

**Predicting Personality from Trait- Relevant vs Trait- Irrelevant Questions using Trait
Activation Theory in Asynchronous Video Interviews**

Fleur van Kempen

SNR: 2042391

ANR: 817854

Tilburg University

Msc Work and Organizational Psychology

2023-2024

Supervisor: Antonios Koutsoumpis

April 7, 2024

Abstract

The current study investigates Trait Activation Theory's assertion that personality traits in job candidates are easier to measure in trait- relevant instead of trait- irrelevant interview questions. To examine this, 509 English- speaking adults from the United States, recruited via the Prolific panel, answered both trait- irrelevant and trait- relevant interview questions in a simulated Asynchronous Video Interview. Afterwards, their personality was predicted by trained raters and their spoken responses were analyzed using the Linguistic Inquiry and Word Count (LIWC) software. Multiple regression analyses were performed to analyze the correlations between verbal behavior and observer reports of personality. The results indicate that trait- relevant questions do not predict personality more accurately than trait- irrelevant questions do. Therefore, these findings suggest that both types of questions can be used to predict personality in asynchronous video interviews. However, future research is needed to increase the generalizability of the results beyond the U.S. population and analyze confounding variables.

Keywords: Trait Activation Theory, Asynchronous Video Interviews, Linguistic Inquiry and Word Count (LIWC), Personality Assessment.

Trait activation Theory is a theory on personality- job fit, stating that personality traits in humans can be activated by certain environmental cues, at the time when someone is exposed to cues that relate to these specific traits. An environment rich of these situational cues, leads to easy expression of a person's unique personality traits (Tett & Bunnett, 2003). Trait activation is the process of individuals expressing their traits when presented with trait-relevant situational cues (Tett & Bunnett, 2003). The theory argues in favor of situational specificity. Whether a trait predicts performance is dependent on the context. Therefore, the relevance of a trait and situation must correspond, so the trait enables someone to respond (Judge & Zapata, 2015). Therefore, a work environment in which people can frequently express their personality traits, is attractive to people (Simonet & Tett, 2013). An illustrative example could be an extrovert person working in customer service, where he or she has a lot of interaction with customers. Trait Activation Theory would suggest this person to show good work performance, since activation of the trait- relevant traits, meaning traits relating to Extraversion (such as sociability) lead to positive outcomes. Trait relevance occurs when a situation is relevant to a trait and offers the opportunity for the trait to be expressed (Tett et al., 2021). However, an extrovert person might also show negative work behavior if he or she is placed in an IT environment, lacking social interaction and having low personality- job fit. In this situation, Extraversion could be activated by excessive communication with coworkers, leading to negative work outcomes and hinder job tasks. This situational cue, or any cues that are not related to the specific personality- trait are considered to be trait- irrelevant (Tett et al., 2021). Trait- relevant situations expose individual differences on relevant traits, because they have a main effect on behavior: social situations increase social behavior in the people that are present and therefore reveal who is sociable and who is not. Trait relevance makes a trait observable in terms of individual differences.

Trait Activation Theory in Job Candidate Selection

Given the increasing competition for human capital, personality has become a popular topic in organizational research (LePine et al., 2011). Trait Activation Theory can be applied to job candidate selection, in a way where the interview questions asked are the situational cues that trigger specific traits in a job candidate. Thus, the way an interview question is designed, can help to activate the desired personality traits in a job interviewee. Asking trait- relevant questions would be a way to see if a job candidate possesses a desired personality trait, based on the given response (Judge & Zapata, 2015). Activation of personality traits that are valued on the job, will lead to increased job performance. Trait- irrelevant questions on the other hand, are not intentionally designed to measure personality. Exploring the match between the required traits for a specific job and the traits a job candidate possesses, can help in the assessment of job fit, job satisfaction and training success (Heimann et al., 2020). For example, an extrovert person might excel and deliver higher levels of performance in a role where they interact with people frequently, since their social skills are activated, ultimately enhancing their performance and job satisfaction. The current research will test Trait Activation Theory by measuring personality using an asynchronous video interview (AVI) that contains both trait-relevant and trait irrelevant interview questions.

Asynchronous Video Interviews (AVIs)

Asynchronous video interviews (AVIs) are a recent technological development used in assessing the hireability of job candidates (Dunlop et al., 2022). AVIs are one-way video interviews, in which potential job candidates answer job-related questions while recording themselves on their computer, without the live presence of a job interviewer (Lukacik et al., 2022). Afterwards, these videos are watched and rated by employers or analyzed automatically (Guchait et al. 2014). The psychometric properties of AVIs have not been established yet, even though scientific research on reliability and validity has just started receiving attention. This new

technological development has multiple potential benefits for employee selection in organizations, such as saving organizations time and money, since no live interview has to be scheduled. The AVI can be completed at any time and any place, which makes it easy for employers to reach candidates from all over the world. The same goes for personnel that have to review the AVI (Dunlop et al., 2022). AVI vendors and developers additionally often claim that AVI's predict job performance, are free of bias, are more engaging for job applicants than traditional assessment and have a good reliability (Mulfinger et al., 2020). This high reliability and validity are due to the inherent structure and standardization of the interviews (Lukacik et al., 2022). Since the live presence of an interviewer is lacking, there can be no deviation from interview protocol (Levashina et al., 2014). Taken together, these advantages improve the efficiency in high volume personnel selection projects (Dunlop et al., 2022). Therefore, as an addition to traditional job interviews, organizations are increasingly implementing the use of asynchronous video interviews in the early stages of applicant screening (Hickman et al., 2022).

It is important that the AVIs gather as much information about a job candidate as possible, to ensure there is a high job fit between the job and job candidate and that the candidate possesses the desired qualities. Previous research has shown that personality traits can be seen as valid predictors of performance in a wide range of jobs (Barrick & Mount, 1991; Judge & Zapata, 2015) and that they are considered to be major predictors of motivational performance criteria like organizational citizenship behavior (Cortina & Luchman, 2013). Personality traits determine habits and therefore shape behavior at work (Dudley & Cortina, 2008). Because of this, measuring personality characteristics is relevant in job interviews.

The Current Study

Previous research has shown that Trait Activation Theory might facilitate personality measurement, by adding trait- relevant interview questions to job interviews (Heimann, 2020;

Holtrop et al., 2022). However, this is not always done in real- life, since research by Heimann et al. (2020) shows that measuring the desired personality would add about 18 minutes of interviewing to a job interview. This can be considered to be a burden to both job candidates and interviewers and will also lead to an increase in costs. Therefore, it would be convenient to measure personality in job interviews through trait- irrelevant questions. However, it has not been studied before whether trait- irrelevant interview questions can measure personality to the same extent as trait- relevant questions do. So, this identifies a gap in the current literature available on AVI's. The current study identifies this gap in the literature by researching whether trait- irrelevant questions can be used to predict personality from job candidates in AVI's, since measuring personality through trait- irrelevant questions will save costs and time. This is because no trait- relevant questions have to be added to job interviews. If the results of the study show that trait- irrelevant questions explain the same amount of variance in personality characteristics as trait- relevant questions, AVI vendors would not have to add additional questions to find out whether a candidate's personality matches the specific requirements of the job, leading to a more time- and money efficient AVI.

