 Social Judgement Theory

The SJT (Cooksey, 1996) was established from Egon Brunswik's functionalist psychology ideas and regression analysis. The theoretical framework proposes that one's judgement is associated with the reality of a social environment and could be conceptualised as a "lens" (Dowding & Thompson, 2003). Furthermore, this notion of a lens may be utilised to mould the manner in which several types of knowledge and information correlate with the actuality of a given judgement setting, as well as the manner in which people tend to use knowledge and information to make judgements (Dowding & Thompson, 2003).

Through its illustrative mechanism, the Lens Framework, STJ is known as the most broadly used, systems-orientated approach for interpreting human judgement in explicit ecological conditions (Cooksey, 1996). The Lens Model proposes that the ecological condition (e.g. the problem of the individual) is situated on the left division of the model. There are different types of cues which are theoretically associated with this ecological condition (e.g. the individual's symptoms or signs), with distinct significance attached to them (Dowding & Thompson, 2003). Moreover, the judges then utilise these cues to construct judgements (the right-hand division of the Lens Framework). If the cues are considered equally important by the judge as they are connected to the ecological condition, then the judgement is considered more accurate (Dowding & Thompson, 2003). If the judge considers one cue more significant than the other, then their judgement may be less accurate. Interestingly, this approach has been utilised in various fields such as finance, weather forecasting and health (Harries & Harries, 2001).

2.4 Individual Differences of the Good Judge

Various individual difference constructs which influence judgement accuracy have been explored, even though findings are not convincing. Conventional studies tend to centre their research on the main effects of broad personality traits and judgement accuracy (Christiansen et al., 2005; Funder, 1999), as well as general intelligence and rating accuracy (Borman, 1979; Christiansen et al., 2005; Lippa & Dietz, 2000). However, the inconsistencies relating to this area of research require more research in order to present firm conclusions. Because most research on individual differences has been assessed in relative isolation, analysing how these variables interact in the prediction of rating accuracy may lead to promising results.

2.5 Intelligence

Intelligence can be operationalised as the ability of an individual to deliberately alter their thinking to new requirements (Salgado, 2017). It is considered a complex system which integrates many levels of analysis, including biological, geographic, computational, sociological, and others (Wagner, 2000). Intelligence may also be referred to as GMA which contains a vast spectrum of precise cognitive abilities, such as abstract reasoning, planning, problem-solving and learning from experience (Di Domenico & Fournier, 2015).

Furthermore, General Mental Ability (GMA), as well as specific intelligence assessments, is known to be the significant predictors of employee performance, task performance, as well as training ability (Salgado, 2017). Therefore, mental ability assessments occupy the most prominent position among the personnel selection processes (Salgado, 2017). Individual differences in intelligence are considered relatively stable over time, for example Deary et al. (2002) assessed participants at age 10 and then again at 80 and discovered that general intelligence presented a rank-order stability coefficient of .63. Moreover, judging individuals is known to be a highly complex task which places more strain on the information processing load of the rater (De Kock et al., 2020; Lance et al., 2004). General mental abilities may be a pertinent element to producing accurate judgements.

2.5.1 Research Findings on Intelligence and Rating Accuracy

Results from De Kock et al. (2020) demonstrate that overall, rater general intelligence was found to be a consistent predictor of rating accuracy compared to all the other individual differences they reviewed. From their meta-analysis of all individual differences reviewed, uncorrected validity coefficients were average .54 (Schneider & Bayroff, 1953); .31 (Borman, 1979); .24 (Borman & Hallam, 1991); .23–.34 (Hauenstein & Alexander, 1991); .36 (Lippa & Dietz, 2000); .25 (Christiansen et al., 2005). Results demonstrated that effect sizes were often rather modest (e.g. uncorrected $.10 < r < .30$) and certain studies, such as Letzring (2008) and (Powell, 2008) discovered no relationship between intelligence and accuracy (De Kock et al., 2020).

The model formulated by De Kock et al. (2020) proposes that the availability of contextual information could impede or encourage cue detection and utilisation. Moreover, the model predicted that boundary conditions, namely interview structures or situational attributes, could affect the type of abundance of cues accessible to the judge (De Kock et al., 2020). Rationally, the influence of GMA on rating accuracy could develop with job complexity. Additionally, the social cognitive theory proposes that intelligence could be presumed to correlate more with accuracy when it simulates a prominent indisputable responsibility in providing accurate judgements, for example when information processing loads are high (Ambady & Rosenthal, 1992).

For example, cue-rich situations, such as strictly structured interviews, could bestow less mental pressure on judges compared to cue-poor situations, such as low structured interviews, where very limited good information is retrieved (De Kock et al., 2020). Therefore, it is anticipated that intelligence might be a stronger predictor of accuracy in less structured interviews. The results of Christiansen et al. (2005) were congruent to the results of Lippa and Dietz (2000), who establish that GMA was correlated to accuracy of personality judgements made in terms of non-verbal cues of outsiders, and expanded the results to judgments of naturally occurring acquaintances. Additionally, findings from Christiansen et al. (2005) also demonstrate that GMA was positively associated with intelligence and openness to experience. Generally, the results illustrated that higher elevations of GMA and openness tended to make more accurate judgements (Christiansen et al., 2005) Similarly, Furnham et al. (2007) found agreeableness to be significantly correlated to intelligence.

2.6 Personality

Personality is commonly referred to as predispositions to react to stimuli in a particular manner (De Kock et al., 2015; John et al., 2008). Essentially, personality includes an intense behavioural propensity focus (Mayer et al., 1999). The Five-Factor Model (FFM) of personality (Digman, 1990) serves as a conceptual model for organising traits which judges are asked to perceive in others and is also used to evaluate the personality traits of judges (Lippa & Dietz, 2000). The FFM of personality has developed into a broadly used model for classifying personality traits (Powell & Goffin, 2009). Labels, as well as common characteristics related to attributes in the model, include Extraversion (sociable or active), agreeableness (considerate or trusting), conscientiousness (dependable, organised or persistent), openness to experience (intellectual or imaginative), and neuroticism (moody, tense) (Barrick et al., 2000; Powell & Goffin, 2009). One of the most common techniques used to measure the Big Five personality traits is the International Personality Item Pool (Goldberg, 1992).

2.7 Personality and Judgement Accuracy

The dimensions within the FFM have been associated with numerous organisational objectives, such as job performance and training success (Powell & Goffin, 2009). Judges' personalities could stabilise their social functioning in the organisation, as well as factors of interpersonal judgment (De Kock et al., 2020). Precisely, personality traits could influence an individual's ability to create accurate notions of others because theoretically they may possibly be connected to the stages of information processing in RAM, specifically, cue detection and utilisation (De Kock et al., 2020).

2.7.1 Research Findings

Previous findings have presented that individuals who demonstrate higher openness are considered more inquisitive and often find joy in working with theoretical notions and beliefs (Goldberg, 1992). It is therefore rational to assume that they are also more inclined to actively construct cognitive depictions of individuals' traits and actions, discover trends of congruencies and formulate and assess propositions regarding other individuals' behaviour (De Kock et al., 2020; Kruglanski & Ajzen, 1983). Conscientiousness commonly presents itself in greater detail orientation; however, it could also influence the way in which we shape impressions concerning others (Goldberg, 1992). For example, in cue detection, judges who are highly conscientious are predicted to be more observant compared to low conscientious judges and also demonstrate more consistency in cue utilisation (De Kock et al., 2020).

Based on earlier studies, no trait is found to be a consistent predictor of rating accuracy. Studies that have analysed these correlations have discovered small to medium effect sizes if they reported significant effects (De Kock et al., 2020). Furthermore, Christiansen et al. (2005) discovered that only openness to experience out of the Big Five personality traits demonstrated a small to medium effect size (r =.23, p < .05) when predicting interview accuracy. Some researchers have proposed analysing specific personality facets rather than broad personality spheres, which might provide a more explicit depiction of the association between personality and judgement accuracy.

2.7.2 How Personality may Moderate Intelligence and Rating Accuracy

A fundamental admonition to results concerning the individual attributes of the good judge is that there is no specific trait that has appeared to be consistently related to judgement accuracy (Davis & Kraus, 1997; Letzring, 2008); some have shown to predict rating accuracy and some have presented no relationship. A possible explanation for these mix findings may be due to the presence of moderator variables. Most researchers have failed to focus on the indirect effects personality may have on judgment accuracy, and perhaps the effect of personality on intelligence and rating accuracy is indirect rather than direct.

For example, Bergold and Steinmayr (2018) explored the moderating effects between general intelligence and the Big Five personality traits in predicting academic performance. Certain theoretical frameworks predicting academic performance conclude that GMA and personality significantly predict academic performance in relative isolation, that is both variables' contribution to academic performance is perceived as incremental (Chamorro-Premuzic & Arteche, 2008). However, earlier frameworks relating to the prediction of performance on the relationship between intelligence and motivation proposed that both measures should interact indirectly with one another in a statistical manner (Bergold & Steinmayr, 2018). Specifically, they anticipated that the relationship between performance and intelligence should be moderated by motivation, such that when motivation is low, intelligence is not an important factor in the prediction of performance compared to when motivation is high (Bergold & Steinmayr, 2018). This may be due to the fact that intellectual individuals who present low levels of motivation could possibly dispose their intellectual skills rather than working on their overall performance (Sackett et al., 1998). Because motivation and personality are related (Steinmayr & Spinath, 2007), a theoretical assumption arises that performance is an interactive function of intelligence and personality (Bergold & Steinmayr, 2018). Thus, personality could possibly moderate the relationship between intelligence and performance.

Various researchers have posited that these mixed finding may also be as a consequence of the use of different criteria for accuracy and measures of accuracy (Letzring, 2008), which could have significant outcomes for how accuracy is associated with the traits and behaviours of the judge. Another argument is that certain practitioners present uncertainty about whether judgment accuracy is an individual difference (Kenny & Albright, 1987), predominantly in terms of low reliability accuracy results across the sample used.

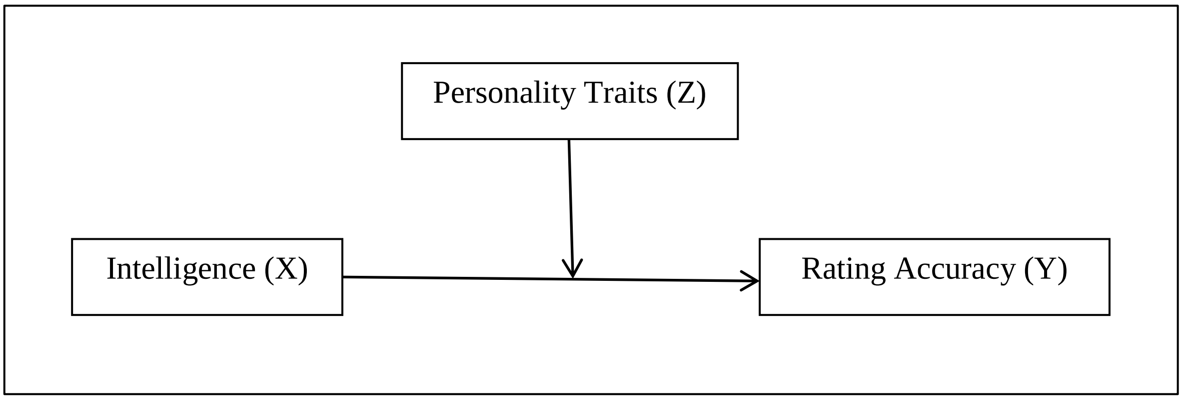
2.7.3 Does Personality Moderate the Effect of Intelligence on Accuracy?

2.7.3.1 Moderation Theoretical Model

Drawing from the theoretical frameworks discussed above, to contribute to the model by De Kock et al. (2020), which illustrates the direct effects of accuracy, this present study aimed to construct a model which demonstrated how personality may moderate the influence of other individual differences constructs on rating quality (showing interaction effects). Importantly, the conceptual diagram proposed that the relationship between intelligence and judgement accuracy is dependent on one level on another variable (e.g. a personality trait) (see Figure 1). This complex interaction between variables is referred to as a moderator effect.

Figure 1

Conceptual Diagram of the Moderator Effect



2.8 Moderation

In the field of psychology, the use of moderator analyses has a long history as it is often used in the analysis of variance, commonly referred to as an interaction effect (Shevlin et al., 2015). Furthermore, interaction, also known as moderating, effects hypothesises that the relationship between two variables is dependent on the value of a third variable (Aguinis & Gottfredson, 2010). Traditionally, several researchers have utilised the term interaction to describe the combined effect of manipulated independent variables within experimental designs (Gardner et al., 2017), whereas “moderation” has been used to explain the integrated effect of continuous variables in non-experimental designs (Pedhazur & Schmelkin, 1991). However, both notions define the multiplicative association between independent or predictor variables as related to a dependent variable.

Statistically the calculations are identical, therefore researchers use interaction and moderation terms interchangeably (Gardner et al., 2017). Furthermore, any variable could be a moderator, whether measured independently or continuously. Identifying and determining pertinent and significant interaction effects relating to relationships between independent and dependent variables is central to theory in social science (Aiken et al., 2003; Andersson et al., 2020). Interactions assist researchers with the capacity to enhance their perspectives of socioeconomic relationships by discovering the circumstances under which they apply (Andersson et al., 2020). Importantly, analysing interaction effects allow the extension or expansion of common relationships to contexts which the primary research may not have considered, as well as provide more precise predictions concerning the relationships (Andersson et al., 2020).

When a moderating effect is continuous, researchers generally rely on Moderated Multiple Regression (MMR) (Aiken et al., 1991). Specifically in this study, when testing the association between intelligence and judgement accuracy, this particular relationship is dependent on one level of another variable (a specific personality trait). Interestingly, Christiansen et al. (2005) illustrate that individuals' conscientiousness and agreeableness traits moderated the relationship between intelligence and acquaintance accuracy. More specifically, their results indicate that in individuals who presented high levels of conscientiousness and agreeableness, intelligence was found to be a more significant predictor of accuracy when individuals expressed these traits (Christiansen et al., 2005).

Since the notion that performance on a specific task is considered an interactive function of personality and competency (Maier, 1955), one could assume that the association between ability and performance may be less strong when a specific personality trait is not present. To understand this complex interaction between the variables of interest, one could refer to the key judgement processes of RAM (Funder, 1995), namely cue detection and cue utilisation, as mentioned previously. The Good Judge Model (De Kock et al., 2020) illustrates that the availability of contextual information could impede or encourage cue detection and utilisation. Comprehending how situations influence behavioural expressions is necessary for discovering the degree to which behavioural information is pertinent to the target's personality (Funder, 1995). It is predicted that judges who are cue sensitive have the ability to detect both verbal and non-verbal stimuli as soon as they happen (De Kock et al., 2020). Due to the limited amount of research on how personality may moderate the relationship between GMA and rating accuracy, future research should delve deeper into this topic. The empirical findings of Christiansen et al. (2005) fail to explain why the moderation effect might have occurred, which raises the question regarding the generalisability of the findings to new samples.

