Role of Ability to Identify Criteria in Self–Observer Personality Agreement in Asynchronous Video Interviews

K.T. Bui

SNR: 2066188

550012-B-12: Bachelor's Thesis: Work and Organizational Psychology

Work and Organizational Psychology, Department of Social Psychology, Tilburg University

Supervisor: Dr. Antonios Koutsoumpis

Second assessor: Dr. Djurre Holtrop

June 30, 2025

Abstract

Asynchronous video interviews (AVIs) are becoming more common in personnel selection; however, the role of the ability to identify criteria (ATIC), which refers to a candidate's ability to understand what qualities or skills are being evaluated during an interview, remains understudied concerning self-observer personality agreement within this format. The study examined whether ATIC moderates self–observer agreement on personality traits in AVIs, focusing on two traits from the HEXACO model: Extraversion (high visibility) and Honesty-Humility (low visibility). A sample of 646 participants (M_{Age}= 36.69, SD = 11.92) completed an AVI with the personality-related questions and a self-reported personality questionnaire. A group of trained raters evaluated participants' personality traits based on their video responses, and a second group evaluated participants' ATIC based on their written responses. Findings indicated that self-reported Honesty-Humility, Extraversion, and ATIC each had significant positive effects on observer ratings. Self–observer agreement was significantly weaker for Honesty-Humility (r=.22) than for Extraversion (r=.42). Regarding moderation effects, ATIC strengthened the alignment between self- and observer ratings for Extraversion, such that candidates with higher ATIC scores showed better agreement. However, no moderation effect was found for Honesty-Humility. These findings suggest that visible traits are more accurately inferred in AVIs and that ATIC enhances agreement only when relevant behavioral cues are detectable. While AVIs can be effective for evaluating visible traits, they may be insufficient for assessing low-visibility traits, highlighting the need for complementary assessment methods in the hiring process.

Keywords: asynchronous video interviews (AVIs), ability to identify criteria (ATIC), self–observer agreement, Extraversion, Honesty-Humility, personnel selection.

The Shift to Asynchronous Video Interviews in Personnel Selection

Recent technological advancements have transformed recruitment practices, increasing the use of alternative interview formats such as videoconferencing, telephone, and asynchronous video interviews (AVIs), replacing traditional face-to-face interviews (Oliphant et al., 2008). In AVIs, candidates record their responses to pre-set interview questions using digital devices such as computers, laptops, or smartphones, without direct interaction with interviewers (Lukacik et al., 2022). These formats offer scalability, efficiency, and the potential for standardized assessments, making them increasingly attractive in personnel selection, especially in light of the COVID-19 pandemic, which accelerated their adoption (Dunlop et al., 2022; Levashina et al., 2014; Lukacik et al., 2022). AVIs are used to either automatically or manually evaluate characteristics of the candidates, including personality traits. However, little is known about the factors that might affect the accuracy of self-other agreement. The present study addresses this gap by exploring self–other agreement in AVIs, focusing on whether the accuracy of this agreement depends on a candidate's ability to identify the qualities or skills that interview assessors aim to measure.

Research Question: Does the relationship between self-reported and observer-reported personality traits depend on whether a candidate correctly identifies what assessors measure in asynchronous video interviews?

Personality Assessment in Digital Interviews

One growing area of interest in personnel selection science is the assessment of personality traits via AVIs. Traits outlined in the HEXACO model, which include Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness, are robust predictors of job performance, organizational citizenship behavior, and ethical conduct

(Chamorro-Premuzic & Furnham, 2010; Pletzer et al., 2019). As AVIs enable the evaluation of candidates' behaviors by human raters, they present an opportunity to reduce socially desirable responses and capture behavioral expressions of personality (Chamorro-Premuzic & Furnham, 2010). However, these technology-mediated contexts raise important concerns, particularly regarding how personality ratings obtained through AVIs align with traditional self-report measures.