HEXACO Model of Personality

The HEXACO Model of Personality is a six- dimensional framework used to structure personality. The model contains six personality factors: Honesty- Humility (H), Emotionality (E), Extraversion (X), Agreeableness (A), Conscientiousness (C) and Openness to Experience (O) (Ashton & Lee, 2007). Four of these personality traits will be used in the current research, so their definition will be elaborated on. Honesty- Humility can be defined as the extent to which someone is fair and genuine in dealing with others, in the sense of cooperation and refraining from manipulation, exploitation and fraud (Ashton & Lee, 2007). Extraverted people can be described as outgoing, energetic, warm and friendly. They seek to experience more positive

emotions and are optimistic, along with seeking excitement and stimulation (Zhao et al., 2009). Furthermore, Agreeableness is a dimension that characterizes trust, altruism, cooperation and modesty. Agreeable people show sympathy, concern for others needs and are trustworthy (Zhao et al., 2009). Lastly, Conscientiousness describes someone's work motivation, self- control, organization and planning and level of achievement and responsibility towards others (Roberts et al., 2005).

Conscientiousness in Predicting Job Performance

Personality traits in general relate to different work outcomes (Barrick & Mount, 1991). Conscientiousness shows consistent relations with all job performance criteria for all occupational groups and is regarded as the most consistent personality predictor of job performance across all types of occupations (Barrick & Mount, 1991). This personality trait has been shown to predict task performance (Ones et al., 1993), is highly correlated with long- term career success (Judge et al., 1999) and is considered to be the strongest predictor of organizational citizenship behavior (Borman et al., 2001), which relates to the behavior that serves to facilitate organizational functioning (Bourdage et al., 2018). Other studies consider Agreeableness, next to Conscientiousness, to be the top- ranked predictors of performance in a wide range of jobs (Barrick & Mount, 1991; Dunn et al., 2005; Sackett & Walmsley, 2014; Vinchur et al., 1998).

Agreeableness in Predicting Job Performance

Wilmot & Ones (2022) performed a large analysis on the effects of Agreeableness. This study showed that Agreeableness has been found to have an overall positive relationship with job performance. Agreeable people exhibit behaviors such as cooperation and positive interpersonal relationships, that are valued in many work settings. Additionally, the study showed that Agreeableness is negatively related to counterproductive work behaviors, such as aggression,

absenteeism and workplace deviance and that Agreeable people will likely show higher levels of leadership effectiveness. Leaders who score high on Agreeableness are often perceived as more effective, supportive and charismatic by their followers. Additionally, Agreeableness is positively related to various work-related attitudes, such as organizational commitment. Individuals high in Agreeableness tend to show more positive attitudes towards their job and have higher job satisfaction.

Extraversion in Predicting Job Performance

Extraversion is associated with high work performance, because this trait is associated with higher order goals that match well to the workplace. Extraverted people are more positive, adjust better and have higher interpersonal- and leadership skills, but are also good at influencing others, gaining status and achieving socially recognized success and helping others (Wilmot et al., 2019). So, an Extraverts' proactive behavior consisting of creativity, job crafting, advocating initiatives for organizational change and seizing career opportunity, contributes to a higher probability of success in receiving higher performance and desired rewards in the workplace. Additionally, Extraverts are found to receive more promotions and higher salary, since rewards reinforce their motivation and behavior. Findings of Wilmot et al. (2019) show that Extraversion is associated with higher overall job performance and lower counterproductive work behavior.

Honesty- Humility in Predicting Job Performance

Honesty- Humility is a relatively new predictor of individual differences in the applied psychology literature (Lee et al., 2019), but has been proposed as a valid predictor of multiple dimensions of job performance, including organizational citizenship behavior (Bourdage et al., 2018) and task performance (Oh et al., 2014), next to counterproductive work behavior (Sackett & Walmsley, 2014). The addition of this dimension improves the predictive validity of the HEXACO model, especially in predicting maladaptive behaviors, such as workplace deviance,

because low levels of Honesty- Humility are a strong predictor of workplace deviance (Pletzer et al., 2020). Research of Lee et al., 2005, shows that this trait can predict workplace deviance stronger than any personality trait.

Variable Relationships Between Personality Dimensions and Job Performance

The previous shows that personality dimensions show variable relationships with job performance. In occupations that involve frequent interaction or cooperation, Extraversion will be a valid predictor of performance criteria of occupations such as management and sales, but not for production workers or engineers (Barrick & Mount, 1991). On the contrary, Agreeableness has little effect on job performance in sales and management occupations, that involve goal-oriented achievement and emphasize production (Wilmot & Ones, 2021). Agreeableness has its strongest effects for occupations like customer service and health care, that involve helping. Research by Dunn et al. (1995) showed that there are negative relations between Conscientiousness and Agreeableness and counterproductivity. Later research showed that Honesty- Humility is also negatively related to counterproductive work behavior (Sackett & Walmsley, 2014). Research from Judge & Zapata (2015) showed that Conscientiousness is more important in jobs that require independence, whereas Agreeableness and Extraversion are more predictive of performance in jobs that require strong social skills. When competition is highly valued, Extraverted job candidates perform better in the job than people who score higher on Agreeableness. Extraversion and Agreeableness are also shown to be important when dealing with angry people (Judge & Zapata, 2015). These are just a few examples that reveal the relevance of general and specific situational conditions in facilitating the relevance of personality to job performance and thus, the importance of measuring personality traits in job interviews.

Furthermore, personality traits can interact with job characteristics to determine motivation at work (Smith & Denunzio, 2020). Volitional behavior is experienced as meaningful

when someone's personality traits align with job characteristics in terms of implicit motivational strings. An individual that scores high on conscientiousness is diligent and hardworking. Therefore, for this person, receiving feedback about their effort will lead to the individual experiencing positive emotional states and performing effective behaviors, leading to a positive impact of their work behaviors and performance.

Because of this, measuring personality is relevant in job interviews. The personality traits in the current research were collected using the HEXACO model. However, as mentioned before, there are differences in the extent to which specific personality traits predict job performance. Therefore, the current study will focus on the four most relevant personality traits in work settings: Honesty- Humility, Extraversion, Agreeableness and Conscientiousness (Koutsoumpis et al., 2024).

Measuring Personality from Spoken Text in AVIs

The current research aims to measure personality from spoken text in AVI, thereby focusing on verbal features of job candidates. Verbal features show the strongest correlations with personality traits compared to audio and visual features (Koutsoumpis et al., 2024). Research has shown that textual content of answers to standard interview questions can be used to infer personality traits reliably (Jayaratne & Jayatilleke, 2020; Saucier & Goldberg, 1996). Furthermore, the HEXACO model is grounded in the Lexical Hypothesis, which claims that personality characteristics are encoded in language, therefore showing the impact of language in defining personality traits (Jayaratne & Jayatilleke, 2020). The amount of information captured in syntax and semantics used in an interview, is greater than from a personality test using a limited set of multiple-choice questions (Jayaratne & Jayatilleke, 2020).

The theoretical framework that explains how and why verbal features might relate to personality inferences is Brunswik's Lens Model (1955). This framework explains why we make

judgements about other people and why and how a perception is accurate, based on behavioral cues. It is assumed that words people use are not randomly chosen, but are a behavioral manifestation of personality, along facial expressions (Hickman et al., 2022) and daily behavior (Tackman et al., 2020). This information is used by others to infer personality traits. We know someone is extraverted because we see them do activities related to extraversion, like going to parties and having a lot of social interaction. When observing this behavior from someone, people make inferences about this person being extraverted. These behaviors like words, tone of voice and gestures from which observers can make inferences about a person are called behavioral cues (Brunswik, 1955). The name of the model can be derived from the behavioral cues acting as a lens through which other people make inferences about a target. Besides the Lens Model, the focus on verbal features can also be explained by practical reasons: the Linguistic Inquiry Word Count Tool (LIWC) doesn't require machine learning skills.