2.8.1 Agreeableness

Agreeableness is a broad personality trait, defined by cooperativeness, tolerance, soft-heartedness, good-natured and friendliness (Barrick & Mount, 1991; Goldberg, 1999; O'Connor & Athota, 2013). Some researchers have proposed agreeableness to be a fundamental concept in the evaluation of individual differences when interviewing a candidate (Witt et al., 2002). However, agreeableness appears to be the most pertinent to employee performance in conditions in which collaboration and team work are needed. Job contexts which involve high levels of interpersonal interaction require more tolerance, flexibility, friendliness and courteous behaviour. Thus, it is anticipated that agreeable individuals tend to deal with conflict more cooperatively, seek common understanding and maintain social relations (Digman, 1990).

Some researchers discovered that more agreeable individuals were more accurate judges (Vogt & Colvin, 2003). However, literature relating to this topic has yielded mixed results (Taft, 1955). Even though early research indicated a positive association between interpersonal orientation and judgement accuracy, technical limitations render these results inconclusive (Cronbach, 1955). Moreover, imperative support on the personality traits related to judgement accuracy was provided by Davis and Kraus (1997). Their research findings concluded that individuals who trust and are sociable provide more accurate judgements compared to those less invested in building interpersonal relations. However, one concern is that most of this research involved focusing on the accuracy of judgements of emotions rather than of personality characteristics.

In contrast, certain findings indicate that individuals who are lower in sociability and confidence are more accurate judges (Ambady et al., 1995). Interestingly, Christiansen et al. (2005) found the association between agreeableness and accuracy to be more abstruse. When the judge presented high levels of agreeableness, intelligence was a more efficient predictor of accuracy. Similarly, another study analysing interaction effects demonstrated a positive relationship between agreeableness and job performance among those high in cognitive ability, but a negative association among those low in cognitive ability (Hollenbeck et al., 1988).

Furthermore, it is proposed that performance on certain job activities is best when people are driven to succeed on the job and demonstrate increased levels of pertinent abilities (Wright et al., 1995). Recently, studies on intelligence and rating accuracy have demonstrated that they may not be associated in a bivariate sense, but that their relationship might be moderated by certain personality traits, accounting for the mixed results. For example, intelligence might only matter for accuracy when the judge is difficult or less sociable (low on agreeableness). A possible explanation for why intelligence might only matter for difficult people could be due to the fact that individuals who have the propensity toward being difficult are known to maintain greater levels of intelligence (Baker & Bichsel, 2006). Because they present high cognitive abilities, they are therefore able to develop impressions of others accurately through cue utilisation and cue detection.

Therefore, it is proposed that:

Hypothesis 1: The relationship between intelligence and rating accuracy is moderated by agreeableness, such that it is stronger for judges high in agreeableness than for judges low in agreeableness.

2.8.2 Openness to Experience

Openness to experience can be conceptualised as the propensity to get involved in intellectual tasks, as well as a desire for novel ideas and experiences (Furnham et al., 2007). Additionally, traits related to this dimension include being cultured, authentic and broad-minded (Barrick & Mount, 1991). Furthermore, it also involves a cognitive component, which specifies the way in which knowledge or information is processed and organised (Zimprich et al., 2009). Individuals who demonstrate higher openness are considered more inquisitive and often enjoy working with theoretical notions and beliefs (Goldberg, 1992). Therefore, it is expected that judges high in openness are more likely to actively construct mental depictions of individuals' traits and actions and interpret cues effectively (De Kock et al., 2020; Kruglanski & Ajzen, 1983). Generally, according to literature, individuals who present higher levels of general intelligence and openness tended to be more accurate judges (Christiansen et al., 2005).

Personality traits might be related to job performance; however, only for certain jobs or criteria. One would expect openness to experience to be a significant predictor of a specific performance criterion, such as training competency (Barrick & Mount, 1991). This personality trait is anticipated to be related to training competency as it evaluates personal attributes, such as broadmindedness, intelligence and inquisitiveness (Barrick & Mount, 1991), which are all characteristics associated with positive perceptions towards learning experiences. Moreover, DeYoung et al. (2014) have proposed that openness to experience is fundamentally related to general intelligence as they share very similar traits, therefore it is expected that intelligence and judgement accuracy may be weaker when openness to experience is absent. According to RAM (Funder, 1999), accuracy is an interactive function of components such as the availability of utilisation of behavioural information. Consequently, cognitive and personality differences in judges may influence different parts of the framework which interact to determine accuracy. Christiansen et al. (2005) found that judges with higher levels of openness to experience were no more accurate judges than those who were lower on these traits. Because this study included police managers involved in a training programme, it may be interesting to investigate whether intelligence mattered for accuracy when the judge was less curious or creative compared to when the judge was more broad-minded and inquisitive.

Thus, it is proposed that:

Hypothesis 2: The relationship between intelligence and rating accuracy is moderated by openness to experience such that it is stronger for judges high in openness to experience than for judges low in openness to experience.

2.8.3 Conscientiousness

Conscientiousness is an individual difference construct, known as the tendency to pursue socially prescribed norms for impulse control, to be goal-driven and dependable (Di Domenico & Fournier, 2015). Conscientious employees tend to demonstrate more discipline and ambition, whereas employees who present low levels of conscientiousness tend to be deceptive, impetuous and disorganised (Byrne et al., 2005). Furthermore, conscientious individuals are expected to perform successfully as they are more confident in their skills and more persevering (Barrick & Mount, 1991). It is anticipated that in cue detection, judges who are more conscientious are predicted to be more observant compared to those who are less conscientious and demonstrate more consistency in cue utilisation (De Kock et al., 2020). When a judge is highly conscientious, they are expected to have the ability to effectively integrate and interpret cues due to their high cognitive abilities, as they are able to recall and manipulate cues successfully.

Moreover, Bergold and Steinmayr’s (2018) findings illustrate that the personality trait conscientiousness consistently interacted with intelligence when predicting academic achievement. In accordance with previous research, their study also demonstrates that the predictive value of intelligence is even higher when participants present higher scores on conscientiousness (Bergold & Steinmayr, 2018; Ziegler et al., 2009). Additionally, Christiansen et al. (2005) also found that interviewers’ conscientiousness moderated the relationship between intelligence and acquaintance accuracy. Because conscientiousness is expected to be associated with job performance due to its ability to evaluate personal characteristics such as being dependable, responsible and persistent (Barrick & Mount, 1991), individuals with high conscientiousness might be more successful at judging others as they are considered more detail-orientated and hardworking, which might enable them to make use of their full cognitive abilities (Jensen & Patel, 2011).

As a result, conscientious judges may be more successful at identifying important priorities which may enhance organisational objectives. Therefore, it may be interesting to investigate whether conscientiousness is a moderator between intelligence and rating accuracy across different samples, and not only within an educational setting as most previous studies have done. Additionally, it may also add to limited research on why intelligence predicts or drives accuracy for hard-working, dependable individuals as opposed to unreliable, lazy individuals.

Considering these arguments, it is proposed that:

Hypothesis 3: The relationship between intelligence and rating accuracy is moderated by conscientiousness, such that it is stronger for judges high in conscientiousness than for judges low in conscientiousness.

2.9 Summary and Conclusion

The literature review has provided a theoretical and practical context to the moderating effects of personality on the relationship between judgement accuracy and intelligence. Examining how personality could possibly moderate the effect of other individual differences attributes on rating quality is still theoretically underdeveloped. Therefore, the hypotheses put forward have been based on Christiansen et al.'s (2005) study, as well as existing literature that have investigated related constructs.

Chapter 3: Method

The following chapter describes the method followed to conduct the study. The present study utilised secondary data from a study by De Kock et al. (2015). Firstly, the research design will be discussed, followed by materials used in the original study. Secondly, participants, as well as measures used in the secondary study, are explained. Lastly, statistical procedures followed by data analysis are discussed in detail.

3.1 Research Design

This study implemented a quantitative research approach using a cross-sectional design. This particular design allows researchers to establish both the risk factor and the outcome simultaneously, also known as point-in time surveys (Omair, 2015). Furthermore, the research study employed a secondary research design which is defined as research that involves the use of pre-existing data (Trzesniewski et al., 2011). The dataset was approximately 10 years old; however, it was still deemed acceptable for further examination.

There are various benefits of adopting a secondary design. Firstly, the most common advantage is that it is less time-consuming and more cost-effective, allowing researchers to delve deeper into data analysis (Trzesniewski et al., 2011). Secondly, a secondary analysis gives researchers the opportunity to obtain high quality data, due to the fact that most of the primary data sets are well-funded projects with a large sample size (Johnston, 2017). Thirdly, re-examining data and interpreting it from a new perspective could perhaps yield new findings and add to the original study, creating a more in-depth understanding (Johnston, 2017).

3.1.1 Research Design of Original Study by De Kock et al. (2015)

The original study employed a descriptive cross-sectional design (Babbie & Mouton, 2006) which was followed by a correlational approach. Furthermore, this present study was a reanalysis of an existing published study by De Kock et al. (2015), as a way to constructively replicate the findings relating to moderator effects in Christiansen et al. (2005). Data collection was completed in a single session at the end of 2011. After highlighting that the research project was integrated with the interview training, the researchers discussed the rating process and materials. The next step included the researchers presenting five video-recorded interview sections to the sample of participants through the use of a projector (De Kock et al., 2015). The first video was a practice run, followed by a brief overview discussing the ratings and final clarification of any concerns or questions regarding the rating procedure (De Kock et al., 2015). The managers were then required to rate the four remaining video segments and complete their interview rating sheets independently. Lastly, they completed the individual difference measures such as the Wonderlic Personnel Test (2002) and the International Personality Item Pool Big Five factor markers (Goldberg, 1992), before being debriefed and thanked for their participation. Because the study was comparable with Christiansen et al. (2005) in terms of the design as well as choice of variables, and the requisite data for moderation analysis were available, it lent itself to further study of the generalisability of the moderator effects observed in Christiansen et al. (2005).

3.1.2 Materials

Development of interview videos and materials. In the original study, the researchers used videotaped portions of interviewee performance as a catalyst, because it allowed for the presentation of similar stimuli to all participants (De Kock et al., 2015). This process involved researchers video-recording five graduate participants in an interview setting which could essentially assist them with future job applications. The interviews were semi-structured with a competency-based format.

Furthermore, the interview questions covered two distinct aspects, namely communication and people management. These particular aspects were decided upon due to their extensive use in employee interviews (De Kock et al., 2015; Huffcutt et al., 2001) and were therefore deemed appropriate to the fabricated job position for which they were applying. In addition to this, the researchers made use of eight questions rated on a 7-point Likert scale, with 1 being a poor response to 7 being an excellent response (De Kock et al., 2015). The real-life interviews that were recorded involved a professional interviewer asking applicants to respond to the same set of questions in a logical order and the final video segments were shortened to approximately 5 minutes each.

3.2 Participants

In the original study (De Kock et al., 2015), a non-probability purposive sampling approach was employed as the sample was required to meet certain criteria. Police managers from an organisation in South Africa undergoing a 7- week managerial training programme for promotional purposes were recruited as participants in the original study (De Kock et al., 2015). Furthermore, to increase the external validity of the study, all participants who were recruited were required to have prior (at least 5 years) comparable work experience as managers.

There were a total of 146 managers in the sample (24.6% female, 75.4% male). This was considered a strength of the original study, as most psychology research is laboratory research involving university students, therefore one cannot be certain whether these findings generalise to the field. Additionally, field studies like this are potentially useful to help address the science-practice divide. The original sample consisted of 71.2% Black African, 17.3% White, 9.4% mixed race, and 2.2% Asian participants (De Kock et al., 2015). The average age of participants was 43.7 (SD = 5.36) years. A large proportion of the sample was ranked as Captains (57, 6%) and the rest were ranked as either Warrant Officers (35.3%) or Lieutenant (7.2%). Overall, the aforementioned officers represent the most likely interviewers in common organisations as they fall between junior first-line supervisors and senior management (De Kock et al., 2015). In the original study, missing data were deleted pairwise to maximise the available data for analyses; there was no discernible pattern in the missingness. In terms of academic qualifications, some participants had post-secondary school certificates (37.9%) and the rest had only a senior secondary school certificate (62.1 %). Furthermore, English was the most commonly spoken work language in the participating company. More information about the study design and sample can be found in the published paper by De Kock et al. (2015).

3.3 Measures

The following section discusses the specific measures used in this study. Responses from three measures were extracted from the primary dataset. For more information on demographic items, one could refer to the original paper by De Kock et al. (2015).

3.3.1 Accuracy Criterion Measure

In the original study, accuracy scores operated as the dependent variable. Moreover, this study calculated an accuracy score known as the Borman Differential Accuracy (DA) score (Sulsky & Balzer, 1988). The score was calculated within-personal profile correlations at the dimensional level, using an r-to-Fisher’s-z transformation (De

Kock et al., 2015). Additionally, Borman's (1979) DA index is generated by correlating a rater's ratings for each dimension with corresponding true scores across ratees, producing a DA score for each dimension (Sulsky & Balzer, 1988). The distinction between Cronbach's DA Score and Borman's DA measure is that Borman's DA yields only correlation information without taking into account the actual differences between subject ratings and true scores (Sulsky & Balzer, 1988). Furthermore, it is important to note that Borman's DA has been referred to as correlational accuracy, and should avoid being muddled with the correlational component of Cronbach's DA (Sulsky & Balzer, 1988). Consequently, because Borman's (1979) DA only provides correlational information, it is best to consider it an index of rater validity (Sulsky & Balzer, 1988). This particular measure was selected as it provides crucial preliminary information for accuracy. In addition, this method evaluates the similarity between the entire set of judgements made by a judge and the target (De Kock et al., 2015). In this study, the judgements are the interview competency judgements of the raters in the study of video targets responding to interview questions.

3.3.2 Predictor Measures

3.3.2.1 General Cognitive Ability

To test for GMA, the Wonderlic Personnel Test (2002) was used. All participants were required to complete this test, which was revised at the start of the testing session. The Wonderlic Personnel Test is a 50-item timed test (approximately 12 minutes), including items such as word comparisons, disarranged sentences, number comparisons, analysis of geometric figures and problems requiring mathematical and logical solutions (De Kock et al., 2015). For example, spatial reasoning questions are used to assess logical thinking, and participants are asked to apply their critical reasoning skills in this section of the questionnaire. Speed questions involve participants engaging quickly in questions which are thrown at them, and questions could be as simple as "what is not a fruit in this list of items?"; the participant would then have to choose from a list of five possibilities. The Wonderlic Personnel Test demonstrates good predictive validities for a broad range of criteria (Wonderlic, 1998) and reliability estimates[1] commonly vary between .82 and .95 (Wonderlic, 2002). Generally, across various samples, reliability coefficients for the

[1] The original study calculated internal consistency reliability of the Wonderlic measure as .75, but acknowledged that it is not appropriate for speeded measures (Nunnally & Bernstein, 1994).