Self-reports are widely used in recruitment due to their methodological simplicity, practical convenience, and administrative efficiency. However, they are susceptible to biases such as faking, lack of self-insight, and social desirability (McCrae, 1994). In contrast, observer reports, in which external raters judge a target's personality based on observed behavior, tend to correlate moderately with self-reports and can offer additional and sometimes more accurate information (Connelly & Ones, 2010; Kim et al., 2019). This is largely attributed to their greater objectivity and reduced susceptibility to biases such as defensiveness and impression management (McCrae, 1994). Research indicates that observer reports provide incremental predictive validity over self-reports in occupational settings, particularly for externally observable traits (Oh et al., 2011). However, such convergence may be reduced in structured, short-duration formats like AVIs.

Focal Traits and Trait Visibility: Integrating the SOKA Model

In line with assumptions of the Self–Other Knowledge Asymmetry (SOKA) model (Vazire, 2010) and to simplify the study design, the present study focuses on two personality traits from the HEXACO model, namely Extraversion and Honesty-Humility, which differ in their levels of behavioral visibility. The SOKA model distinguishes traits by (1) observability or visibility (i.e., a trait can be easily detected by an external observer) and (2) evaluativeness (i.e.,

a trait is socially desirable or undesirable). These dimensions affect the accuracy of personality judgments and the degree of self–other agreement.

Extraversion, characterized by sociability, assertiveness, and energy, is highly visible and generally low in evaluativeness, resulting in higher self-observer agreement (Oh et al., 2011; Connelly & Ones, 2010). Conversely, Honesty-Humility, reflecting sincerity, fairness, greed avoidance, and modesty (Ashton & Lee, 2007), is more difficult to infer from behavior in short, structured settings (Howard & Van Zandt, 2020). As a less visible and more evaluative trait, it often results in lower self-observer alignment (Varize, 2010). Research confirmed this distinction: self–other agreement is typically weaker for Honesty-Humility (r=.42) than for Extraversion (r=.56), consistent with differences in trait visibility (de Vries, Lee, & Ashton, 2008). Nevertheless, even for a low-visibility trait, observer ratings can meaningfully predict workplace behaviors, including integrity and counterproductive work behavior (de Vries, 2015; Zettler & Hilbig, 2010; Dietl & Kombeiz, 2025).

Theoretical Frameworks: Realistic Accuracy Model and Ability to Identify Criteria

Given the study's focus on the accuracy of personality inference in asynchronous video interviews (AVIs), the Realistic Accuracy Model (RAM; Funder, 1995) provides a comprehensive theoretical framework to understand how and when accurate personality judgments occur. The RAM model outlines the four key conditions necessary for accurate personality judgments: (1) the target must exhibit behavioral cues that are relevant to and consistent with the trait being evaluated; (2) those cues must be available in external contexts; (3) the observer must focus on the target to detect the relevant and available cues; and (4) the observer must correctly utilize the detected cues to make an accurate judgment. If any condition fails, judgment accuracy is undermined. In the context of AVIs, where interactions are short and

structured, meeting this sequence of events may be more challenging unless candidates deliberately engage in behaviors that highlight the trait being assessed (Letzring & Funder, 2021).

This leads to the importance of Ability to Identify Criteria (ATIC), a candidate's capacity to infer what is being assessed during an interview (Kleinmann et al., 2011). Research shows that candidates often attempt to anticipate desirable traits and adjust their behaviors accordingly (Levashina et al., 2014). Candidates with high ATIC, as a good target, could potentially tailor their responses and behaviors to highlight trait-relevant information, increasing cue availability and facilitating observer detection and utilization of such cues. Thus, ATIC serves as a crucial moderator within the RAM framework, potentially strengthening the alignment between self-reports and observer ratings by enhancing the availability of good information. Based on this reasoning, it is hypothesized that this moderating role will influence both Extraversion, a highly observable trait, and Honesty-Humility, a more subtle and less visible trait.