Research of Mairesse et al. (2007) explored the possibility of acquiring personality traits by observing individual linguistic outputs in text and conversation. This study shows that personality can be observed from language in multiple ways. Using Extraversion as an example, extraverts talk more and louder, with fewer pauses and higher speech rates (Gill & Oberlander, 2002). Observers therefore predict someone producing long texts to be extraverted (Koutsoumpis et al., 2022). Also, extraverts use more informal language, while introverts use a broader vocabulary (Gill & Oberlander, 2002). Extraverts also use more words related to positive emotions, social behaviors and show more compliments and agreement compared to introverts (Pennebaker & King, 1999).

Linguistic Inquiry and Word Count (LIWC) tool

Pennebaker and Francis (1999) published the Linguistic Inquiry and Word Count (LIWC) tool, which can be used to analyze spoken text, by counting word categories and identifying

many linguistic features associated with Big Five personality traits (Pennebaker et al., 2015). It is a popular text analysis software that can be used to understand whether someone's use of linguistic categories is dependent on their personality traits (Koutsoumpis et al., 2022). The LIWC software has been used to find interesting correlations among language patterns and personality. This software has an internal dictionary and the latest version (LIWC2022) contains 12000 words and word stems that cover 90% of words that are commonly used in speech and writing (Koutsoumpis et al., 2022). These words can be placed in 93 linguistic categories, containing for example the number of words someone speaks or the amount of positive and negative emotions used in spoken text. These linguistic categories are considered to be verbal behavioral manifestations that indicate personality traits. When this software analyzes a text, it calculates the proportion of words that can be placed in each of the linguistic categories (Koutsoumpis et al., 2022). Therefore, this software can be used to determine what the percentage of trait- relevant words used in spoken text of job candidates in AVIs is. Research of Koutsoumpis et al. (2022) showed that both self- and observer- reported personality traits are significantly correlated with LIWC's linguistic categories.

Observer Reported Personality

The focus on the current study is on observer reports, because previous research has shown this type of reports account for significant variance in predicting job- relevant personality constructs compared to self- ratings (Koutsoumpis et al., 2024; Mount & Barrick, 1994) and that these personality constructs are only valid predictors of actual personality when they are assessed from the observer's perspective, but that this is not the case for assessment from the individual's perspective (Mount et al., 1994). Research Fast & Funder (2008) found that word use has a stronger relationship with observed behavior than with self- reported behavior, as they discovered a large number of correlations between word use and directly observed behavior. Furthermore,

observer reports are used because of practical and validity concerns, as they decrease response bias that is associated with self- reports, such as social desirability or impression management. This is the process through which someone influences how others perceive them in social interactions (Leary & Kowalski, 1990). Observers do not experience this pressure to present themselves favorably, therefore providing more unbiased personality assessment. Using self-reports would therefore decrease the validity of the results (Furnham & Henderson, 1982).

Research Question and Proposed Hypotheses

This research examines whether trait- irrelevant interview questions that are not intentionally designed to measure personality characteristics are able to offer the same validity as trait- relevant questions do in predicting personality in job candidates. The research question therefore is: *“Can personality traits from job candidates in asynchronous job interviews be accurately predicted from trait- irrelevant interview questions compared to trait- relevant interview questions?”* Drawing on Trait Activation Theory, the predictions of the current research are that observer reports of personality will be stronger correlated with participants’ verbal behavior in trait- relevant compared to trait- irrelevant interview question. This will be tested for the traits of Extraversion (Hypothesis 1), Conscientiousness (Hypothesis 2), Agreeableness (Hypothesis 3) and Honesty- Humility (Hypothesis 4). The current research contributes to the literature in examining whether trait- irrelevant interview questions can measure personality traits just as well as trait- relevant questions. If the results support the hypotheses, it will be a direct support for Trait Activation Theory and trait- relevant interview questions should be used specifically for assessing personality characteristics in AVI. If the hypotheses are not supported, it might mean that AVI developers could also rely on trait- irrelevant interview questions when predicting personality in AVI, saving time and costs in job interviews.

Method

Participants

The participants originally are 768 English- speaking adults, recruited through the panel Prolific. To be included in the research, participants had to be native English speakers, residing in the United States. The participants received 10 euro for their participation, which lasted for about 65 minutes. This is because video recordings that were marked as missing or incomplete, were excluded from the analysis. This is the case for people who provided non- compliant responses (people who did not take the study seriously from self- report or observer report or participants that participated twice), incomplete recordings, missing audio and video. Therefore, the data of 509 participants was used in the data analysis, with ages ranging from 18 to 65 years old ($M = 38.20$, $SD = 12.12$). From these participants, 237 were men (46.3%), 254 were women (49.6%) and 17 people identified as non- binary (3.3%). Three participants filled in the wrong value and therefore, their gender has not been identified.

Power analysis

A power analysis was conducted with use of the online calculator tool G*Power (Faul et al., 2009), to calculate the required sample size that is needed to achieve statistical power of 0.80, thereby detecting statistically significant effects in the current study. The results of this analysis provide insight into the reliability of the current study's findings. The outcome of this analysis shows that a sample size of 153 participants is needed to have a power of 0.80, with 71 predictors in the model, a desired F^2 value of 0.35 (which indicates a large effect), with a significance level set at 0.05. The current sample size is 509, which is larger than the required value. Therefore, the sample size is adequate for detecting true effects in the regression analyses.

Procedure

Prior to the start of the interview, participants read the description of a management

traineeship for a fictitious organization and were asked to participate in the selection procedure for this job. A participant answered questions that showed up on their computer screen, while recording themselves using their webcam and microphone. Every participant received the same set of questions, two generic or trait- irrelevant questions and four personality- related or trait- relevant questions measuring Honesty-Humility, Extraversion, Agreeableness and Conscientiousness. This happened without the live presence of an interviewer. Later on, trained raters watched the AVIs and gave specific scores on the participant's personality traits, to get scores on observer reports. Additionally, LIWC was used to analyze the number of linguistic categories used in spoken text of participants.

Included LIWC categories

Not all of the LIWC categories were used in the analysis. Firstly, the LIWC output categories based on punctuation were excluded, since the text data has been automatically transcribed. Therefore, some marks might be missing and transcripts might not have the right accuracy. Secondly, the LIWC output is designed hierarchically, meaning it is built up out of broad and narrow categories, where narrow categories together make up the broader category (Pennebaker et al., 2015). An example is the broader category "*Perceptual processes*" that is made up of the narrower categories "*see*", "*hear*" and "*feel*". Therefore, research of Koutsoumpis et al. (2022) pointed out the problem of "double- dipping" in the same category when measuring both the broad and narrow categories in the analysis, since this would lead to a double analysis of the broader category. Therefore, only the narrow categories were analyzed in the current research, to avoid the negative effects of "double- dipping". Analyzing only the broader categories would decrease the effect of narrow categories that offer unique personality insights, since these effects might be masked when they are averaged with other narrow categories in the broader category. The broad categories: "*Total function words*", "*Personal pronouns*",

“Affective processes”, “Social processes”, “Cognitive processes”, “Perceptual processes”, “Biological processes”, “Drives”, “Relativity” and *“Informal language”* have therefore not been analyzed. This left a total of 71 categories that were included as predictors in the analysis. Lastly, in preparation of the data, the variable ‘word count’ was divided by two for the answers that participants gave to trait- irrelevant questions, since there are two questions answered, compared to the one question that was answered per personality trait for the trait- relevant questions.