Wonderlic Personnel Test are frequently high, for example in a study by Rosopa and Schroeder (2009), test-retest reliability coefficients were above.82 and internal consistency estimates above .88.

3.3.2.2 Personality

Congruent to the original study by De Kock et al. (2015), to measure the personality of participants, this study also made used of the 20-item short version of the International Personality Item Pool Big Five factor markers (Goldberg, 1992) (see Appendix A). This process involved participants assessing how each item in the scale explained itself on a 5-point Likert scale, from very inaccurate to very accurate (Goldberg, 1992). According to De Kock et al. (2015), the mean inter-item correlations for each scale were similar to those in earlier published research studies (Donnellan et al., 2006). Cronbach alpha reliability estimates in a previous South African study presented acceptable reliability coefficients, such as .88 for conscientiousness, .85 for openness to experience, and .81 for agreeableness (Morgan & De Bruin, 2010). Additionally, these reliability estimates compare well with those presented by Taylor and De Bruin (2006), as well as Khorramdel and Von Davier (2014), with estimates from .77 to .86.

3.4 Statistical Procedures

There are particular processes that are the “gold standard” for conducting secondary research. Importantly, the research method comprises how the researcher collects, analyses and interprets the data (Creswell, 2009; Johnston, 2017). Secondary analysis is a systematic technique with substantive and evaluative steps; however, there is limited literature that defines a specific method, thus Johnston (2017) has suggested a process that begins with the construction of the research question, the identification of a relevant data set, followed by elaborate evaluation of the data set.

The first step is that researchers should choose an existing dataset, which is congruent with their research question (Greenhoot & Dowsett, 2012). Finding an existing dataset that contains similar constructs of interest is not always easy. The main aim of this study was to investigate the moderating effects of personality on intelligence and rating accuracy. The research question that directed this study was: Do rater-personality traits moderate the effect of the relationship between intelligence and rating accuracy? Most research has investigated the direct effects of personality on rating accuracy and intelligence; however, De Kock et al. (2020) recommend further research to investigate the moderating effects that personality may have on intelligence and rating accuracy. Consequently, because there was previously collected data on this specific topic, it was deemed appropriate to use De Kock et al.'s (2015) dataset as it contained similar constructs addressing the research question of this study. Once the Commerce Department of the University of Cape Town approved the approach, ethical clearance was provided to conduct the secondary analysis. Access to the dataset was provided by the original authors in SPSS format.

The next phase involved the evaluation of the dataset to confirm its appropriateness for the research topic (Johnston, 2017). Once an appropriate dataset is selected, the researcher should spend extensive time familiarising themselves with the dataset by analysing the codebooks or other relevant documents relating to sampling design and processes (Greenhoot & Dowsett, 2012). Additionally, during the preparation phase of the dataset for analysis and interpretation, it is necessary for the researcher to have a carefully thought-out theoretical framework or model along with an explicit idea of the different types of variables required to test the framework (Greenhoot & Dowsett, 2012). In this study, before using the secondary dataset, a conceptual framework was developed involving three key variables, i.e. rater-personality traits, intelligence, and rating accuracy, to test the model. Once the variables were identified, the relevant dataset was provided for secondary analysis. The original data set by De Kock et al. (2015) consisted of five measures, data which were not relevant to this study were deleted, i.e. the study only made use of agreeableness, conscientiousness and intellectance subscales from the Big-Fiver Personality traits, Borman Differential Accuracy Scale for rating accuracy and, lastly, GMA Scale from the secondary dataset. Furthermore, intellectance was relabelled as openness to experience. Once the data were examined and relevant adjustments were made, the dataset was stored in a secure folder which only the researchers had access to, ensuring protection and security of the dataset.

3.5 Data Analysis

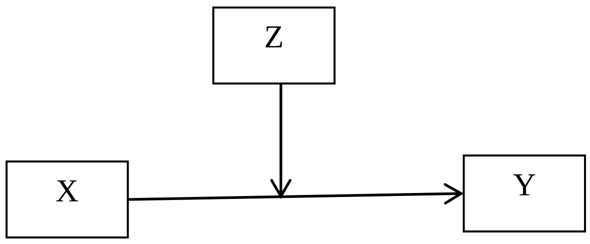
To analyse the secondary data, the Statistical Package for Social Scientists (SPSS) version 23 was used. Descriptive statistics were calculated to determine the characteristics of the sample group, such as means, standard deviations and frequencies (Field, 2018). A bivariate correlation analysis was performed using Pearson product moment correlation coefficients (Cohen, 1988). Lastly, a moderated multiple regression was used to test the moderator hypotheses.

3.5.1 Moderation Analysis

Moderation analysis is utilised when a researcher is particularly fascinated by examining whether the magnitude of a variable's effect on an outcome variable of interest is dependent on a third variable (Hayes, 2012). Importantly, moderation effect is also commonly referred to as interaction (Hayes & Rockwood, 2017), for example if X's effect on Y is moderated by Z, then X and Z interact (see Figure 2). Moreover, any variable could be a moderator, whether measured independently or continuously.

Figure 2

Simple Moderation Model



3.5.2 Moderated Multiple Regression (MMR)

Importantly, when a moderating effect is continuous, researchers generally rely on MMR (Aiken et al., 1991). MMR appears to be the most common statistical tool of choice for estimating moderating effects within organisational sciences (Aguinis & Gottfredson, 2010). Furthermore, MMR involves developing a regression model that predicts the outcome based on a predictor X, a second predictor Z hypothesised to be a moderator, and the product term between X and Z, which carries information on the moderating effect of Z on the X-Y relationship (Aguinis et al., 2017). The regression coefficient for the product term XZ from which X and Z have been partialed out represents information on the presence and magnitude of the moderating effect (Aguinis et al., 2017). MMR analysis provides researchers with vital information which is not provided by tests of the equality of correlation coefficients (Aguinis, 2004).

3.5.3 Best Practices for MMR

Due to the fact that MMR is known to lack statistical power, it is crucial to consider design, instrument and analysis concerns in order to enhance the likelihood that existing moderators will be discovered if they do exist (Aguinis & Gottfredson, 2010). Therefore, this study followed best practising guidelines outlined by Aguinis and Gottfredson (2010) for estimating the moderating effects using MMR. Firstly, before collecting data, it is recommended that researchers develop a rationale for why they assume a moderating effect should exist. Moreover, they should also consider design features such as the sample size and reliability of variables (Aguinis & Gottfredson, 2010).

Based on sample size, because the moderator variables in this study where continuous sample size becomes a concern, the smaller the sample size, the weaker the power (Aguinis & Gottfredson, 2010). However, the sample size of this study was deemed acceptable based on the normative sample size guidelines by Aguinis and Gottfredson (2010). A sample size of 146 participants with a three-predictor variable equation was used for each test. Furthermore, regarding reliability issues, statistical power is enhanced when the reliability of measures (or variables) is improved (Aguinis & Gottfredson, 2010).

The variables in this study presented appropriate construct reliability for all measures. Secondly, another instrumental concern to account for is scale coarseness (Aguinis, Pierce et al., 2009). A scale is considered coarse when a variable is continuous and measured using items where different true scores are disintegrated into the same group or category (Aguinis & Gottfredson, 2010). In the present study scale coarseness was minimised, as continuous scales were used to measure continuous variables, enhancing statistical power.

Post-data collection best practices include, avoiding the dichotomisation of continuous variables (Aguinis & Gottfredson, 2010), and this was avoided in the present study. Additionally, when dealing with interaction terms, mean-centring of predictor scores was required, and continuous variables were mean-centred in this study. This was relevant as mean-centring achieves the objective of making the interpretation of the first-order coefficients meaningful by the technique of re-scaling (Aguinis & Gottfredson, 2010). It is also recommended that researchers create graphs to demonstrate the nature of the moderation effect. A simple slope analysis was created in this study to visualise interaction effects and to provide a more robust understanding of the moderating effects.

3.6 Additional Analyses

A power analysis was conducted as MMR is known to suffer from low statistical power (Aguinis et al., 2001). One concern raised was that several studies using MMR lacked sufficient power to detect the interaction effects (Memon et al., 2019). Power refers to the possibility of obtaining a statistically significant outcome (Cohen, 1988). The common convention is that power of a statistical test is required to be above .80 (Cohen, 1988; Goodhue et al., 2007). Consequently, it is recommended that researchers compute a power analysis in their studies when using MMR to dismiss the idea that their analysis is underpowered (Memon et al., 2019). Low statistical power means that researchers may conclude that no significant interaction effect is present, when in fact there is one. Therefore, it was deemed appropriate to perform a statistical power analysis to investigate whether no statistically significant interaction may be due to a lack of statistical power.

3.6.1 Earlier Approaches to Estimating Interaction Effects

Based on earlier approaches to testing moderation effects, Baron and Kenny (1986) highlight the importance of adopting an appropriate analytical technique for assessing moderation and considered four distinct cases. This approach to testing moderation hypothesis is known as the causal steps approach (Baron & Kenny, 1986). Even though this approach has been used by many researchers, it has been criticised on various grounds. Firstly, it is known to have low statistical power (Hayes, 2009). Secondly, their approach produces no test that the indirect effect (moderation) has occurred (Hayes, 2009). Instead, an indirect effect is implied logically by the results of hypothesis tests. Moreover, the Baron and Kenny (1986) approach is centred on testing null hypotheses, in line with techniques established during a period when the fundamental emphasis of statistical test was significance testing of null hypotheses (Chmura Kraemer et al., 2008).

Modern statistical tests currently place importance on meaningful effect sizes. Using PROCESS (Hayes, 2012), an extension for SPSS, allows researchers to generate an output for the indirect effect, including effect sizes as well as confidence intervals. Although Baron and Kenny (1986) encourage using the Sobel test to estimate indirect effects, they failed to explicitly emphasise it as part of their technique, and as a result it has often been overlooked by some researchers (Hayes, 2009). The Sobel test is an inferential method which uses a specialised t-test to establish a moderating effect (Chmura Kraemer et al., 2008). Furthermore, because the Sobel test lacks statistical power and relies on normal sampling distributions, Hayes (2012) suggests the utilisation of bootstrapping approaches to estimate indirect effects which are accessible via PROCESS. Bootstrapping is considered a form of robust statistic which prompts how a study would be replicated by resampling from a population (LaFlair et al., 2015), which does not suffer from the aforementioned limitations.

3.7 PROCESS

The following section will discuss PROCESS macro constructed by Hayes (2012), which was used to examine the moderating effect of personality on the relationship between intelligence and rating accuracy. PROCESS is a computation tool for SPSS which implements moderation or mediation analysis (Hayes, 2012). Importantly, it covers several of the analytical concerns researchers interested in conducting moderation or mediation analysis commonly confront (Hayes, 2012).

Based on the researcher’s theoretical framework and statistical diagrams that define each regression model, the researcher inserts each model that they want to estimate into PROCESS. For example, each variable serves a distinct role in the model (e.g. independent variable, dependent variable and moderator) and PROCESS evaluates all the path coefficients, standard errors, t- and p-values, confidence intervals and numerous other statistics (Hayes & Rockwood, 2017). Moreover, PROCESS makes use of ordinary least squares regression to estimate the parameters of each of the equations, a popular practice in observed variable path analysis (Hayes & Rockwood, 2017).

Using this technique may be the most efficient option as PROCESS integrates many of the functions of commonly used tools or programs such as SOBEL, RSQUARE and MBESS into one user-friendly method (Hayes, 2012). Furthermore, it also eliminates researchers having to educate themselves on various tools which perform only one specialised task. Statistical researchers might find PROCESS a valuable teaching tool, making it both convenient and easy to illustrate traditional and modern techniques to moderation and mediation analysis.

Moreover, inferences about moderation analysis statistics are based on bootstrapping techniques; because many of these statistics have irregular sampling distributions, using ordinary methods is perceived as more challenging (Hayes, 2012; Hayes & Rockwood, 2017). However, PROCESS may not be able to do everything a researcher might want to do. In certain instances, a structural equation-modelling program might be a better option to address a specific analytical problem (Hayes, 2012). In addition, using any structural equation-modelling program requires more coding as well as the competency to write that code, whereas PROCESS generates those statistics automatically. PROCESS was therefore considered the most effective approach as it generates output that would otherwise require great effort and programming skills to implement.

Chapter 4: Results

This chapter reports the results of analyses of secondary data from a prior study (De Kock et al., 2015). The data analytic strategy was guided by statistical good practice guidelines (Greenhoot & Dowsett, 2012) for secondary data analysis. Statistical analyses in this study included descriptive statistics, correlations, multiple regression, and MMR analyses (for moderator analysis). Statistical analyses were conducted with SPSS version 24 (Field, 2018). Moderator analysis was done using the PROCESS macro (Hayes, 2012) in SPSS. Following a discussion of the assumptions, the results of the statistical tests for each hypothesis are discussed.

4.1 Data Preparation

Prior to the statistical analyses, data preparation and cleaning were conducted. The original data set by De Kock et al. (2015) consisted of a broader set of variables than those relevant to this study. As such, after deleting irrelevant study variables, the data frame for this study consisted of measures of the independent (predictor) variables, namely personality traits (including agreeableness, conscientiousness and intellectance subscales, all from the Big-Five Personality framework), the intelligence measure, and demographic variables. The data frame also included the dependent variable measure of rating accuracy (Borman Differential Accuracy Scale). It is recommended that researchers avoid the urge to include non-focal variables into the analysis, searching for interesting correlations which are not theoretically driven, as this approach may increase the possibility of a Type 1 error (Greenhoot & Dowsett, 2012). Following a preliminary scanning of the data and relevant adjustments required specifying variable names and properties, descriptive statistics were inspected in order to ensure that all values fell within-range and no coding errors were evident. Given that validity and reliability analyses relevant to the study measures are reported in the original study (De Kock et al., 2015), these are not repeated here.

4.2 Testing Assumptions: Multiple Regression Analysis

Most statistical analyses rely upon specific assumptions concerning the variables used in the analysis (Field, 2018). Various statistical assumptions are required to be met in order for the data to be considered appropriate for a multiple regression analysis – the statistical analyses relied upon to test the main hypotheses in the present study.

Furthermore, it is important to test whether these assumptions are being met as results may be unreliable or untrustworthy, resulting in a Type 1 or Type 2 error, or over- or under-estimation of significance or effect sizes (Osborne & Waters, 2002). MMR follows the same set of assumptions as a multiple regression analysis as moderation is considered a particular type of regression analysis (Hayes, 2012).