Hypothesis 1: ATIC will moderate the positive correlation between self- and observer-reported Honesty-Humility, such that the correlation will be stronger for individuals with higher ATIC scores.

Hypothesis 2: ATIC will moderate the positive correlation between self- and observer-reported Extraversion, such that the correlation will be stronger for individuals with higher ATIC scores.

The Present Study

Despite increasing use of AVIs in personnel selection, little is known about how well observer ratings derived from AVIs align with self-reports, and more importantly, whether this alignment is influenced by candidates' awareness of the assessment criteria. This study

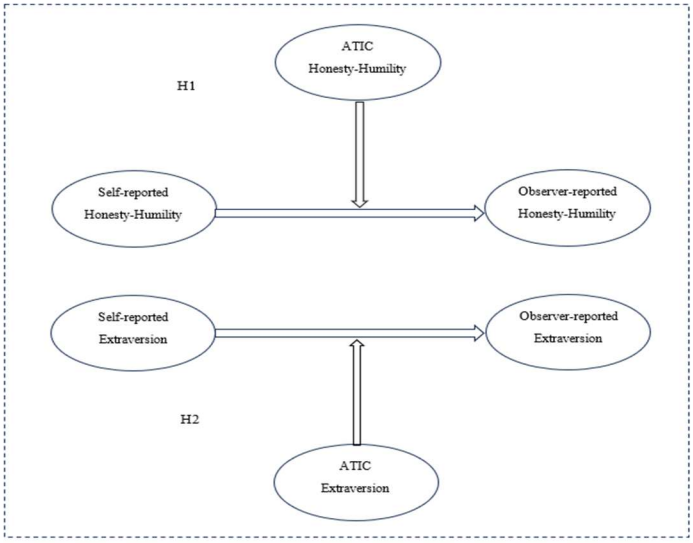
investigates the moderating role of ATIC on the relationship between self- and observer-rated personality traits in AVIs. To maintain focus and avoid an overly complex set of hypotheses, the study concentrated on two personality traits that differ in visibility: Extraversion and Honesty-Humility.

An online experimental study was conducted in which participants were presented with a fictitious management traineeship and completed an AVI comprising six structured questions. They also completed the HEXACO-100 personality inventory and open-ended questions measuring ATIC. Based on the AVIs and participants' responses to two open-ended questions, personality traits of Extraversion and Honesty-Humility, along with ATIC were rated by trained observers. This enables an examination of ATIC's moderating effect on the alignment between self- and observer ratings. The hypotheses formulated to address the research question are shown in Figure 1.

By exploring the moderating role of candidate awareness in AVI-based personality assessment, this study may contribute to a deeper theoretical understanding of self–observer agreement in personality judgment and offer practical insights for enhancing selection practices in contemporary recruitment settings.

Figure 1

Hypotheses and Variables



Note. H1, H2: Hypothesis 1,2. This is a conceptual model of hypothesized moderation effects of ATIC on self–observer agreement.

Methods

The dataset utilized in this study is sourced from Koutsoumpis et al. (2025). For the current analyses, this study focused exclusively on the variables relevant to the research hypotheses, including some participants' demographic information, self- and observer-reported measures of Extraversion and Honesty-Humility, and participants' Ability to Identify Criteria (ATIC) related to these two traits.

Participants

793 participants were initially accessed via Prolific's online platform (Peer et al., 2017). Following data checking, some participants were excluded for some reasons, such as incomplete responses, lack of consent for data sharing, outlier levels of response variability on HEXACO items, corrupted audio in video recordings, and other issues. The final sample comprised 646 participants, native English speakers from the United States (309 men, 309 women, and 28 nonbinary). Most of the participants were predominantly White (n=467), Black or African

American (*n* = 73), Hispanic or Latino (*n* = 46), Asian (*n* = 30), other (*n* = 25), or did not report ethnicity (*n* = 5). The average age was 36.69 years (SD = 11.92), with 15.96 years of work experience on average (*SD* = 11.36). Education levels ranged from incomplete high school (n = 6) to doctoral degrees (n = 16), with most holding a Bachelor's (n = 233) or Master's degree (n = 89); moreover, 6 participants did not report their educational background (n = 6). The study took approximately 60 minutes to complete, and participants received £10.00 (approximately €11.50 or US$12.50) as compensation.