Measures

The data used in the current study is part of a data collection for a larger research project and therefore only the parts relevant in the current study will be mentioned. The project has received ethical approval from the Review Board of the Vrije Universiteit Amsterdam. The participants were asked six questions in the AVI, two of these are trait- irrelevant questions and the remaining four are questions that were designed to measure a specific personality trait. The specific questions asked in the research can be found in Appendix A. An example question that measures Agreeableness is: *“Think of situations when someone annoyed you. How do you usually react in such situations? Why do you think that is?”* Only four of the HEXACO personality traits have been examined. Openness to experience and Emotional Stability have been left out to minimize the burden for candidates, as these two personality traits are less relevant in work related environments (Koutsoumpis et al, 2024).

Observer Reports

Observer reports were generated of Extraversion for both trait- relevant ($M = 3.10$, $SD = .45$) and trait- irrelevant questions ($M = 3.22$, $SD = .41$), of Conscientiousness for both trait- relevant ($M = 3.34$, $SD = .39$) and trait- irrelevant questions ($M = 3.45$, $SD = .34$), of Agreeableness for both trait- relevant questions ($M = 3.32$, $SD = .40$) and trait- irrelevant

questions ($M = 3.37$, $SD = .35$) and lastly, for Honesty- Humility for both trait- relevant ($M = 3.27$, $SD = .35$) and trait- irrelevant questions ($M = 3.30$, $SD = .26$) with the use of Behavioral Anchored Rating Scales (BARS, Levashina et al., 2014). This is a scale that helps to assess performance of job applicants based on well- defined behavioral patterns. The use of BARS is associated with greater predictive validity and reliability and less bias in interviews (Kell et al., 2017).

The inter-rater agreement was calculated using Intraclass Correlation Coefficient (ICC) estimates in SPSS (Version 27), based on a mean- rating ($k = 4$), consistency, 2 way- mixed effects model. For Extraversion, the ICC value was .84 for trait- relevant questions and .67 for trait- irrelevant questions. For Conscientiousness, the ICC value was .83 for trait- relevant questions and .64 for trait - irrelevant questions. For Agreeableness, the ICC value was .77 for trait- relevant questions and .62 for trait - irrelevant questions, and for Honesty- Humility, the ICC value was .77 for trait- relevant questions and .48 for trait - irrelevant questions. Research shows that ICC values less than 0.50 are indicative of poor reliability, values between 0.50 and 0.75 indicate moderate reliability, values between 0.75 and 0.90 indicate good reliability and values above 0.90 indicate excellent reliability (Koo & Li, 2016).

Spoken Text

The spoken text of participants that participants answered to both trait- relevant and trait- irrelevant questions was analyzed using the Linguistic and Inquiry and Word Count (LIWC) text analysis software. When putting a written text into the software, it provides an output for a number of linguistic categories, which gives more information on the percentage of a certain personality trait used in the spoken text of the participants. The text length participants provided for answering the question related to Extraversion was ($M = 179.60$, $SD = 71.83$), the provided

text length for answer measuring Conscientiousness was ($M = 175.91$, $SD = 67.60$), the text length related to Agreeableness was ($M = 187.85$, $SD = 74.43$), for Honesty Humility, text length was ($M = 184.44$, $SD = 77.60$) and the average text length of the generic questions was ($M = 146.35$, $SD = 61.26$).

Statistical analysis

This is an observational study using a correlational design and no manipulation has taken place. For every personality trait, two multiple regression analyses were performed in SPSS Statistics (Version 27) to generate correlations between verbal behavior and observer reports of personality. For the first analysis, the trait- relevant interview questions were used and for the second analysis, trait- irrelevant questions were used. 71 LIWC categories were used as predictors in the regressions, which is the entire set, with exception of the categories that were previously mentioned. The adjusted R^2 was used to correct for the high numbers of predictors in the analysis. Prior to the analysis, all of the scores were converted to Z- scores in SPSS, to account for scale differences in predictors. A Z- test was used to compare if the two regressions differ significantly from each other, using the online Comparing Correlations tool (Diedenhofen & Musch, 2015). This tool uses capital R as a multiple correlation coefficient and converts it to Z- scores. The tool reports different kinds of Z- tests. Because the current sample is dependent and non-overlapping, Dunn and Clark's (1969) Z- test was to present the output, since this test is reported to have the best overall statistical properties (Heimann et al., 2020). The Z- score was based on a two- tailed test, with a significance level of $\alpha = .05$.

Assumptions

The assumptions of a regression analysis were tested prior to all analyses to see whether the results of the regression analysis are reliable and they will be discussed briefly. The first two assumptions are checked: the study uses continuous dependent variables and there are two or

more independent variables. Thirdly, the assumption of independence states that the observations should be independent of each other. There was independence of residuals, as assessed by plotting the Durbin- Watson statistic in SPSS (scores ranging between 1.66 and 2.05). A value of approximately 2 indicates that there is no correlation between residuals, meaning that this assumption is not violated.

Thirdly, the assumption of homoscedasticity states that the variance of residuals should be constant across all levels of the independent variables. This was tested by plotting normal P-P plots. For all analyses, the data points appear to follow the line and therefore this assumption was not violated. Also, the assumption of normality was checked by plotting histograms in SPSS, along with running normality tests that examine the skewness of the dependent variables. Research shows that a skewness of 0 indicates a perfectly normally distributed sample, but any value between -1 and 1 is seen as normally distributed (Mishra et al., 2019). This analysis shows that the data used in the current study is normally distributed and therefore, this assumption is not violated.

The assumption of outliers is not violated, but requires some elaboration. The predictors containing the LIWC categories do not represent outliers in the traditional way and do not require special treatment in the analysis. In regression analyses, extreme values can influence the estimation of regression coefficients and therefore make it more difficult to interpret the results. In the current study, the output of LIWC categories is a representation of the percentage of words a participant used in each linguistic category when providing their answers to interview questions, ranging from 0% to 100%. Because of this, high percentages should not be treated as if they are outliers, because they reflect meaningful variations in language use across different linguistic categories in the responses of participants. Therefore, the regression analysis remains robust to high values in predictors.

Furthermore, the assumption of multicollinearity was initially violated, because the results showed that multiple predictors show high correlations amongst each other ($r > .70$). In the current study, the use of LIWC categories as independent variables introduces a large number of predictors (71 in total), leaving a higher risk of multicollinearity issues. After generating the VIF values in SPSS (Version 27), 8 predictors with a VIF value larger than 10 were deleted from the model, since this indicates multicollinearity issues (Hair et al., 2014). After this, the regression analyses and Z- test were performed again for all analyses, leaving no multicollinearity issues. However, this did not yield any differences in the study's results.

Lastly, the assumption of linearity states that there should be a linear relationship between the dependent and independent variables. In the study, linearity between the dependent and independent variables collectively, was examined by plotting a scatterplot of the studentized residuals against the (unstandardized) predicted values. These scatter plots showed somewhat linear relationships, thereby not violating this assumption. Additionally, it was examined whether a linear relationship exists between the dependent variable and each of the independent variables, using partial regression plots between each independent variable and the dependent variable. However, not all of these regression plots showed a strictly linear pattern, meaning that this assumption is violated. This violation will be further elaborated on in the discussion section.