However, because the independent variables in this study were continuous, mean centring was required. According to Field (2018), when one includes an interaction term in a model, if the independent variables are continuous, they must be centred. Three multiple regression models were then tested, where each model consisted of intelligence, one personality trait, as well as the interaction effect between the two (intelligence and one personality trait) as predictor variables, and rating accuracy as the dependent variable.

The first assumption for multiple regression involves the sample size. Greene and d'Oliveira (2005, as cited in Tabachnick & Fidell, 2007) propose a rule-of-thumb formula as a benchmark to determine the appropriateness of the sample size: $N > 50 + 8m$ (where m = the number of independent variables). The regression analysis in this study consisted of a maximum number of six independent variables, which would require a minimum sample size of 98 respondents; this study had a total number of 146 participants, therefore this assumption has been met. Secondly, for the assumption of multicollinearity, the tolerance values were used. All tolerance values were above the minimum benchmark of .10 for all three models (the smallest value was .909 and the highest .995).

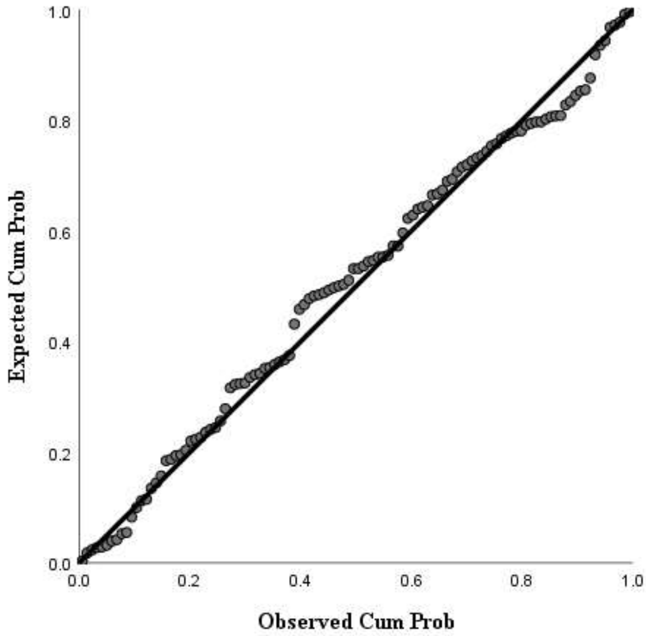
All continuous variables were centred to address the issue of multicollinearity (Field, 2018). Furthermore, outliers were tested, defined as a data point with standardised residual values above 3.3 or less than -3.3 (Tabachnick & Fidell, 2007). Standardised residuals all fell within the accepted range. The minimum standardised residual value out of the three models was -2.47 and the maximum was 2.64. Lastly, normality, linearity and homoscedasticity were tested by visual inspection of the Normal Probability Plot (P-P) of the regression standardised residual (refer to Figure 3). Moreover, this assumption is important as Non-normally distributed variables (e.g. highly skewed variables or substantial outliers) may distort relationships as well as significance tests (Osborne & Waters, 2002). The closer the dots lie towards the diagonal line, the closer to normal the residuals are distributed (Field, 2018). All assumptions were met for all three models and therefore data were considered appropriate for MMR analysis. Various assumptions for MMR were tested before performing the analysis, and all were met.

Figure 3

Normal Probability Plot (P-P) of the Regression Standardised Residuals of the Three Regression Models

Normal P-P Plot of Regression Standardized Residual between General mental ability and Openness to experience

Dependent Variable: Rating accuracy (Borman DA)



Normal P-P Plot of Regression Standardized Residual between General mental ability and Concientiousness

Dependent Variable: Rating accuracy (Borman DA)

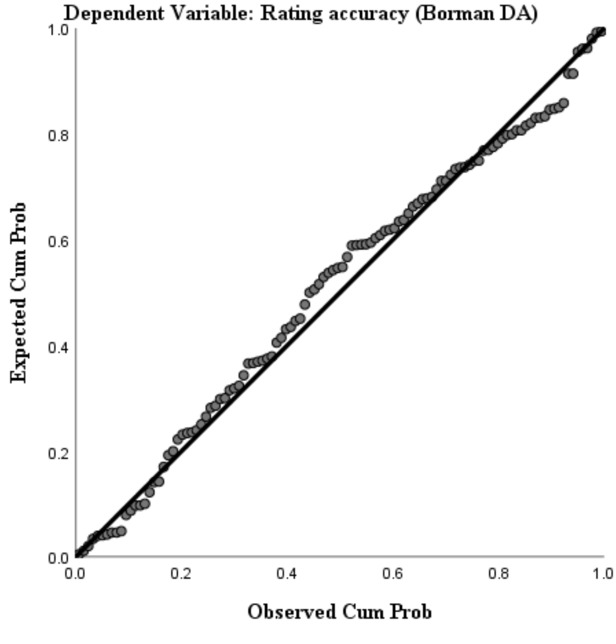
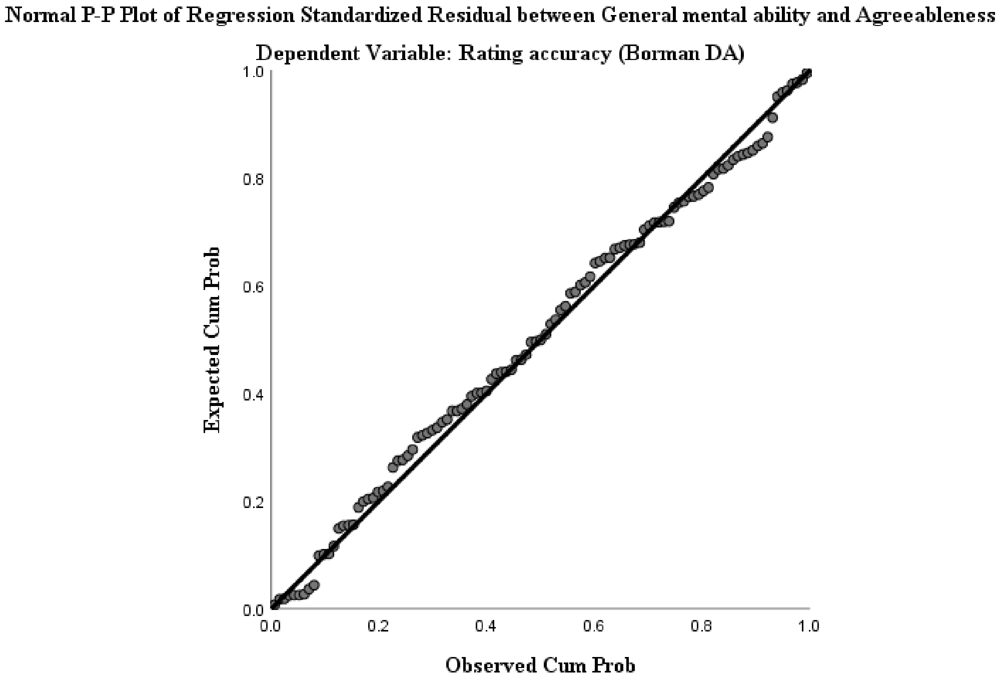


Figure 3 (continued)

Normal Probability Plot (P-P) of the regression standardised residuals of the three regression models (*continued)*



4.3 Skewness and Kurtosis

Skewness of the dataset can be referred to as the symmetry of the distribution; the degree to which the distribution is skewed to have either too many high scores or too many low scores within a distribution (Field, 2018). Moreover, kurtosis provides information concerning the “peakedness” of the distribution (Pallant, 2011). The further the kurtosis and skewness values are from zero, the greater the possibility that the dataset is not normally distributed (Pallant, 2011). Even though several statistical analyses are formulated on the assumption of normal distribution of data, in the field of social sciences research, it is accepted that this assumption is not always met (Egboro, 2015). Additionally, many of the parametric tests in SPSS are robust and may still be useful when this assumption of normality is not being strictly met (Egboro, 2015). The GMA scale skewness value was above .5, suggesting a moderate positive skewness. Kurtosis value for this scale was above zero, suggesting that the data was leptokuric (Field, 2018). For the personality measure, the overall skewness values for all three factor markers were below.5, suggesting a moderate negative skewness, in other words, managers scored relatively high on the personality scale. Additionally, the kurtosis values for all three variables were less than zero, indicating that the data were platykuric (Field, 2018). For the Borman's DA accuracy measure, the skewness value was above -.5, indicating a moderate negative skewness, i.e. participants generally scored high on the accuracy scale. Furthermore, the kurtosis value was slightly above zero, demonstrating that the data is somewhat leptokuric (Field, 2018).

4.4 Descriptive Statistics

Table 1 reports the descriptive statistics. The table reflects the means, standard deviations, skewness and kurtosis, and intercorrelation of the study variables

Table 1

Descriptive Statistics and Correlations for Study Variables

Variable	*M*	*SD*	Skewness	Kurtosis	1	2	3	4	5	5	7	8
1. GMA	12.42	4.40	.63	.72	-							
2. Agreeableness	13.03	4.37	.02	-.84	.28**	—						
3. Conscientious	14.02	4.07	-.04	-.83	.16	.40**	—					
4. OTE	12.76	4.15	-.08	-.69	.11	.40**	.54**	—				
5. DA[2]	.985	.649	-.30	.46	.20*	.15	.13	.04	—			
6. GMA x Agree	5.82	17.96	.32	1.82	.05	-.05	.09	-.15	-.19*	—		
7. GMA x Consc	3.20	17.54	.23	2.45	-.06	.10	.14	-.02	-.11	.48**	—	
8. GMA x OTE	2.48	18.12	.24	4.25	-2.2*	-.13	-.02	.03	-.20*	.23*	.30**	—

Note. *N*s range from 113 to 146; *M* = mean; *SD* = standard deviation; Conscientious = conscientiousness; GMA= General Mental Ability; OTE= openness to experience; DA= Differential Accuracy; GMA x Agree = interaction term for GMA and agreeableness; GMA x Consc = interaction term for GMA and conscientiousness; GMA x OTE = interaction term for GMA and openness to experience.

[2] Differential Accuracy = Borman DA scores are Fisher transformed (*r* to *z*) profile correlations between participants' ratings and true scores at dimension level.
*p < .05. **p < .001 (two-tailed).

For the GMA measure, the mean score was *M* =12.42 (*SD* = 4.40) (refer to Table 1), with a minimum value of 2 and the maximum was 28. Overall, similarly to the original study by De Kock et al. (2015), the GMA scores of managers within the sample were somewhat lower than those published in previous research studies using university students as samples. For the personality measure, agreeableness scores ranged from 4 to 20, with a mean for this subscale of *M* =13.03 (*SD* = 4.37). For conscientiousness, scores ranged from 6 to 20, the mean score for this subscale was *M* = 14.02 (*SD* = 4.07). Intellectance scores ranged from 4 to 20, the mean score for this subscale was *M* = 12.76 (*SD* = 4.15). For the Borman's DA accuracy measure, participants scores ranged from -.81 to 2.72, with a mean score of .985 (*SD* = .649). Participants generally scored moderately high on this scale, demonstrating high levels of accuracy. A high score on this scale indicates greater accuracy levels (Borman, 1979; Powell & Goffin, 2009).

4.5 Correlation Analysis

A bivariate correlation analysis was performed using Pearson product moment correlation coefficients. Table 1 illustrates the correlation coefficients. Furthermore, Cohen's (1988) recommendations for interpreting the strength of the correlation coefficients were used as a guideline. Correlations close to .1 are considered small (weak), .3 as medium (moderate) and .5 as large (strong) (Cohen, 1988).

A significant, moderate effect was found between GMA and agreeableness (r = .28, p < .001, N = 113), which is consistent with prior research (Christiansen et al., 2005). A significant, weak to moderate positive effect was found between GMA and rating accuracy (r = .20, p < .05, N = 141). Furthermore, the strength, direction and significance of the relationship were the same between agreeableness and openness to experience and agreeableness and conscientiousness (r = .40, p < .001, N = 113). Additionally, a significant, strong positive relationship was found between openness to experience and conscientiousness (r = .54, p < .001, N = 116).

A significant weak negative effect was also found between Borman's DA and the interaction term GMA and agreeableness (r = -.19, p < .05, N = 109), as well as the interaction term GMA and intellectance (r = -.20, p < .05, N = 112). A significant weak effect was discovered between the interaction term GMA and agreeableness and the interaction term GMA and intellectance (r = -.23, p < .05, N = 113). A significant strong effect was found between the interaction term GMA and agreeableness and the interaction term GMA and conscientiousness ($r = .48$, $p < .001$, $N = 116$). Additionally, a significant moderate positive effect was found between the interaction term GMA and conscientiousness, and GMA and intellectance ($r = .30$, $p < .001$, $N = 116$).

Since the data generally satisfied the relevant statistical assumptions (Tabachnick & Fidell, 2013), it was deemed acceptable to run multiple regression analysis. MMR was used for this research study as it seeks to examine an interaction effect, in this case whether the effect of intelligence (X) on rating accuracy (Y) is moderated by rater personality traits (Z). The analysis was conducted using the guidelines of Aguinis and Gottfredson (2010) for estimating and interpreting interaction effects. In particular, the analytical strategy relied on the following: First, implementing adequate design and measurement features, for example total sample size and making use of relevant computer programs for continuous moderators. Design choices and practical constraints in terms of statistical power analysis outcomes were considered (Aguinis & Gottfredson, 2010). Additionally, scale coarseness was minimised by using reliable scales (Aguinis, Pierce et al., 2009). Secondly, dichotomisation of variables was avoided and all continuous variables were mean-centred (Aguinis & Gottfredson, 2010). Thirdly, graphs were created to visualise interaction effects (Hayes, 2012), as well as estimation of effect sizes and practical significance.

Field (2018) recommends the use of the PROCESS macro for running moderation analysis (Hayes, 2012). As such, the PROCESS script was added as an extension to SPSS and used to conduct the moderation analysis. All analyses were run at a 95% level of confidence and 5000 bootstrap samples were used to estimate the moderation effect. When running the analysis, it was not necessary to centre the variables or calculate the interaction as PROCESS conducts this automatically (Field, 2018). The Johnson-Neyman (1936) method was used in PROCESS to investigate the significance of the interaction effects. Essentially, this technique involves analysing how the relationship between the predictor variable and outcome changes at various different values of the moderator (Field, 2018).

4.5.1 Tests of Hypotheses: Moderator Effects

Agreeableness. Hypothesis 1 stated that the relationship between intelligence and rating accuracy is moderated by agreeableness, such that it is stronger for judges high in agreeableness than for judges low in agreeableness.

To test this hypothesis, a model was specified with rating accuracy as the Y-variable; intelligence as the X-variable, and agreeableness as the moderator variable.