Procedure

Participants applied for a simulated management traineeship role (see Appendix B) and completed an AVI via a platform developed for the study, a personality questionnaire, a cognitive ability assessment, and provision of demographic information. In the AVI, participants answered two generic questions commonly used in hiring (e.g., "What would you consider among your greatest strengths and weaknesses as an employee?") and four personality traits-related questions targeting Honesty-Humility, Extraversion, Agreeableness, and Conscientiousness based on the HEXACO framework (Lee & Ashton, 2018; Appendix C). For the present study, the analysis only focused on the two questions designed to assess Honesty-Humility and Extraversion. These questions followed a past-behavioral format, with the Honesty-Humility item emphasizing two facets, such as Greed Avoidance and Modesty, and the Extraversion item targeting Liveliness and Sociability. Participants were instructed to respond to a question within 1-2 minutes. Responses shorter than 30 seconds prompted a request for elaboration, while those exceeding 150 seconds triggered a prompt to conclude. All participants' responses could be recorded only once, without any opportunity to delete or resend a response. After completing the AVI, participants reviewed the six interview questions and then answered

two open-ended items asking what they believed each question was intended to assess. This task aimed to measure participants' ability to identify criteria (ATIC; Kleinmann et al., 2011).

Measures

Self-reports

Self-reports. Participants accomplished the HEXACO-100 personality inventory (Lee & Ashton, 2018), which includes 100 self-reported items across six domains: Honesty-Humility, Emotionality, Extraversion, Agreeableness, Conscientiousness, and Openness to Experience. Each domain consists of four facets, with four items measuring each facet, plus four additional items assessing the facet of Altruism. Responses were rated on a 5-point scale (1 = *Strongly Disagree*; 5 = *Strongly Agree*). Since this study only paid attention to Extraversion and Honesty-Humility scores, the reliability of the domains ranged from .84 (Honesty-Humility) to .90 (Extraversion), as indicated by Cronbach's alpha values (Table 2).

Observer reports

Personality. A group of 12 raters (five men, seven women; average age = 26.72 with 2 Ph.D., 8 Master's, 2 Bachelor's) assessed participants' personalities after completing 9 hours of training as described in Koutsoumpis et al. (2024). Due to high workload, there was a difference in the number of participants rated by each rater. Raters evaluated different batches of video responses: one for generic questions and one for personality-related questions. After watching all responses per batch, the raters provided their scores. For the present study, the focus was on the personality ratings of Extraversion and Honesty-Humility from the second batch (i.e., two of the four personality questions), so each participant received four independent personality ratings. The ratings were collected using a Behavioral Anchored Rating Scale (BARS; Appendix D) with 4 items per personality domain and a 5-point scale (1 = *Very low*; 5 = *Very high*) for responses,

allowing decimal precision. Personality domain scores were calculated by averaging the ratings across facets. The inter-rater reliability ranged from moderate ($ICC_{1,4}$ = .66– Honesty- Humility) to high ($ICC_{1,4}$ = .82 - Extraversion), see Table 2.

Ability to identify criteria (ATIC).

After completing the interview, participants were shown the interview questions again and, for each question, answered two open-ended questions in writing via the same online platform, including: (a) what kind of qualities or traits they thought each question was meant to assess, and (b) why they thought that (see Appendix E).

Subsequently, a group of four trained recruiters (three women and one man; average age 25.75 years old; *SD* = 7.54) with diverse ethnic backgrounds (Netherlands, Estonia, Croatia, and Vietnam) independently coded and evaluated the ATIC responses. All raters were bachelor's students in Psychology who completed a 7- hour training program that covered: (1) an overview of personality traits, including definitions, facets, domains, and behavioral indicators; (2) how to construct an ATIC rating form, based on guidelines from Kleinmann et al. (2011) and Melchers et al. (2012); and (3) how to evaluate participant responses using the rating form (see Appendix F).