Results

Table 1

Descriptive Statistics and Correlations for Observer Reported Personality Traits

V	M	SD	1	2	3	4	5	6	7	8
1. Ex- <i>p</i>	3.06	.29	1	.						
2. Ex- <i>g</i>	3.12	.32	.35**	1						
3. Co- <i>p</i>	3.34	.40	-.06	-.07	1					
4. Co- <i>g</i>	3.49	.35	.01	.04	.34**	1				
5. Ag- <i>p</i>	3.31	.40	.17**	.04	.14**	.13**	1			
6. Ag- <i>g</i>	3.37	.36	.13**	.45*	.03	.12**	.15**	1		
7. HH- <i>p</i>	3.27	.35	.17**	.07	.20**	.14**	.36**	.11*	1	
8. HH- <i>g</i>	3.30	.27	.06	.40**	.05	.18**	.11*	.50**	.19**	1

Note. $N = 509$. Variables are observer ratings. * $p < .05$. ** $p < .01$. Ex = Extraversion; Co = Conscientiousness; Ag = Agreeableness; HH= Honesty- Humility. *p* = personality related questions; *g* = generic questions.

Table 2*Correlations between Observer Reported Personality Traits and Text Length per Question*

	Ex- <i>p</i>	Ex- <i>g</i>	Con- <i>p</i>	Con- <i>g</i>	Ag- <i>p</i>	Ag- <i>g</i>	HH- <i>p</i>	HH- <i>g</i>
TL- Ex	.29**	.25**	-	-	-	-	-	-
TL- Con	-	-	.05	.08	-	-	-	-
TL- Ag	-	-	-	-	-.01	-.05	-	-
TL- HH	-	-	-	-	-	-	.17**	.09*
TL- <i>g</i>	.24*	.24*	.07	.20**	.06	-.02	.18**	.14**

Note. *N* = 509. Variables are observer ratings. **p* < .05. ***p* < .01. TL = Text Length; Ex = Extraversion; Co = Conscientiousness; Ag = Agreeableness; HH= Honesty- Humility; *p* = personality related questions; *g* = generic questions. Only the relevant correlations have been noted.

Drawing on Trait Activation Theory, the predictions of the current research were that observer reports of personality correlated stronger with participants' verbal behavior in trait-relevant compared to trait- irrelevant interview question. This was tested for the traits of Extraversion (Hypothesis 1), Conscientiousness (Hypothesis 2), Agreeableness (Hypothesis 3) and Honesty- Humility (Hypothesis 4). In the first analysis that measured Extraversion, observer reports correlated with LIWC categories $R = .49$ ($p = <.001$) in trait-relevant and $R = .41$ ($p = <.001$) in trait-irrelevant interview questions, and this difference was not significant

($z = 1.50, p = .133$). As a result, Hypothesis 1 was not supported. Regarding Conscientiousness, observer reports correlated with LIWC categories $R = .38$ ($p = <.001$) in trait-relevant and $R = .45$ ($p = <.001$) in trait-irrelevant interview questions, and this difference was not significant ($z = -1.45, p = .147$). Therefore, Hypothesis 2 was not supported. Regarding Agreeableness, observer reports correlated with LIWC categories $R = .36$ ($p = <.001$) in trait-relevant and $R = .28$ ($p = .002$) in trait-irrelevant interview questions, and this difference was not significant ($z = 1.46, p = .146$). Therefore, Hypothesis 3 was not supported. Lastly, regarding Honesty- Humility, observer reports correlated with LIWC categories $R = .40$ ($p = <.001$) in trait-relevant and $R = .31$ ($p = <.001$) in trait-irrelevant interview questions, and this difference was not significant ($z = 1.68, p = .092$). Therefore, Hypothesis 4 was also not supported.

Table 3*Results of Linear Regression Analyses*

	Trait- relevant		Trait- irrelevant		Z- test results	
	<i>Adj. R²</i>	<i>p</i>	<i>Adj. R²</i>	<i>p</i>	<i>z</i>	<i>p</i>
H1. Extraversion	.239	< .001	.171	< .001	1.50	.133
H2. Conscientiousness	.144	< .001	.206	< .001	-1.45	.147
H3. Agreeableness	.132	< .001	.079	.002	1.46	.146
H4. Honesty- Humility	.163	< .001	.098	< .001	1.68	.092

Note. Dependent variable is observer reports. Independent variables are the LIWC categories. $N = 509$.

Discussion**Description of main findings**

The goal of the study was to test Trait Activation Theory's assertion that personality traits in job candidates are easier to measure in trait- relevant instead of trait- irrelevant interview questions, using the following research question: *"Can personality traits from job candidates in asynchronous job interviews be accurately predicted from trait- irrelevant interview questions compared to trait- relevant interview questions?"* This was examined because previous research has shown that variance in personality characteristics could be explained when asking trait- relevant questions to a job candidate, meaning that Trait Activation Theory might facilitate personality measurement, by adding trait- relevant interview questions to job interviews (Heimann, 2020; Holtrop et al., 2022). However, no research yet examined whether personality can also be predicted from the spoken text to trait- irrelevant questions. The results of the current study indicate that for Extraversion (H1), Agreeableness (H3) and Honesty- Humility (H4), observer reports of personality were stronger correlated with participants' verbal behavior in trait- relevant compared to trait- irrelevant interview question, initially appearing to support TAT's assumption that personality in job candidates is more readily observable from trait- relevant questions for these traits. Z- tests revealed that for none of analyses, the difference between these regressions is significant, meaning that observers can equally assess personality traits either from trait- relevant or trait- irrelevant interview questions. For Conscientiousness (H2), the opposite effect was found. For this specific trait, observer reports of personality were correlated stronger with participants' verbal behavior in trait- irrelevant compared to trait- relevant interview question. However, also for this analysis, a Z- test revealed that the difference between the two regressions was not statistically significant. All of the regression analyses were significant, meaning that according to this study, both trait- relevant and trait- irrelevant questions can be used to predict personality in job interviews. However, because the Z- scores revealed that the regression analyses do not differ from each other significantly, none of the hypotheses are

supported in this research and for none of the personality traits, trait- relevant questions can predict personality more accurately than trait- irrelevant questions can.

Relating the Current Findings to Previous Research

This does not support TAT's main assumption that personality can be more easily predicted from trait- relevant questions compared to trait- irrelevant questions and therefore, these results differ from previous research findings, for example of Tett & Bunnnett (2003) and Judge & Zapata (2014), who showed that trait activation significantly influences the validity of personality traits in predicting job performance. Koutsoumpis et al., (2022) also found support for Trait Activation Theory in AVIs. Additionally, Holtrop et al. (2022) developed a text- based personality assessment tool measuring HEXACO, and found that the construct validity of this tool was stronger when participants replied to trait- relevant questions compared to trait- irrelevant questions. Previous research has shown that asking trait- relevant questions enhances the construct validity of job interviews that rely on human judgment (Lievens et al., 2006) and that this also applies to job interviews that are analyzed by text- analysis techniques (Holtrop et al., 2022). Interestingly, the current study's analyses of inter-rater agreement ratings did reveal higher values for trait- relevant questions compared to trait- irrelevant questions, suggesting there may indeed be a trend toward higher agreement among observers in personality assessment based on trait- relevant questions for all personality traits, leaving some support for TAT's assumption. This particular finding is in line with research of Lievens et al. (2006), who found that inter-rater agreement among judges was stronger for trait- relevant, compared to trait- irrelevant exercises.

The results of this study also showed variability in inter-rater agreement ratings, meaning that some traits were more visible than others. Extraversion shows the highest inter-rater agreement, which is in line with previous research done by Funder (2012) and Funder & Dobroth (1987). These studies showed that Extraversion is easier to observe, because observers can detect

this through high levels of talkativeness and expressiveness (Funder, 2012). This trait is more therefore more visible, available, easier to detect, and able to be judged better than introspection. The more behavioral and relevant cues are visible in a target, the better observers will be able to make predictions about someone's personality (Koutsoumpis et al., 2022, Funder, 2012). These findings show that trait- behavior relations are complex, resulting from an interplay among diverse and competing trait- relevant situational cues (Tett & Guterman, 2000).