According to the moderation analysis, the moderation effect appears to be significant, *b* = -.007, 95% CI [-.014, -.000], *t* = -2.113, *p* = .037, indicating that the relationship between intelligence and rating accuracy is moderated by agreeableness (refer to Table 2). When agreeableness traits are low, there is a significant positive relationship between rating accuracy and intelligence, *b* =.065, 95% CI [.025, .106], *t* = 3.194, *p* = .002. Furthermore, at the mean value of agreeableness traits, there is a significant positive relationship between rating accuracy and intelligence, *b* =.035, 95% CI [.008, .062], *t* = 2.584, *p* = .011. When agreeableness traits are high, there is a non-significant positive relationship between rating accuracy and intelligence, *b* =.005, 95% CI [-.033, .042], *t* =.246, *p* = .806. Importantly, these results indicate that the relationship between intelligence and rating accuracy only really emerges in judges with average or lower levels of agreeableness. In addition, the small to medium effect size (Cohen's f^2=.13) that we observed, supports Hypothesis 1.

Table 2

Moderator Test Results: Agreeableness as a Moderator of Effect between General Mental Ability and Accuracy

Predictors	*b* [95% CI]	*SE B*	*t*	*p*
Constant	1.025 [.903; 1.149]	.062	16.589	.000
General Mental Ability (Centred)	.035 [.008; .062]	.014	2.584	.011
Agreeableness (Centred)	.009 [-.019; .038]	.014	.637	.525
General Mental Ability x Agreeableness	-.007 [-.014;.-.000]	.003	-2.114	.037

Note. R^2= .11

4.5.2 Simple Slope Analysis: Agreeableness

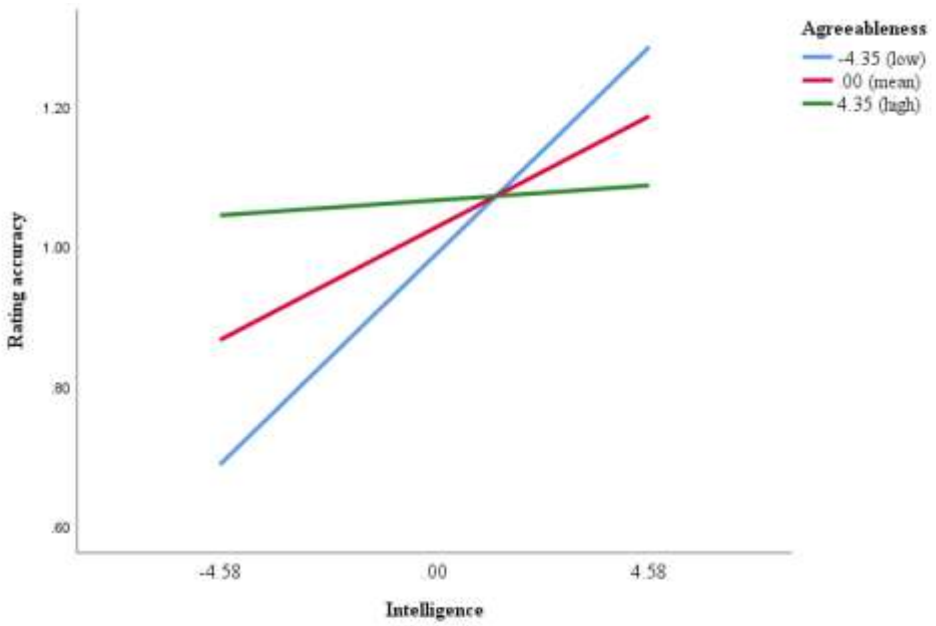
To interpret the moderation effect visually, one could examine the simple slopes (Field, 2018). The simple slope can be defined as the slope of the regression of Y on X at (conditional on) a single value of Z (Aiken et al., 1991). Importantly, the simple slope combines the regression coefficient of Y on X with the interaction coefficient (Aiken et al., 1991). Moreover, if the interaction is significant, this means that a moderation is present

(Field, 2018). If a moderating effect is present, one could follow up the analysis by using the simple slopes analysis, which examines the relationship between the predictor and the outcome at low, mean, and high levels of the moderator variable (Field, 2018). PROCESS uses 1 standard deviation above and below the mean value of the moderator (Field, 2018). The simple slope was generated practically using SPSS. Because the interaction variables were continuous, the variables had to be transformed using the Rank Cases function, in order to split the data into two groups. One could split the data into three or four groups; however, using two groups results in fewer complications (Field, 2018). The numbers in the key of the graph represent the relationship between the predictor and the outcome at low, mean, and high levels of the moderator variable.

The simple slope analysis (refer to Figure 4) illustrates that there is a significant strong positive relationship when agreeableness traits are low; at the mean level of agreeableness (red line), there is a significant a positive relationship between intelligence and rating accuracy, however, this relationship becomes non-significant and less positive at high levels of agreeableness (green line). Moreover, the fact that the lines cross indicates a significant interaction effect (Field, 2018).

Figure 4

Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Agreeableness



Moderation effect of openness to experience. Hypothesis 2 stated that the relationship between intelligence and rating accuracy is moderated by openness to experience, such that it is stronger for judges high in openness than for judges low in openness.

The results illustrated that moderation presents a non-significant interaction effect, *b*= -.005, 95% CI [-.012; .001], t = -1.624, p = .107, indicating that the relationship between intelligence and rating accuracy is not moderated by openness to experience. The interaction term for openness to experience explained only 2.2% of the variance in judgement accuracy after accounting for the main effects. Moreover, although statistically non-significant, based on Cohen's (1988) effect size conventions (.02 = small; .15 = medium; .35 = large), a small to medium effect size (Cohen's f^2 = .10) was observed. As such, no support for Hypothesis 2 was found. Table 3 reports the results of the MMR analysis for the model.

Table 3

Moderator Test Results: Openness to Experience as a Moderator of Effect between General Mental Ability and Accuracy

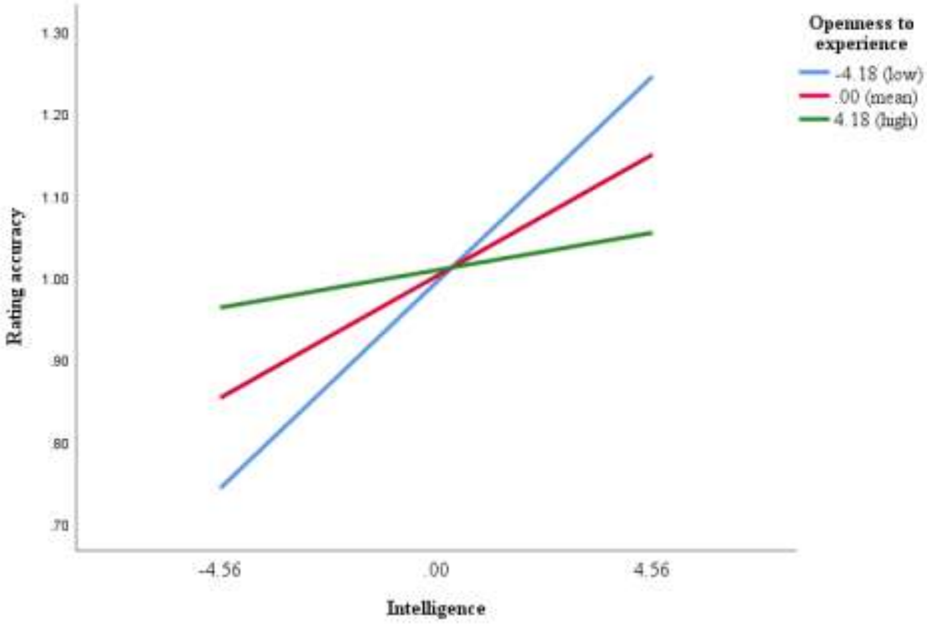
Predictors	*b* [95% CI]	*SE B*	*t*	*p*
Constant	1.001 [.884; 1.118]	.059	16.967	.000
General Mental Ability (Centred)	.033 [.006; .058]	.013	2.444	.016
Openness to experience (Centred)	.002 [-.018; .157]	.014	.125	.901
General Mental Ability x Openness to experience	-.005 [-.012; -.001]	.003	-1.624	.107

Note. R^2 = .09

To establish whether the direction of the interactions was consistent with the hypothesis, the simple slope analysis was used. Furthermore, the slopes (see Figure 5) indicate that when openness to experience is low (blue line), there is a non-significant positive relationship between intelligence and rating accuracy; at the mean level of openness to experience (red line), there is a strong positive relationship between intelligence and rating accuracy, however, this relationship gets weaker at high levels of openness to experience (green line). Additionally, the fact that the lines cross indicates a possible interaction effect (moderation) (Field, 2018). However, these results should be interpreted in conjunction with the statistical test for moderation, which showed that the effect was not statistically significant.

Figure 5

Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Openness to Experience



Moderation effect of conscientiousness. Hypothesis 3 stated that the relationship between intelligence and rating accuracy is moderated by conscientiousness, such that it is stronger for judges high in conscientiousness than for judges low in conscientiousness.

Results from the moderated multiple regression analysis (refer to Table 4) illustrate that moderation is shown up by a significant interaction effect; however, in this case the interaction is non-significant, b = -.004, 95% CI [-.011; .003], t = -1.188, p = .237, indicating that the relationship between intelligence and rating accuracy is not moderated by conscientiousness. The interaction term for conscientiousness explained only 1% of the variance in judgement accuracy after accounting for the main effects. Additionally, the small to medium effect size (Cohen's f^2 =.11) observed does not support Hypothesis 3.

Table 4

Moderator Test Results: Conscientiousness as a Moderator of Effect between General Mental Ability and Accuracy

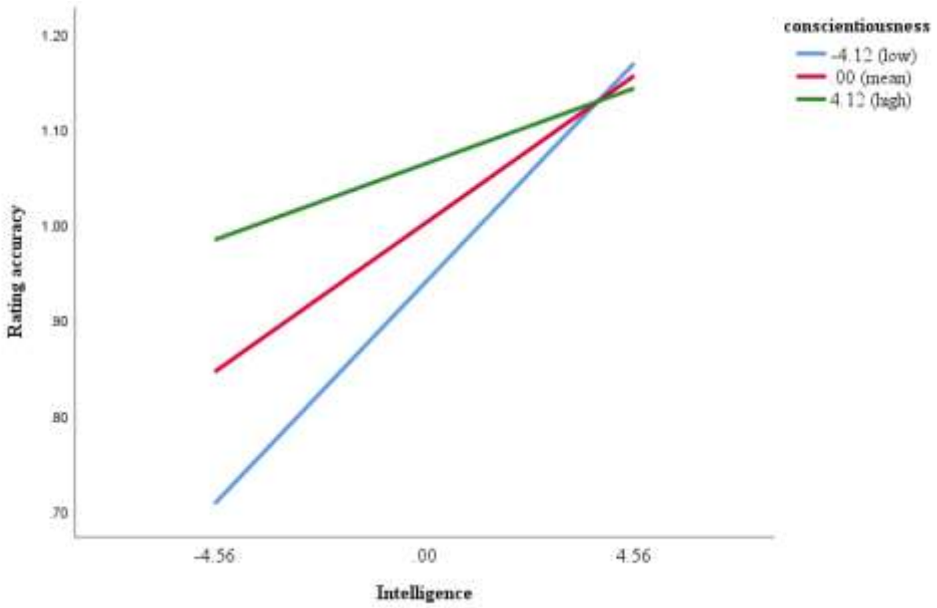
Predictors	*b* [95% CI]	*SE B*	*t*	*p*
Constant	1.001 [.883; 1.118]	.059	16.810	.000
General Mental Ability (Centred)	.034 [.008; .060]	.013	2.590	.011
Conscientiousness (Centred)	.015 [-.014; .044]	.015	1.042	.299
General Mental Ability x Conscientiousness	-.004 [-.011; -.003]	.003	-1.188	.237

Note. R^2 = .09

From the simple slope analysis (Figure 6), it is evident that when conscientiousness traits are low (blue line), there is a non-significant positive relationship between intelligence and rating accuracy; at the mean level of conscientiousness (red line), there is a positive relationship between intelligence and rating accuracy; however, this relationship becomes less positive and weaker at high levels of conscientiousness (green line). Additionally, the crossing of lines indicates a significant interaction effect (moderation) (Field, 2018).

Figure 6

Simple Slope Equations of the Regression of Intelligence on Rating Accuracy at Three Levels of Conscientiousness



4.6 Summary of Moderation Results

The findings of this study reported that rater-personality traits generally do not seem to act as moderators, but they may for more isolated traits. Agreeableness was the only trait out of the three personality traits analysed which was found to be a significant moderator of the relationship between intelligence and rating accuracy, with a small to moderate effect size.

4.7 Further Analyses

Statistical power. Research has demonstrated that several interaction effects hypothesised on the basis of strong theoretical frameworks are generally small, non-significant or not easily replicated (Marsh et al., 2013). Part of this concern may be due to the fact that tests of interaction effects commonly lack power so that meaningfully large interaction effects are not statistically significant (Marsh et al., 2013). Therefore, an additional analysis was performed to test the power of the interaction effects in this study for the first two regression models. These two models were chosen as their results indicated a non-significant interaction effect; it is therefore necessary to further investigate whether this may be due to a lack of statistical power.

Power refers to the possibility of obtaining a statistically significant outcome (Cohen, 1988). A post hoc power analysis was performed using GPower (Faul et al., 2009). A sample size of 146 participants with a three-predictor variable equation was used for each test. The recommended effect sizes used for this analysis were: small ($f^2 = .02$), medium ($f^2 = .15$), and large ($f^2 = .35$) (Cohen, 1988). Additionally, the alpha level used for this analysis was $p = .05$. The first regression model which examined whether openness to experience moderated the relationship between intelligence and rating accuracy, demonstrated an acceptable power value of .93 (Faul et al., 2009), indicating that there is a 93% chance of rejecting the null hypothesis with an effect size of $f^2 =.10$. The second regression model which examined whether conscientiousness moderated the relationship between intelligence and rating accuracy, also demonstrated an acceptable power of .93, with a small to medium effect size of $f^2 =.10$, indicating that there is a 93% chance of rejecting the null hypothesis. In sum, a lack of statistical power did not seem to represent a plausible rival explanation for the null results.

Chapter 5: Discussion

5.1 Introduction

The following chapter will discuss the aim and findings of the present study in more detail. Specifically, three MMR models were examined, namely the moderating effect of openness to experience on the relationship between intelligence and rating accuracy, the moderating effect of conscientiousness on the relationship between intelligence and rating accuracy, and lastly, the moderating effect of agreeableness on the relationship between intelligence and rating accuracy. Implications for practice as well as suggestions for future research are discussed in this chapter.

The main aim of this research study was to analyse whether rater-personality traits moderate the relationship between intelligence and rating accuracy. Current research regarding individual differences relating to accuracy has failed to produce consistent trends. Therefore, a possible explanation for these mixed findings could be due to moderating effects. The present study extended the research of Christiansen et al. (2005) by demonstrating how personality may moderate GMA on rating accuracy. Individual differences are related to key judgement processes, namely cue detection and cue utilisation thought to cause accuracy (De Kock et al., 2020; Funder, 1999). Literature has indicated that theoretically, there are specific individual constructs which may facilitate better cue detection and utilisation. For example, when a judge is highly conscientious, they are expected to have the ability to effectively integrate and interpret cues due to their high cognitive abilities as they are able to recall and manipulate cues successfully, allowing them to make more accurate judgements. Therefore, intelligence might matter for accuracy only for intellectual people (high on conscientiousness). Additionally, this study aimed to add to the good judge model (De Kock et al., 2020) by analysing how individual differences interact indirectly, influencing accuracy.

5.2 Main Findings

Firstly, the findings of this study demonstrated that rater-personality traits do in fact moderate the relationship between intelligence and rating accuracy; however, only for certain personality traits. De Kock et al.’s (2020) meta-analysis proposed that there might be more complex interactions between personality and judgment accuracy, such as moderating effects which may account for these inconsistent trends. Moreover, the results of this study illustrated that the moderating effect of agreeableness on the relationship between intelligence and rating accuracy was supported, indicating that the relationship between intelligence and rating accuracy only really emerges with lower levels of agreeableness in judges, which is supported by Christiansen et al. (2005).

Theoretically, this indicates that individual characteristics in the good judge model (De Kock et al., 2020) are also linked together in how they affect accuracy. The good judge model (De Kock et al., 2020) illustrates only the direct effects of individual differences and not yet any interactions, therefore the model may require some revision to show these moderating effects, if possible. A possible explanation for why intelligence might only matter for difficult people is that individuals who have the propensity for being difficult are known to maintain greater levels of intelligence (Baker & Bichsel, 2006). They require high IQ to do proper cue detection and cue utilisation, which may perhaps compensate for having less valid cues due to poor rapport or poor relationships.

Secondly, because only one of the three personality traits presented a significant moderating effect, it may be possible that personality has a moderating effect on the relationship between intelligence and rating accuracy; however, only for certain traits and not for others. Therefore, we need replications of the findings for agreeableness, as well as other studies to examine whether the findings are robust. This study is the first to investigate the replicability of the moderator effects in Christiansen et al. (2005) and contributed to this limited research area in various ways. By studying replication, researchers build a stronger science, increasing the validity of the study (Farrar et al., 2020).

Thirdly, the study demonstrated that even though MMR is known to suffer from low statistical power, it is not always the case if the correct methodological techniques are used.

5.2.1 Agreeableness

This study examined the moderating effects of personality traits on the relationship between intelligence and accuracy. The findings demonstrated that out of the three regression models test, only one model showed a significant moderating effect. Importantly, the results of this study illustrated that the moderating effect of agreeableness on the relationship between intelligence and rating accuracy was supported, indicating that the relationship between intelligence and rating accuracy only really emerges with lower levels of agreeableness in judges.

The fact that a moderator effect is present may suggest a boundary effect (De Kock et al., 2020), as intelligence might only matter for accuracy for difficult individuals (low in agreeableness). To establish practically why difficult people (low on agreeableness) require more intelligence in order to be accurate compared to those high in agreeableness who do not really take intelligence into account, one could refer to the RAM model (Funder, 1999). For example, difficult people require high IQ to do proper cue detection and cue utilisation, which may perhaps compensate for having less valid cues due to poor rapport or poor relationships. The findings may also propose different mechanisms to achieve accuracy; perhaps highly agreeable individuals may also be accurate as they get along better with people and are able to establish good relationships, allowing individuals to open up to them. In contrast, difficult people may have a more challenging time achieving accuracy and therefore require higher IQ to perform proper cue detection and cue utilisation.

Additionally, findings from this study are consistent with past research (Hollenbeck et al., 1988; Wright et al., 1995), demonstrating that performance on specific tasks is more efficient when people are motivated to perform well on the job and have greater levels of relevant abilities. Additionally, Wright et al.'s (1995) study support the interactive findings of Hollenbeck et al. (1988), who propose that personality and cognitive ability interactively predict job performance.

5.2.2 Openness to Experience

The second model examined whether openness to experience moderated the relationship between intelligence and rating accuracy. Results from this model were consistent with Christiansen et al. (2005), who found that judges with higher levels of openness to experience were no more accurate judges than those who were lower on these traits. Christiansen et al. (2005) are the only researchers to date to have investigated the interaction effects of personality on rating accuracy and intelligence. Since no moderating effect was detected, it suggests that intelligence does not matter for accuracy, whether the individual presents high or low levels of openness to experience.

One possible explanation could be due to the fact that openness to experience is fundamentally related to general intelligence as they share very similar traits (DeYoung et al., 2014). Based on current research relating to personality and intelligence, general intelligence tends to overlap with openness to experience (Silvia & Sanders, 2010). Some studies have indicated a positive correlation between curiosity (a characteristic of openness to experience) and intelligence (Silvia & Sanders, 2010).

To establish why openness to experience does not really matter for intelligent individuals, one could refer to cue utilisation and cue detection in the RAM model (Funder, 1999). Openness to experience traits, such as intellectance, correlate positively with rating accuracy as it facilitates effective behaviour information processing, a fundamental element in cue utilisation (De Kock et al., 2020). Because general intelligence and openness to experience both deal with the ability to deal with complex tasks, intelligence may not matter for individuals who present higher or lower levels of openness to experience. Curious or inquisitive people may be accurate even if they are not intelligent, as they have the ability to understand complex tasks and concepts, solve problems, and comprehend new ideas. As a result they have more cues at their disposal (because they have the ability to understand people, process knowledge and come with creative solutions). Therefore, they may have other beneficial cues available to them, which may help them to achieve accurate judgement.

5.2.3 Conscientiousness

The third regression model examined the moderating effect of Conscientiousness on the relationship between intelligence and rating accuracy. However, the results from this study were not entirely consistent with Christiansen et al. (2005). Moreover, their findings propose a more complex relationship between conscientiousness and accuracy, with intelligence being a better predictor of judgment accuracy when the judge was more conscientious (Christiansen et al., 2005).

It is anticipated that in cue detection, judges who are highly conscientious are predicted to be more observant compared to low conscientious judges and also demonstrate more consistency in cue utilisation (De Kock et al., 2020). Therefore, conscientious individuals may be accurate even if they are not intelligent, as they have competencies such as the ability to remain focused during long interviews or pay attention to specific details, which facilitate better cue detection, allowing them to make accurate judgements. Consequently, being smart may not be a significant factor for such individuals as they utilise different cues which help them to achieve accuracy.

Another possible explanation for no moderating effect could be due to the different types of samples used. This particular study made use of police managers with a mean age of 44 years. Moreover, Christiansen et al. (2005) used a sample of university undergraduate students and results illustrate a significant moderating effect. Similarly, Ziegler et al. (2009) used a sample of psychology students enrolled at a German university. Results from this study support the moderating effect of conscientiousness on the relationship between intelligence and GPA (school performance) (Ziegler et al., 2009). Generally, their results demonstrate that higher levels of conscientiousness results in higher motivation to work dutifully, present self-discipline, feel capable and avoid being impulsive (Ziegler et al., 2009). Additionally, due to conscientiousness's link to a variety of educational achievements and association with volition (hard-working and persevering), it has also been referred to as the will to achieve and work (Barrick & Mount, 1991). Both studies made use of a student population.

Furthermore, the mixed findings in this study could also be due to job complexity. Individuals in Wright et al. (1995) study illustrated low complexity jobs (academic environments), whereas the sample in Mount et al. (1999) were from higher complexity environments (managers and sale representatives) who found no interaction effect of conscientiousness and GMA on job performance. Additionally, the sample size in certain studies was also relatively small, which resulted in weak statistical power (Hollenbeck et al., 1988), making it difficult to observe an interaction or the interaction could be spurious.

Importantly, another factor which may explain these mixed findings could be due to the use of construct valid measures of conscientiousness (Mount et al., 1999). Some studies made use of facets of conscientiousness, such as persistence, while others used conscientiousness as a broad trait (Mount et al., 1999; Wright et al., 1995; Ziegler et al., 2009). Even though studies have been centred on the broad trait concept of conscientiousness, recent literature has proposed that the facets of conscientiousness are distinctly correlated to academic performance, with achievement-oriented facets being the strongest predictor in university climates (Noftle & Robins, 2007; Paunonen & Ashton, 2013).

Scale coarseness is another factor to consider (Aguinis & Gottfredson, 2010). Statistical power may be enhanced when scale measures are less coarse (Aguinis, Pierce et al., 2009). Even though this fact is rarely recognised, organisational science researchers make use of coarse scales every time continuous variables are measured using Likert-type scales (Aguinis, Pierce et al., 2009). This particular study made use of continuous scales when measuring continuous underlying variables, which may have enhanced each regression model's statistical power. Moreover, in terms of measurement errors, low reliability can have a truncating influence on moderating effects (Cohen, 1988), which may have resulted in the mixed findings; however, both GMA and personality measures produced acceptable reliability estimates.

5.3 Moderated Multiple Regression and Statistical Power

MMR has been commonly known to suffer from low statistical power (Aguinis & Gottfredson, 2010), which suggests that the probability of detecting interaction effects may be low, resulting in insignificant findings. In any research study, statistical power plays a pertinent role and is required to be adequate (Barrick & Mount, 1991). Importantly, this is a primary concern when attempting to observe interaction effects, because more subjects are needed compared to when attempting to observe main effects (Aguinis & Gottfredson, 2010).

The power of observing an interaction of moderate magnitude with an alpha of .05 was sufficient for all three models tested in this study. However, in this study the regression models of the moderating effects of openness to experience and conscientiousness on the relationship between intelligence and rating accuracy both yielded a power value of .93 (Faul et al., 2009), with a medium effect size of f^2=.10, and no moderation was detected. Therefore, weak statistical power is not an acceptable explanation for the failure to observe a moderating effect.

Most research has indicated that when an insignificant interaction is observed, it is necessary to consider design, measurement, as well as analysis issues to increase the probability of finding an interaction effect if it is present (Aguinis & Gottfredson, 2010). This study met most of the conditions detrimental to statistical power in MMR (Aguinis & Gottfredson, 2010). Earlier studies that failed to detect moderating effects did not highlight that low statistical power may have led to a failure to detect interaction effects, even though it may have been present. The present study therefore shed light on this concern, and found that in conditions with sufficient power, generally traits do not seem to moderate the effect of GMA on accuracy, except for possible individual exceptions like agreeableness in this study.

5.4 Limitations

The aim of this study was to examine the moderating effect of rater-personality traits on the relationship between intelligence and rating accuracy. The moderating effect of personality traits on rating accuracy is still a theoretically underdeveloped notion and this study aimed to contribute to this growing field of research. The following section highlights the limitations identified in this research study.

This study made use of secondary data and there are various limitations associated with using this particular type of data source. Secondary data analysis involves data that was collected by a researcher with the intention of answering their primary aim (Johnston, 2017). Because the data were not collected to answer this current study's research question, there may have been some information which could have been useful which was not collected. In terms of the data provided, one has to assume data were collected accurately with no intentional manipulation (Tripathy, 2013). Another limitation of using this type of data source is that the researcher does not know exactly how the data collection procedure was conducted and therefore the secondary researcher may not know if the data were affected by issues, such as low response rates or participants misinterpreting certain survey questions (Johnston, 2017). Additionally, participation in the original study was voluntary, which may have resulted in sample selection bias. Furthermore, measures used to evaluate the interviewer's personality were self-report questionnaires; using questionnaires to collect data increases the risk of social desirability bias and several other types of response bias. However, because responses remained anonymous, one could presume this may have accounted for some of the response bias concerns (Tripathy, 2013).

A cross-sectional research design was implemented, which presents some limitations. Adopting a cross-sectional design limits the researcher to establishing cause-and-effect relationships, as it only establishes associations (Sedgwick, 2014). Future research should consider adopting a longitudinal approach to investigate possible cause and effect (Field, 2018). This study could benefit from using a qualitative approach to explain the significance of the results based on unique contexts. In several cases, small effect sizes may be very valuable for practice, and in contrast, large effect sizes may not be very valuable and influential in certain conditions. Therefore, adopting a qualitative approach could add value to the moderating effects field by providing a more in-depth understanding and critically explaining the relationship in practical terms when dealing with effect sizes (Aguinis & Gottfredson, 2010; Schilling, 2015).

Another limitation of this study may be that the original study did not make use of "actual life" interviews; instead videotaped interviewees were used to account for experimental control (De Kock et al., 2015). Moreover, by using standardised videotaped stimuli in a controlled setting, the primary researchers were able to hold performance constant and reduce error variance associated with real life interviewees (De Kock et al., 2015). Importantly, adopting a standard stimuli technique enabled the researchers to establish true scores for these performances, which might not have been possible in a true field environment. To alleviate the loss of fidelity, the primary researchers tried to create realistic conditions for the participants. All interviewees were in that stage of their lives where they were searching for jobs and used the interview activity to prepare for real life selection conditions. Most of the participants who provided feedback stated that the interviews felt relatively realistic. Furthermore, future research should examine the moderating effect of personality traits on the relationship between intelligence and rating accuracy by interviewing real interviewees.

5.5 Recommendations for Future Research

Various limitations relating to the use of secondary data give rise to recommendations for future research. According to De Kock et al. (2015), caution should be taken regarding overgeneralising the results from a single organisation. Future research could also assess the degree to which the results of this study generalise to other businesses, as results may differ across various other jobs and contexts, therefore replication across diverse samples is necessary. De Kock et al. (2015) decided to expand the generalisability of dispositional reasoning studies to non-student samples; this may have resulted in some limitations because Schmid Mast et al. (2011) found that they are not that different in terms of accuracy. However, future research should consider assessing whether the results would generalise to other types of judges (e.g. psychologists). In addition, because only one of the three personality traits presented a significant moderating effect, it may be possible that personality has a moderating effect on the relationship between intelligence and rating accuracy; however, only for certain traits and not others. The present study only examined three out of the five Big Five traits, therefore future research should replicate the findings of this study, as well as other traits which this study did not include.

Reporting interaction effect size estimates, such as f^2 and R^2, may not always produce information and knowledge on the practical importance of a given effect (Aguinis & Gottfredson, 2010). In several cases, small effect sizes may be very valuable for practice, and in contrast, large effect sizes may not be very valuable and influential in certain conditions (Aguinis, Werner et al., 2009). It is therefore recommended that researchers adopt a customer-centric technique when analysing the findings, which includes performing a qualitative approach that explains the significance of the findings based on unique contexts (Aguinis, Werner et al., 2009). In summary, one should not only report the statistical significance or non-significance of a moderating effect, but in addition to that, critically explain the nature and magnitude of the moderating effect in practical terms (Aguinis & Gottfredson, 2010).