Theoretical Implications

Trait Activation Theory assumes that personality traits in job candidates are easier to measure in trait- relevant instead of trait- irrelevant interview questions, because personality traits in humans can be activated by certain environmental cues, at the time when someone is exposed to cues that relate to these specific traits (Tett & Bunnnett, 2003). The results of the current study show a slight trend supporting TAT's assumption that personality is more readily observable from trait- relevant questions, but no statistically significant differences were found between the two, meaning that the relationship may be more nuanced than was initially hypothesized and that TAT's assumption that personality traits in job candidates are easier to measure in trait- relevant instead of trait- irrelevant interview questions, is not supported by the results. These findings offer insights into the complexities of personality assessment in job interviews and therefore show the importance of considering any contextual factors that possibly influence the activation and observability of personality traits in job interview settings, like the specific wording of questions. Additionally, the nature of the job and the required candidate can also play a role in the extent to which personality traits are activated and observable from job interviews. Furthermore, the findings may show the need for a refinement of TAT's assumptions, accounting for factors that moderate the relationship between the question type and observability of personality traits.

Practical implications

The results of the current study have practical implications for AVI vendors and users. They indicate that trait- relevant questions do not measure personality traits more accurately than trait- irrelevant questions do, meaning that AVI vendors do not have to add additional questions to find out whether a candidate's personality matches the specific requirements of the job, leading to a more time- and money efficient AVI since a broader range of questions may be valuable in assessing candidate suitability. Another practical implication that should be mentioned is the awareness of bias and perceptual differences among raters. Even trained raters might interpret situations differently. Therefore, interviewers should be aware of their own biases and assumptions when assessing personality traits in job candidates, especially for traits that are less easily observable and leave room for interpretation, such as Honesty- Humility. A possibility would be to provide training programs for recruiters, interviewers or hiring managers, including guidance on how to objectively assess and interpret personality traits effectively. By providing clear definitions of personality traits, examples of behaviors that are associated with these traits, as well as strategies for recognizing these traits in job candidates, personality can be assessed in job candidates without providing personality- related questions. Additionally, providing feedback for interviewers can help with improving consistency and accuracy in personality assessments because opportunities for discussion create higher consensus among raters and refinement of assessment criteria.

Even though observer reports account for significant variance in predicting job- relevant personality constructs compared to self- ratings (Koutsoumpis et al., 2024, Mount & Barrick, 1994), research done by Vazire (2010) shows that self- and other knowledge of personality traits will vary by trait, due to individual differences in accessible information and susceptibility to biases or motivations (Bollich- Ziegler, 2020), This discrepancy might lead to differences in the accuracy of judgments of internal and external traits. Therefore, using self- reports and observer

ratings together in personality assessment can help to provide a more comprehensive and reliable assessment of personality traits in job candidates, because using them together can reveal more about personality traits than either method could do alone (McCrae, 1994). This method can help to compensate for possible limitations in single assessment methods, like response bias in self-reports thereby enhancing the validity of personality assessments in job interviews, while still gaining information on internal traits.

Limitations and future research

However, there are limitations that have to be taken into consideration concerning the validity of the results of the current study. Firstly, potential confounding factors have not been analyzed. For example, certain candidate characteristics like communication skills, honesty or interview conditions (like time constraints) could have impacted the observed relationships between spoken text of job candidates and their observer reported scores of personality. Someone familiar with job interviews might be able to express themselves better than an inexperienced candidate. Additionally, some candidates might not have been able to speak freely in their surroundings. This could be a threat to the validity of the information that the observer reports are based on. Therefore, future research should take into account additional confounding variables that may moderate the relationship between personality ratings and spoken text and ensure participants are familiar with the situation and clarify the need to express themselves freely.

Secondly, given that all participants are residents of the USA and had to be native speakers to participate, the generalizability of the study's findings is limited to the population of the United States, meaning that the participants have similar demographic characteristics. However, it is essential to acknowledge the existence of cultural and linguistic differences across countries and even regions, which may influence the communication styles, language used and

behavioral responses in job interviews. Therefore, future research would need to be done in different cultures, to generalize the findings to a broader population.

Thirdly, the current research had four observers that rated the AVI's for job candidate's personality. They received training prior to rating the AVI's in order to enlarge inter-rater reliability and ensure consistency in rating criteria and procedures. However, measurement errors such as inaccuracies in observer ratings should still be taken into account in explaining the observed relationships. The inter-rater agreement was moderate to high for almost all analyses, but was low for the observer reported personality traits from trait- irrelevant questions. There is still a chance that certain observations remain subjective, meaning that they are influenced by an observer's own interpretations and experiences. Different observers may perceive the same behavior differently, leading to variability in personality ratings across observers. Developing clear rating guidelines that outline specific criteria for assessing personality traits will help to increase consistent and reliable ratings across observers. These guidelines could include examples of observable behaviors per personality trait and descriptions of specific trait manifestations. Future research should increase the number of observers, since this allows for a more extensive assessment of personality traits and it will reduce the impact of situational factors and measurement error.

Furthermore, the assumption of linearity was violated. The relationship between language and personality is not always simple, meaning that personality might not change in a predictable way when language changes. This is because language is very complex and it might affect personality traits in different ways. An example of this is that using complex words can be associated with high levels of conscientiousness, since this might indicate intelligence and attention to detail. However, a non- monotonic pattern can be identified when a person uses language that is too complex. In this situation, someone might come across as more pedantic.

Eventually, this threshold would lead to lower ratings of conscientiousness and have an effect on the relation between language and personality traits.

Conclusion

The results of this study indicate that there is no significant difference in predicting personality traits from job candidates through trait- relevant questions and trait- irrelevant questions. These findings implicate that trait- relevant questions do not have to be added to a structured job interview, because trait- irrelevant questions might explain personality similarly without sacrificing predictive validity. Therefore, both organizations using asynchronous video interviews and AVI vendors should take this into consideration when structuring an AVI. Not having to add trait- relevant questions would make the AVI more efficient, since no extra questions have to be added in order to measure personality.

To summarize, the study's findings suggest that both trait- relevant and trait- irrelevant questions can contribute to effective personality measurement in AVIs. The study is the first to explore whether trait- irrelevant interview questions can be used to predict personality characteristics in job candidates during asynchronous video interviews. Therefore, future research that accounts for the current study's limitations is needed in order to increase validity in the results, by increasing the number of raters, diversifying the sample and including confounding variables.