5.6 Implications for Theory and Future Research

This study contributed to literature on the moderating effects of personality traits on rating accuracy and intelligence in interviews by replicating and expanding on the previous research by Christiansen et al. (2005) and De Kock et al. (2015). The findings have implications for issues which are crucial for theory building, as well as research on individual differences in rating accuracy. Overall, the results of this study support mainstream theories of judgement accuracy which propose that judges could be fundamental moderators of accuracy (De Kock et al. 2015; Funder 1995, 2012). Specifically, accuracy derives in part from individual differences in their ability to make use of behavioural cues, as posited by Funder (1995). While Christiansen et al. (2005) demonstrated that agreeableness and conscientiousness moderated the relationship between dispositional intelligence and acquaintance accuracy, this study was able to constructively replicate their findings and offer a more concrete representation of whether their findings were spurious or robust. The present study also extended the good judge model (De Kock et al., 2020), illustrating that individual characteristics in the framework are also linked together in how they affect accuracy. Although only one of the hypotheses was supported in this study, this provides interesting evidence into how design, measurement and analysis techniques play a fundamental role when examining moderator variables.

The power of observing an interaction of moderate magnitude with an alpha of .05 was sufficient for all three models tested in this study. Therefore, weak statistical power in this case was not considered an acceptable explanation for failure to observe an interaction.

One might need to investigate the design of the study, as well as data analysis when detecting interaction effects. It is therefore recommended that future studies using MMR should routinely plan for sufficient power and report the findings to develop a more robust argument. Furthermore, another crucial factor present in this study is the use of valid personality constructs. For example, this study made use of conscientiousness as a broad personality trait, whereas certain studies have made use of personality constructs which aimed to be elements of conscientiousness but failed to provide construct validity data (Barrick & Mount, 1991). Evidently, this may be a possible explanation for the mixed findings in previous studies.

5.7 Implications for Practice

The findings of this study may be pertinent to human resource staffing and recruitment as accurate personality judgements contribute to overall organisational performance. Given that the effects observed for the moderating hypotheses were generally non-existing (or small to moderate), we do not anticipate these to translate to practically significant conditions for the workplace application. However, for the agreeableness trait, the results could provide practical implications for interviewer screening and training. The findings may suggest that for agreeableness, perhaps rater training and screening may approach people low or high in agreeableness differently. One group may need higher levels of GMA to achieve accuracy, whereas those high in agreeableness may not need much intelligence after all as they achieve accuracy despite their intelligence levels. In addition, another possible efficient training technique lies in interventions which are based around each component (De Kock et al., 2015), for example individuals who are agreeable or conscientious.

5.8 Conclusion

Individuals differ in their ability to make interpersonal judgements; some may be more accurate than others, and these differences could be described by a mixture of cognitive as well as personality factors. This research study expanded on the research of Christiansen et al. (2005) on moderating effects by considering three personality traits and how they moderate the relationship between intelligence and rating accuracy. The findings of this study have extended support for personality traits being a moderator variable influencing the relationship between intelligence and rating accuracy.

Importantly, the findings indicated that agreeableness moderated the relationship between intelligence and rating accuracy in interviews. Contrary to expectations, openness to experience and conscientiousness did not moderate the relationship between intelligence and rating accuracy. It may be possible that personality has a moderating effect on the relationship between intelligence and rating accuracy, but only for certain traits and not others. Therefore, we need replications of the findings for agreeableness, as well as other studies to examine whether the findings are robust. However, upon further analysis, the power of the interaction effects between the regression models which indicated a non-significant relationship demonstrated moderately high power values of .93. This may be important aspects to consider when analysing moderation effects. The present study shed light on this concern, and found that in conditions with sufficient power, generally traits do not seem to moderate the effect of GMA on accuracy, except for possible individual exceptions like agreeableness in this study. Additionally, this study may be relevant to rating practices, as the success of rater selection, as well as rater training, could be dependent on a valid model of the specific interviewer constructs that drive accuracy.

References

Adams, H. F. (1927). The good judge of personality. *The Journal of Abnormal and Social Psychology, 22*(2), 172-181. https://doi.org/10.1037/h0075237.

Aguinis, H. (2004). *Regression Analysis for Categorical Moderators*. Guilford Press.

Aguinis, H., & Gottfredson, R. K. (2010). Best-practice recommendations for estimating interaction effects using moderated multiple regression. *Journal of Organizational Behavior*, *31*(6), 776-786. https://doi/org/10.1002/job.686

Aguinis, H., Boik, R. J., & Pierce, C. A. (2001). A generalized solution for approximating the power to detect effects of categorical moderator variables using multiple regression. *Organizational Research Methods, 4*(4), 291-323. https://doi.org/10.1177%2F109442810144001

Aguinis, H., Edwards, J. R., & Bradley, K. J. (2017). Improving our understanding of moderation and mediation in strategic management research. *Organizational Research Methods*, *20*(4), 665-685. https://doi.org/10.1177%2F1094428115627498

Aguinis, H., Pierce, C. A., & Culpepper, S. A. (2009). Scale coarseness as a methodological artifact: Correcting correlation coefficients attenuated from using coarse scales. *Organizational Research Methods*, *12*(4), 623-652. https://doi.org/10.1177/1094428108318065

Aguinis, H., Werner, S., Lanza Abbott, J., Angert, C., Park, J. H., & Kohlhausen, D. (2009). Customer-centric science: Reporting significant research results with rigor, relevance, and practical impact in mind. *Organizational Research Methods*, *13*(3), 515-539. https://doi.org/10.1177%2F1094428109333339

Aiken, L. S., West, S. G., & Pitts, S. C. (2003). Multiple linear regression. In L. B. Weiner (ed.), *Handbook of Psychology* (pp. 481-507). Wiley. https://doi.org/10.1002/0471264385.wei0219

Aiken, L. S., West, S. G., & Reno, R. R. (1991). *Multiple Regression: Testing and Interpreting Interactions*. Sage.

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin 111*, 256–274. https://doi.org/10.1037/0033-2909.111.2.256.

Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology*, *69*(3), 518. https://doi.org/10.1037/0022-3514.69.3.518.

Andersson, U., Cuervo-Cazurra, A., & Nielsen, B. B. (2020). Explaining interaction effects within and across levels of analysis. In *Research Methods in International Business* (pp. 331-349). Palgrave Macmillan.

Babbie, E., & Mouton, J. (2006). *The Practice of Social Research* (SA Edition). Oxford Press.

Baker, T. J., & Bichsel, J. (2006). Personality predictors of intelligence: Differences between young and cognitively healthy older adults. *Personality and Individual Differences*, *41*(5), 861-871. http://dx.doi.org/10.1016/j.paid.2006.02.017

Baron, R. M., & Kenny, D. A. (1986). The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, *51*(6), 1173-1182. https://doi.org/10.1037/0022-3514.51.6.1173

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1-26. https://doi.org/10.1111/j.1744-6570.1991.tb00688.x

Barrick, M. R., Patton, G. K., & Haugland, S. N. (2000). Accuracy of interviewer judgments of job applicant personality traits. *Personnel Psychology*, *53*, 925–951. https://doi.org/10.1111/j.1744-6570.2000.tb02424.x

Bergold, S., & Steinmayr, R. (2018). Personality and intelligence interact in the prediction of academic achievement. *Journal of Intelligence*, *6*(2), 1-27. http://dx.doi.org/10.3390/jintelligence6020027

Borman, W. C. (1979). Individual difference correlates of rating accuracy using behavior scales. *Applied Psychological Measurement*, *3*, 103–115. https://doi.org/10.1177/014662167900300111

Borman, W.C., & Hallam, G.L., 1991. Observation accuracy for assessors of work-sample performance: Consistency across task and individual-differences correlates. *Journal of Applied Psychology*, *76*, 11–18. https://doi.org/10.1037/0021-9010.76.1.11.

Brunswik, E. (1956). *Perception and the Representative Design of Psychological Experiments*. University of California Press.

Byrne, Z. S., Stoner, J., Thompson, K. R., & Hochwarter, W. (2005). The interactive effects of conscientiousness, work effort, and psychological climate on job performance. *Journal of Vocational Behavior, 66*(2), 326-338. https://doi.org/10.1016/j.jvb.2004.08.005

Chamorro-Premuzic, T., & Arteche, A. (2008). Intellectual competence and academic performance: Preliminary validation of a model. *Intelligence*, *36*(6), 564-573. https://doi.org/10.1016/j.intell.2008.01.001

Chmura Kraemer, H., Kiernan, M., Essex, M., & Kupfer, D. J. (2008). How and why criteria defining moderators and mediators differ between the Baron & Kenny and MacArthur approaches. *Health Psychology*, *27*(2S), S101. https://doi.org/10.1037/0278-6133.27.2(Suppl.).S101

Christiansen, N. D., Wolcott-Burnam, S., Janovics, J. E., Burns, G. N., & Quirk, S. W. (2005). The good judge revisited: Individual differences in the accuracy of personality judgments. *Human Performance*, *18*(2), 123-149. https://doi.org/10.1207/s15327043hup1802_2.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. (2nd ed.).Lawrence Erlbaum Associates.

Cooksey, R. W. (1996). The methodology of social judgement theory. *Thinking & Reasoning*, *2*(2-3), 141-174. https://doi.org/10.1080/135467896394483

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). Sage.

Cronbach, L. J. (1955). Processes affecting scores on" understanding of others" and" assumed similarity". *Psychological Bulletin*, *52*(3), 177-193. https://doi.org/10.1037/h0044919.

Davis, M. H., & Kraus, L. A. (1997). Personality and empathic accuracy. In W. Ickes, (Ed.), *Empathic Accuracy* (pp. 144–168). Guilford Press.

Dawson, J. F. (2014). Moderation in management research: What, why, when, and How. *Journal of Business Psychology*, *29*, 1–19. https://doi.org/10.1007/s10869-013-9308-7

De Kock, F. S., Lievens, F., & Born, M. P. (2015). An in-depth look at dispositional reasoning and interviewer accuracy. *Human Performance*, *28*(3), 199-221. https://doi.org/10.1080/08959285.2015.1021046

De Kock, F. S., Lievens, F., & Born, M. P. (2020). The profile of the 'Good Judge' in HRM: A systematic review and agenda for future research. *Human Resource Management Review,* Art. 100667. https://doi.org/10.1016/j.hrmr.2018.09.003

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish mental survey. *Intelligence*, *28*, 49–55.

DeYoung, C. G., Quilty, L. C., Peterson, J. B., & Gray, J. R. (2014). Openness to experience, intellect, and cognitive ability. *Journal of Personality Assessment*, *96*(1), 46-52. https://doi.org/10.1080/00223891.2013.806327

Di Domenico, S. I., & Fournier, M. A. (2015). Able, ready, and willing: Examining the additive and interactive effects of intelligence, conscientiousness, and autonomous motivation on undergraduate academic performance. *Learning and Individual Differences*, *40*, 156-162. https://doi.org/10.1016/j.lindif.2015.03.016

Digman, J.M. (1990) Personality structure: Emergence of the five-factor model. *Annual Review of Psychology*, *41*, 417–440. https://doi.org/10.1146/annurev.ps.41.020190.002221.

Donnellan, M. B., Baird, B. M., Lucas, R. E., & Oswald, F. L. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the big five factors of personality. *Psychological Assessment*, *18,* 192-203. https://doi.org/10.1037/1040-3590.18.2.192

Dowding, D., & Thompson, C. (2003). Measuring the quality of judgement and decision-making in nursing. *Journal of Advanced Nursing*, *44*(1), 49-57. https://doi.org/10.1046/j.1365-2648.2003.02770.x

Egboro, F. O. (2015). The implications of parametric and non-parametric statistics in data analysis in marketing research. *International Journal of Humanities and Social Science*, *5*(6), 74-83.

Farrar, B. G., Boeckle, M., & Clayton, N. S. (2020). Replications in comparative cognition: What should we expect and how can we improve. *Animal Behavior and Cognition*, *7*(1), 1-22. https://doi.org/10.26451/abc.07.01.02.2020

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/BRM.41.4.1149

Field, A. (2018). *Discovering Statistics using IBM SPSS Statistics.* (5th ed.). Sage.

Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science, 21,* 177–182. https://doi.org/10.1177/0963721412445309.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*, 652–670. https://doi.org/10.1037/0033-295X.102.4.652.

Funder, D. C. (1999). *Personality Judgment: A Realistic Approach to Person Perception.* Academic Press.

Funder, D. C, & West, S. G. (1993). Consensus, self-other agreement and accuracy in personality judgment: An introduction. *Journal of Personality*, 61, 457-476. https://doi.org/10.1111/j.1467-6494.1993.tb00778.x

Funder, D. C., Letzring, T. D. & Wells, S. M. (2004). Information quantity and quality affect the realistic accuracy of personality judgment. *Journal of Personality and Social Psychology*, *91*(1), 111.

Furnham, A., Dissou, G., Sloan, P., & Chamorro-Premuzic, T. (2007). Personality and intelligence in business people: A study of two personality and two intelligence measures. *Journal of Business and Psychology*, *22*(1), 99-109. https://doi.org/10.1007/s10869-007-9051-z

Gardner, R. G., Harris, T. B., Li, N., Kirkman, B. L., & Mathieu, J. E. (2017). Understanding "it depends" in organizational research: A theory-based taxonomy, review, and future research agenda concerning interactive and quadratic relationships. *Organizational Research Methods*, *20*(4), 610-638. https:///doi.org/10.1177/1094428117708856

Goldberg, L. R. (1992). The development of markers for the big-five factor structure. *Psychological Assessment*, *4*, 26-42. https://doi.org/10.1037/1040-3590.4.1.26.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, &F. Ostendorf (Eds.), *Personality Psychology in Europe,* Vol. 7 (pp. 7–28). Tilburg University Press.

Goodhue, D., Lewis, W., & Thompson, R. (2007). Research note: Statistical power in analyzing interaction effects: Questioning the advantage of PLS with product indicators. *Information Systems Research*, *18*(2), 211-227. https://www.jstor.org/stable/23015403

Greenhoot, A. F., & Dowsett, C. J. (2012). Secondary data analysis: An important tool for addressing developmental questions. *Journal of Cognition and Development*, *13*(1), 2-18. http://dx.doi.org/10.1080/15248372.2012.646613

Guion, R. M., & Highhouse, S. (2011). *Essentials of Personnel Assessment and Selection.* Routledge.

Harries, P. A., & Harries, C. (2001). Studying clinical reasoning, part 2: Applying social judgement theory. *British Journal of Occupational Therapy*, *64*(6), 285-292. https://doi.org/10.1177/030802260106400604

Hauenstein, N. M. A., & Alexander, R. A. (1991). Rating ability in performance judgments: The joint influence of implicit theories and intelligence. *Organizational Behavior and Human Decision Processes, 50*, 300–323. https://doi.org/10.1016/0749-5978(91)90024-N

Hayes, A. F. (2009). Beyond Baron and Kenny: Statistical mediation analysis in the new millennium. *Communication Monographs*, *76*(4), 408-420. https://doi.org/10.1080/03637750903310360

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation, and conditional process modeling.[White Paper]. http://www.afhayes.com/public/process2012.pdf

Hayes, A. F., & Rockwood, N. J. (2017). Regression-based statistical mediation and moderation analysis in clinical research: Observations, recommendations, and implementation. *Behaviour Research and Therapy*, *98*, 39-57. https://doi.org/10.1016/j.brat.2016.11.001

Hollenbeck, J. R., Brief, A. P., Whitener, E. M., & Pauli, K. E. (1988). An empirical note on the interaction of personality and aptitude in personnel selection. *Journal of Management*, *14*, 441–451. https://doi.org/10.1177/014920638801400307

Huffcutt, A. I., Conway, J. M., Roth, P. L., & Stone, N. J. (2001). Identification and meta-analytic assessment of psychological constructs measured in employment interviews. *Journal of Applied Psychology, 86*, 897-913. https://doi.org/10.1037/0021-9010.86.5.897

Jensen, J. M., & Patel, P. C. (2011). Predicting counterproductive work behavior from the interaction of personality traits. *Personality and Individual Differences*, *51*(4), 466-471. https://doi.org/10.1016/j.paid.2011.04.016

John, O. P., Robins, R. W., & Pervin, L. A. (2008). *Handbook of Personality: Theory and Research* (3rd ed.). Guilford Press.

Johnson, P. O., & Neyman, J. (1936). Tests of certain linear hypotheses and their application to some educational problems. *Statistical Research Memoirs*.

Johnston, M. P. (2017). Secondary data analysis: A method of which the time has come. *Qualitative and Quantitative Methods in Libraries*, *3*(3), 619-626. http://www.qqml-journal.net/index.php/qqml/article/view/169

Jones, R.G., & Born, M.P., 2008. Assessor constructs in use as the missing component in validation of assessment center dimensions: A critique and directions for research. *International Journal of Selection and Assessment*, *16*, 229–238. https://doi.org/10.1111/j.1468-2389.2008.00429.x

Kenny, D. A., & Albright, L. (1987). Accuracy in interpersonal perception: a social relations analysis. *Psychological Bulletin*, *102*(3), 390-398. https://doi.org/10.1037/0033-2909.102.3.390

Khorramdel, L., & von Davier, M. (2014). Measuring response styles across the big five: A multiscale extension of an approach using multinomial processing trees. *Multivariate Behavioral Research*, *49*(2), 161-177. https://doi.org/10.1080/00273171.2013.866536

Kruglanski, A. W., & Ajzen, I. (1983). Bias and error in human judgment. *European Journal of Social Psychology*, *13*(1), 1-44. https://doi.org/10.1002/ejsp.2420130102.

LaFlair, G. T., Egbert, J., & Plonsky, L. (2015). A practical guide to bootstrapping descriptive statistics, correlations, t tests, and ANOVAs. *Advancing Quantitative Methods in Second Language Research* (pp. 46-77). Routledge.

Lance, C. E., Foster, M. R., Gentry, W. A., & Thoresen, J. D. (2004). Assessor cognitive processes in an operational assessment center. *Journal of Applied Psychology*, *89*, 22–35. https://doi.org/10.1037/0021-9010.89.1.22.

Letzring, T. D. (2008). The good judge of personality: Characteristics, behaviours, and observer accuracy. *Journal of Research in Personality*, *42*, 914-932. https://doi.org/10.1016/j.jrp.2007.12.003.

Li, M., Sharp, B. M., Bergh, D. D., & Vandenberg, R. (2019). Statistical and methodological myths and urban legends in strategic management research: The case of moderation analysis. *European Management Review*, *16*(1), 209-220. http://dx.doi.org/10.1111/emre.12319

Lippa, R. A., & Dietz, J. K. (2000). The relation of gender, personality, and intelligence to judges' accuracy in judging strangers' personality from brief video segments. *Journal of Nonverbal Behavior*, *24*(1), 25-43. https://doi.org/10.1023/A:1006610805385.

Maier, N. R. F. (1955). *Psychology in Industry*. (2nd ed.). Houghton-Mifflin.

Marsh, H. W., Hau, K.-T., Wen, Z., Nagengast, B., & Morin, A. J. S. (2013). *Moderation.* In T. D. Little (Ed.), *Oxford Library of Psychology. The Oxford Handbook of Quantitative Methods: Statistical Analysis* (pp. 361–386). Oxford University Press.

Mayer, J. D., Caruso, D. R., & Salovey, P. (1999). Emotional intelligence meets traditional standards for an intelligence. *Intelligence*, *27*(4), 267-298. https://doi.org/10.1016/S0160-2896(99)00016-1

Memon, M. A., Cheah, J. H., Ramayah, T., Ting, H., Chuah, F., & Cham, T. H. (2019). Moderation analysis: Issues and guidelines. *Journal of Applied Structural Equation Modeling*, *3*(1), 1-11.

Morgan, B., & De Bruin, K. (2010). The relationship between the big five personality traits and burnout in South African university students. *South African Journal of Psychology*, *40*(2), 182-191. https://doi.org/10.1177/008124631004000208

Mount, M. K., Barrick, M. R., & Strauss, J. P. (1999). The joint relationship of conscientiousness and ability with performance: Test of the interaction hypothesis. *Journal of Management*, *25*(5), 707-721. https://doi.org/10.1177/014920639902500505

Noftle, E. E., & Robins, R. W. (2007). Personality predictors of academic outcomes: Big five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, *93*(1), 116-120. https://doi.org/10.1037/0022-3514.93.1.116

Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory.* McGraw-Hill.

O’Connor, P. J., & Athota, V. S. (2013). The intervening role of agreeableness in the relationship between trait emotional intelligence and Machiavellianism: Reassessing the potential dark side of EI. *Personality and Individual Differences*, *55*(7), 750-754. https://doi.org/10.1016/j.paid.2013.06.006

Omair, A. (2015). Selecting the appropriate study design for your research: Descriptive study designs. *Journal of Health Specialties*, *3*(3), 153.

Osborne, J. W., & Waters, E. (2002). Four assumptions of multiple regression that researchers should always test. *Practical Assessment, Research, and Evaluation*, *8*(1), 1-5.

Pallant, J. (2011). *SPSS Survival Manual: A Step-by-step Guide to Data Analysis using IBM SPSS*. Routledge.

Paunonen, S. V., & Ashton, M. C. (2013). On the prediction of academic performance with personality traits: A replication study. *Journal of Research in Personality*, *47*(6), 778-781. https://doi.org/10.1016/j.jrp.2013.08.003

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design, and Analysis: An Integrated Approach.* Psychology Press.

Powell, D. M. (2008). *Assessing Personality in the Employment Interview: The Impact of Rater Training and Individual Differences on Rating Accuracy.* [Doctoral Thesis (NR39315), The University of Western Ontario].

Powell, D. M., & Goffin, R. D. (2009). Assessing personality in the employment interview: The impact of training on rater accuracy. *Human Performance*, *22*(5), 450-465. https://doi.org/10.1080/08959280903248450.

Rosopa, P. J., & Schroeder, A. N. (2009). Core self-evaluations interact with cognitive ability to predict academic achievement. *Personality and Individual Differences*, *47*(8), 1003-1006. https://doi.org/10.1016/j.paid.2009.07.028

Sackett, P. R., Gruys, M. L., & Ellingson, J. E. (1998). Ability-personality interactions when predicting job performance. *Journal of Applied Psychology*, *83*(4), 545-550. https://doi.org/10.1037/0021-9010.83.4.545

Salgado, J. F. (2017). Using ability tests in selection. In H. W. Goldstein, E. D. Pulakos, J. Passmore, & C. Semedo, (Eds.), *Wiley Blackwell Handbooks in Organizational Psychology. The Wiley Blackwell Handbook of the Psychology of Recruitment, Selection and Employee Retention* (pp. 115–150). Wiley Blackwell.

Schilling, E. (2015). ‘Success is satisfaction with what you have’? Biographical work–life balance of older female employees in public administration. *Gender, Work & Organization*, *22*(5), 474-494. https://doi.org/10.1111/gwao.12097

Schmid Mast, M., Bangerter, A., Bulliard, C., & Aerni, G. (2011). How accurate are recruiters' first impressions of applicants in employment interviews? *International Journal of Selection and Assessment*, *19*(2), 198-208. https://doi.org/10.1111/j.1468-2389.2011.00547.x

Schneider, D.E., & Bayroff, A.G. (1953). The relationship between rater characteristics and validity of ratings. *Journal of Applied Psychology 37*, 278–280. https://doi.org/10.1037/h0062458.

Sedgwick, P. (2014). Cross sectional studies: Advantages and disadvantages. *British Medical Journal, 348*, 2276-2277. https://doi.org/10.1136/bmj.g2276

Shevlin, M., Miles, J. N., Kulesza, M., Ewing, B., Shih, R. A., Tucker, J. S., & D'Amico, E. J. (2015). Moderated mediation analysis: An illustration using the association of gender with delinquency and mental health. *Journal of Criminal Psychology, 5*(2), 99–123.

Silvia, P. J., & Sanders, C. E. (2010). Why are smart people curious? Fluid intelligence, openness to experience, and interest. *Learning and Individual Differences*, *20*(3), 242-245. https://doi.org/10.1016/j.lindif.2010.01.006

Steinmayr, R., & Spinath, B. (2007). Predicting school achievement from motivation and personality. *Zeitschrift für Pädagogische Psychologie* (German Journal of Educational Psychology), *21*(3-4), 207–216. https://doi.org/10.1024/1010-0652.21.3.207

Sulsky, L. M., & Balzer, W. K. (1988). Meaning and measurement of performance rating accuracy: Some methodological and theoretical concerns. *Journal of Applied Psychology*, *73*(3), 497-506. https://doi.org/10.1037/0021-9010.73.3.497

Tabachnick, B. G., & Fidell, L. S. (2007). *Using Multivariate Statistics* (Vol. 5, pp. 481-498). Pearson.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using Multivariate Statistics*. International edition. Pearson.

Taft, R. (1955). The ability to judge people. *Psychological Bulletin, 52*, 1–23. https://doi.org/10.1037/h0044999

Taylor, N., & De Bruin, G. P. (2013). The Basic Traits Inventory. In S. Laher & K. Cockcroft (Eds.), *Psychological Assessment in South Africa: Research and Applications* (pp. 232–243). Wits University Press. https://doi.org/10.18772/22013015782.21

Tripathy, J. P. (2013). Secondary data analysis: Ethical issues and challenges. *Iranian Journal of Public Health, 42*(12), 1478-1479

Trzesniewski, K. H., Donnellan, M., & Lucas, R. E. (2011). *Secondary Data Analysis: An Introduction for Psychologists*. American Psychological Association.

Vernon, P. E. (1933). Some characteristics of the good judge of personality. *Journal of Social Psychology, 4*, 42–57. https://doi.org/10.1080/00224545.1933.9921556

Vogt, D. S., & Colvin, C. R. (2003). Interpersonal orientation and the accuracy of personality judgments. *Journal of Personality*, *71*, 267-295. https://doi.org/10.1111/1467-6494.7102005

Wagner, R. K. (2000). Practical intelligence. In *Handbook of Intelligence* (pp. 380-395). Cambridge University Press.

Witt, L. A., Burke, L. A., Barrick, M. R., & Mount, M. K. (2002). The interactive effects of conscientiousness and agreeableness on job performance. *Journal of Applied Psychology*, *87*(1), 164. https://doi.org/10.1037/0021-9010.87.1.164

Wright, P. M., Kacmar, K. M., McMahan, G. C., & Deleeuw, K. (1995). P= f (MXA): Cognitive ability as a moderator of the relationship between personality and job performance. *Journal of Management*, *21*(6), 1129-1139. https://doi.org/10.1177/014920639502100606

Wonderlic, E. F. (1998). *Manual of the Wonderlic Personnel Test & scholastic level exam*. Wonderlic Personnel Test, Inc.

Wonderlic, E. F. (2002). *Wonderlic Personnel Test & scholastic level exam user's manual*. Wonderlic Personnel Test, Inc.

Ziegler, M., Knogler, M., & Bühner, M. (2009). Conscientiousness, achievement striving, and intelligence as performance predictors in a sample of German psychology students: Always a linear relationship? *Learning and Individual Differences*, *19*(2), 288-292. https://doi.org/10.1016/j.lindif.2009.02.001

Zimprich, D., Allemand, M., & Dellenbach, M. (2009). Openness to experience, fluid intelligence, and crystallized intelligence in middle-aged and old adults. *Journal of Research in Personality*, *43*(3), 444-454. https://doi.org/10.1016/j.jrp.2009.01.018

Appendix A: A 20-item short version of the International Personality Item Pool Big Five Factor Markers

A 20-item short version of the International Personality Item Pool Big Five factor markers (Goldberg, 1992).

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to other people you know of the same sex as you are, and roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in absolute confidence. Indicate for each statement whether it is 1. Very Inaccurate, 2. Moderately Inaccurate, 3. Neither Accurate Nor Inaccurate, 4. Moderately Accurate, or 5. Very Accurate as a description of you.

	Very Inaccurate	Moderately Inaccurate	Neither Accurate Nor Inaccurate	Moderately Accurate	Very Accurate	
1. Am the life of the party.	O	O	O	O	O	(1+)
2. Feel little concern for others.	O	O	O	O	O	(2-)
3. Am always prepared.	O	O	O	O	O	(3+)
4. Get stressed out easily.	O	O	O	O	O	(4-)
5. Have a rich vocabulary.	O	O	O	O	O	(5+)
6. Don't talk a lot.	O	O	O	O	O	(1-)
7. Am interested in people.	O	O	O	O	O	(2+)
8. Leave my belongings around.	O	O	O	O	O	(3-)
9. Am relaxed most of the time.	O	O	O	O	O	(4+)
10. Have difficulty understanding abstract ideas.	O	O	O	O	O	(5-)
11. Feel comfortable around people.	O	O	O	O	O	(1+)
12. Insult people.	O	O	O	O	O	(2-)
13. Pay attention to details.	O	O	O	O	O	(3+)
14. Worry about things.	O	O	O	O	O	(4-)
15. Have a vivid imagination.	O	O	O	O	O	(5+)
16. Keep in the background.	O	O	O	O	O	(1-)
17. Sympathize with others' feelings.	O	O	O	O	O	(2+)

18. Make a mess of things.	O	O	O	O	O	(3-)
19. Seldom feel blue.	O	O	O	O	O	(4+)
20. Am not interested in abstract ideas.	O	O	O	O	O	(5-)