References

- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11(2), 150–166. <https://doi.org/10.1177/1088868306294907>
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance. *Personnel Psychology*, 44(1), 1–26.
<https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Borman, W. C., Penner, L. A., Allen, T. D., & Motowidlo, S. J. (2001). Personality predictors of citizenship performance. *International Journal of Selection and Assessment*, 9(1–2), 52–69. <https://doi.org/10.1111/1468-2389.00163>
- Brunswik, E. (1955). Representative design and probabilistic theory in functional psychology. *Psychological Review*, 62(3), 193–217. <https://doi.org/10.1037/h0047470>
- Cortina, J. M., & Luchman, J. N. (2013). Personnel selection and employee performance. W. Schmitt, S. Highhouse, I. B. Weiner, N. W. Schmitt, S. Highhouse, & I. B. Weiner (Eds.), *Handbook of psychology: Industrial and organizational psychology*, Vol. 12, 2nd ed. (pp. 143–183). John Wiley & Sons Inc.
- Diedenhofen, B. & Musch, J. (2015). Cocor: A Comprehensive Solution for the Statistical Comparison of Correlations. *PLOS ONE*, 10(4): e0121945.
[doi:10.1371/journal.pone.0121945](https://doi.org/10.1371/journal.pone.0121945)
- Dudley, N. M., & Cortina, J. M. (2008). Knowledge and skills that facilitate the personal support dimension of citizenship. *Journal of Applied Psychology*, 93(6), 1249–1270.
<https://doi-org.tilburguniversity.idm.oclc.org/10.1037/a0012572>.
- Dunlop, P. D., Holtrop, D., & Wee, S. (2022). How asynchronous video interviews are used in practice: A study of an Australian- based AVI vendor. *International Journal of Selection*

- and Assessment*, 30(3), 448–455. <https://doi.org/10.1111/ijasa.12372>
- Dunn, O. J., & Clark, V. A. (1969). Correlation coefficients measured on the same individuals. *Journal of the American Statistical Association*, 64(325), 366–377. <https://doi.org/10.1080/01621459.1969.10500981>
- Dunn W. S., Mount M. K., Barrick M. R., Ones D. S. (1995). Relative importance of personality and general mental ability in managers' judgments of applicant qualifications. *Journal of Applied Psychology*, 80(4), 500–509. <https://doi.org/10.1037/0021-9010.80.4.500>
- 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, 1149–1160.
- Fast, L., & Funder, D. (2008). Personality as manifest in word use: correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology*, 94(2), 334–346. <https://doi.org/10.1037/0022-3514.94.2.334>.
- Funder, D. C. (2012). Accurate personality judgment. *Current Directions in Psychological Science*, 21(3), 177–182. <https://doi.org/10.1177/0963721412445309>
- Funder, D. C. 2001. Personality. *Annual Review of Psychology*, 52, 197–221.
- Funder D. C., & Dobroth K. M. (1987). Differences between traits: Properties associated with interjudge agreement. *Journal of Personality and Social Psychology*, 52(2), 409–418. <https://doi.org/10.1037/0022-3514.52.2.409>
- Furnham, A., & Henderson, M. (1982). The good, the bad and the mad: Response bias in self report measures. *Personality and Individual Differences*, 3(3), 311–320. [https://doi.org/10.1016/0191-8869\(82\)90051-4](https://doi.org/10.1016/0191-8869(82)90051-4).
- Gill, A. J., & Oberlander, J. (2002). Taking care of the linguistic features of extraversion. In

- Proceedings of the 24th Annual Conference of the Cognitive Science Society, 363-368/
- Guchait, P., Ruetzler, T., Taylor, J., & Toldi, N. (2014). Video interviewing: A potential selection tool for hospitality managers—A study to understand applicant perspective. *International Journal of Hospitality Management*, 36, 90–100.
<https://doi.org/10.1016/j.ijhm.2013.08.004>
- Hair, J. F., Sarstedt, M., Hopkins, L., & Kuppelwieser, V. G. (2014). Partial least squares structural equation modeling (PLS-SEM). *European Business Review*, 26(2), 106–121.
<https://doi.org/10.1108/eb-10-2013-0128>
- Heimann, A. L., Ingold, P. V., Debus, M. E., & Kleinmann, M. (2020). Who will go the extra mile? Selecting organizational citizens with a personality-based structured job interview. *Journal of Business and Psychology*, 36(6), 985–1007.
<https://doi.org/10.1007/s10869-020-09716-1>
- Hickman, L., Bosch, N., Ng, V., Saef, R., Tay, L., & Woo, S. E. (2022). Automated video interview personality assessments: Reliability, validity, and generalizability investigations. *Journal of Applied Psychology*, 107(8), 1323–1351.
<https://doi.org/10.1037/apl0000695>
- Holtrop, D.J., Oostrom, J.K., Van Breda, W.R.J., Koutsoumpis, A., & De Vries, R.E. (2022). Exploring the application of text-to-personality algorithms in job interviews. *European Journal of Work and Organizational Psychology*, 31(6), 1–18.
<https://doi.org/10.1080/1359432X.2022.2051484>
- Jayarathne, M., & Jayatilake, B. (2020). Predicting Personality Using Answers to Open-Ended Interview Questions. *IEEE Access*, 8, 115345-115355.
<https://doi.org/10.1109/ACCESS.2020.3004002>

- Judge, T. A., & Higgins, C.A., Thoresen C.J, Barrick, M.R. (1999). The Big Five Personality traits, general mental ability, and career success across the lifespan. *Personnel Psychology*, 52(3), 621- 652. <https://doi.org/10.1111/j.1744-6570.1999.tb00174.x>
- Judge, T. A., & Zapata, C. P. (2015). The person-situation debate revisited: Effect of situation strength and trait activation on the validity of the big five personality traits in predicting job performance. *Academy of Management Journal*, 58(4), 1149–1179. <https://doi.org/10.5465/amj.2010.0837>
- Kell, H. J., Martín-Raugh, M. P., Carney, L., Inglese, P., Chen, L., & Feng, G. (2017). Exploring methods for developing behaviorally anchored rating scales for evaluating structured interview performance. *ETS Research Report Series*, 2017(1), 1–26. <https://doi.org/10.1002/ets2.12152>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal Of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Koutsoumpis, A., Ghassemi, S., Oostrom, J. K., Holtrop, D., Van Breda, W., Zhang, T., & De Vries, R. E. (2024). Beyond traditional interviews: Psychometric analysis of asynchronous video interviews for personality and interview performance evaluation using machine learning. *Computers in Human Behavior*, 154, 108128. <https://doi.org/10.1016/j.chb.2023.108128>
- Koutsoumpis, A., Oostrom, J. K., Holtrop, D., Van Breda, W., Ghassemi, S., & De Vries, R. E. (2022). The kernel of truth in text-based personality assessment: A meta-analysis of the relations between the Big Five and the Linguistic Inquiry and Word Count (LIWC). *Psychological Bulletin*, 148(11–12), 843–868. <https://doi.org/10.1037/bul0000381>
- Leary, M., & Kowalski, R. (1990). Impression management: A literature review and two

- component model. *Psychological Bulletin*, 107(1), 34–47.
<https://doi.org/10.1037/0033-2909.107.1.34>
- Lee, K., & Ashton, M. C. (2004). Psychometric Properties of the HEXACO Personality Inventory. *Multivariate Behavioral Research*, 39(2), 329–358.
http://dx.doi.org/10.1207/s15327906mbr3902_8
- Lee, K., Ashton, M. C., & Shin, K. H. (2005). Personality correlates of workplace anti-social behavior. *Applied Psychology*, 54(1), 81–98.
- Lee, Y., Berry, C. M., & Gonzalez-Mulé, E. (2019). The importance of being humble: A meta-analysis and incremental validity analysis of the relationship between honesty-humility and job performance. *Journal of Applied Psychology*, 104(12), 1535–1546.
<https://doi.org/10.1037/apl0000421>
- Lievens, F., Chasteen, C. S., Day, E. A., & Christiansen, N. D. (2006). Large-scale investigation of the role of trait activation theory for understanding assessment center convergent and discriminant validity. *Journal of Applied Psychology*, 91(2), 247–258.
<https://doi.org/10.1037/0021-9010.91.2.247>
- LePine J.A., Buckman B.R., Crawford E.R., & Methot J.R. (2011). A review of research on personality in teams: accounting for pathways spanning levels of theory and analysis. *Human Resource Management Review*, 21, 4311–30.
- Letzring, T. D., & Funder, D. C. (2021). The realistic accuracy model. In T. D. Letzring & J. S. Spain (Eds.), *The Oxford Handbook of Accurate Personality Judgment* (pp. 9–22). Oxford University Press.
- Levashina, J., Hartwell, C. J., Morgeson, F. P., & Campion, M. A. (2014). The structured employment interview: Narrative and quantitative review of the research literature,

- Personnel Psychology*, 67(1), 241–293. <https://doi.org/10.1111/peps.12052>
- Lukacik, E. R., Bourdage, J. S., & Roulin, N. (2022). Into the void: A conceptual model and research agenda for the design and use of asynchronous video interviews. *Human Resource Management Review*, 32(1), 100789. <https://doi.org/10.1016/j.hrmr.2020.100789>
- McCrae, R. (1994). The Counterpoint of Personality Assessment: Self Reports and Observer Ratings. *Assessment*, 1(2), 159 - 172. <https://doi.org/10.1177/1073191194001002006>.
- Mishra, P., Pandey, C. K., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, 22(1), 67–72. https://doi.org/10.4103/aca.aca_157_18
- Mount, M. K., Barrick, M. R., & Strauss, J. P. (1994). Validity of observer ratings of the big five personality factors. *Journal of Applied Psychology*, 79(2), 272–280. <https://doi.org/10.1037/0021-9010.79.2.272>
- Mulfinger, E., Wu, F., Alexander, L., & Oswald, F. L. (2020). AI technologies in talent management systems: It glitters but is it gold? *Work in the 21st Century: Automation, Workers, and Society*, Houston, TX.
- Oh, I.-S., Le, H., Whitman, D. S., Kim, K., Yoo, T.-Y., Hwang, J.-O., & Kim, C.-S. (2014). The incremental validity of honesty–humility over cognitive ability and the Big Five personality traits. *Human Performance*, 27(3), 206–224. <http://dx.doi.org/10.1080/08959285.2014.913594>
- Ones D.S, Viswesvaran C, Schmidt F.L. (1993). Comprehensive meta- analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78(4), 679–703.

<https://doi.org/10.1037/0021-9010.78.4.679>

- Organ, D. W., Podsakoff, P. M., & Podsakoff, N. P. (2011). Expanding the criterion domain to include organizational citizenship behavior: Implications for employee selection. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology*. (pp. 281–323). American Psychological Association
- Pennebaker, J. W., Boyd, R. L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC2015. (pp. 1–26). The University of Texas at Austin.
<https://dx.doi.org/10.15781/T29G6Z>
- Pennebaker, J. W., & Francis, M. E. (1999). Linguistic inquiry and word count (LIWC): A computer-based text analysis program. Lawrence Erlbaum Publishers.
- Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology*, 77, 1296–1312.
- Pletzer J. L., Oostrom J. K., Bentvelzen M., de Vries R. E. (2020). Comparing domain- and facet level relations of the HEXACO personality model with workplace deviance: A meta-analysis. *Personality and Individual Differences*, 152(1), 109539.
- Rule N. O., & Ambady N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science*, 19(2), 109–111.
<https://doi.org/10.1111/j.1467-9280.2008.02054.x>
- Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, 9(5), 538–551.
<https://doi.org/10.1177/1745691614543972>

- Saucier, G., & Goldberg, R. L. (1996). The language of personality: Lexical perspectives on the five- factor model. *The Five- Factor Model of Personality: Theoretical Perspectives*. New York, NY, USA: Guilford Press. Pp. 21-50.
- Simonet, D. V., & Tett, R.P. (2013). 5 perspectives on the leadership-management relationship: a competency-based evaluation and integration. *Journal of Leadership and Organizational Studies*, 20(2), 199–213, doi:10.1177/1548051812467205
- Smith, R., & Denunzio, M. (2020). Examining personality—Job characteristic interactions in explaining work outcomes. *Journal of Research in Personality*, 84, 103884.
<https://doi.org/10.1016/j.jrp.2019.103884>.
- Ones D.S, Viswesvaran C, Schmidt F.L. (1993). Comprehensive meta- analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78, 679–703.
- Tackman, A. M., Baranski, E. N., Danvers, A. F., Sbarra, D. A., Raison, C. L., Moseley, S. A., Polsinelli, A. J., & Mehl, M. R. (2020). ‘Personality in its natural habitat’ revisited: A pooled, multi-sample examination of the relationships between the Big Five personality traits and daily behaviour and language use. *European Journal of Personality*, 34(5), 753–776. <https://doi.org/10.1002/per.2283>
- Tett, R. P., & Burnett, D. D. 2003. A personality trait-based interactionist model of job performance. *Journal of Applied Psychology*, 88, 500–517.
<https://doi-org.tilburguniversity.idm.oclc.org/10.1037/0021-9010.88.3.500>.
- Tett, R., & Guterman, H. (2000). Situation Trait Relevance, Trait Expression, and Cross Situational Consistency: Testing a Principle of Trait Activation. *Journal of Research in Personality*, 34, 397-423. <https://doi.org/10.1006/JRPE.2000.2292>.

- Tett, R.P., Toich, M.J., Ozkum, S.B. (2021). Trait Activation Theory: A Review of the Literature and Applications to Five Lines of Personality Dynamics Research. *Annual Review of Organizational Psychology and Organizational Behavior*, 8(1), 199–233.
[doi:10.1146/annurev-orgpsych-012420-062228](https://doi.org/10.1146/annurev-orgpsych-012420-062228)
- Vazire, S. (2010). Who knows what about a person? The self–other knowledge asymmetry (SOKA) model. *Journal of Personality and Social Psychology*, 98(2), 281–300.
<https://doi.org/10.1037/a0017908>
- Vinchur A. J., Schippmann J. S., Switzer F. S., Roth P. L. (1998). A meta-analytic review of predictors of job performance for salespeople. *Journal of Applied Psychology*, 83, 586–597.
- Wilmot M. P., Ones D. S. (2021). Occupational characteristics moderate personality performance relations in major occupational groups. *Journal of Vocational Behavior*, 131, 103655. <https://doi.org/g4dd>
- Wilmot, M., & Ones, D. (2022). Agreeableness and its consequences: A quantitative review of meta-analytic findings. *Personality and Social Psychology Review*, 26(3), 242–280.
<https://doi.org/10.1177/10888683211073007>.
- Wilmot, M., Wanberg, C., Kammeyer-Mueller, J., & Ones, D. (2019). Extraversion advantages at work: A quantitative review and synthesis of the meta-analytic evidence. *The Journal of Applied Psychology*, 104(12), 1447–1470. <https://doi.org/10.1037/apl0000415>.
- Zhao, H., Seibert, S. E., & Lumpkin, G. T. (2009). The Relationship of Personality to Entrepreneurial Intentions and Performance: A Meta-Analytic Review. *Journal of Management*, 36(2), 381–404. <https://doi.org/10.1177/0149206309335187>

Appendix

Appendix A

The Interview Questions used in the AVI

1. What would you consider among your greatest strengths and weaknesses as an employee?
[Trait- irrelevant question]
2. How would your best friend describe you? [Trait- irrelevant question]
3. Think of situations when you made professional decisions that could affect your status or how much money you make. How do you usually behave in such situations? Why do you think that is? [Honesty-Humility]
4. Think of situations when you joined a new team of people. How do you usually behave when you enter a new team? Why do you think that is? [Extraversion]
5. Think of situations when someone annoyed you. How do you usually react in such situations? Why do you think that is? [Agreeableness]
6. Think of situations when your work or workspace were not very organized. How typical is that of you? Why do you think that is? [Conscientiousness]