Due to the high workload, participants were randomly assigned to two pairs of raters, ensuring that each participant received two independent ATIC ratings. Raters reviewed a set of four personality-related questions and evaluated the open-ended responses using a rating form (Appendix F) with a 4-point scale: 0 = No or poor understanding, 1= Moderate understanding, 2= Good understanding, 3= Excellent understanding, with half-point increments allowed (e.g., '2.5' was possible). The overall ATIC score was calculated as the average of the two raters' scores across the four questions, and particularly ATIC scores related to Extraversion and

Honesty-Humility were used in this study. The inter-rater reliability ranged from notably poor ($ICC_{1,2}$ = .07– Extraversion) to good ($ICC_{1,2}$ = .61– Honesty- Humility), also see Table 2.

Statistical Analyses

All analyses were conducted in SPSS (Version 28), using the PROCESS macro (Version 4.2; Hayes, 2022) for moderation tests. This study employed a cross-sectional correlational design to examine whether ATIC moderated the relationship between self- and observer-reported Extraversion and Honesty-Humility. While self-reported Extraversion and Honesty-Humility served as independent variables, observer-reported ratings are treated as dependent variables. ATIC was tested as a moderator, and no covariates were included. Prior to analysis, data were screened for outliers and missing data. No extreme outliers were identified based on Cook's Distance and Mahalanobis Distance, and no missing data were present due to complete responses. Self-reported traits and ATIC scores were mean-centered to reduce multicollinearity before creating interaction terms. All relevant assumptions of linear regression, including linearity, normality of residuals, independence of observations, and absence of multicollinearity, were tested and met. Descriptive statistics and Pearson correlations were calculated. Hypotheses 1 and 2 were tested using linear regression with interaction terms (PROCESS Model 1).

Results

Descriptive Statistics

Descriptive statistics for the main study variables are reported in Table 1. The average participant age was 36.69 years (*SD* = 11.92), and participants were generally well-educated (M = 3.61, SD = 1.03). The sample was predominantly White (72.3%), with 11.3% identifying as Black or African American. Self-reported Honesty-Humility scores ranged from 1.56 to 5.00 (*M* = 3.65, *SD* = .64), while self-reported Extraversion ranged from 1.13 to 5.00 (*M* = 3.23, *SD* =

.73). Observer-reported Honesty-Humility and Extraversion had means of 3.26 (*SD* = .43) and 3.02 (*SD* = .51), respectively. The average ATIC score was 1.48 (*SD* = .79) for Honesty-Humility and 1.84 (*SD* = .53) for Extraversion.

Correlations and the baseline of self-other agreement

Pearson correlations between key variables are displayed in Table 2. Self- and observer-reported Honesty-Humility were moderately correlated ($r = .22$, $p < .01$), while the correlation for Extraversion was stronger ($r = .42$, $p < .01$). A Steiger's Z-test confirmed that the self–other agreement was significantly weaker for Honesty-Humility than for Extraversion ($Z = -4.25$, $p < .001$), consistent with the SOKA model. In addition, ATIC was positively associated with observer ratings for both traits (Honesty-Humility r =.29, p<.01, and Extraversion r =.10, p<.05). Self-reported Honesty-Humility and Extraversion were also slightly correlated ($r = .10$, $p < .01$).

Self–Observer Agreement for Honesty-Humility and the Role of ATIC (Hypothesis 1)

A linear regression analysis was conducted to examine whether ATIC moderated the relationship between self- and observer-rated Honesty-Humility. All predictors, including self-reported Honesty-Humility and ATIC scores for Honesty-Humility, were mean-centered, and an interaction term was computed using these centered variables.

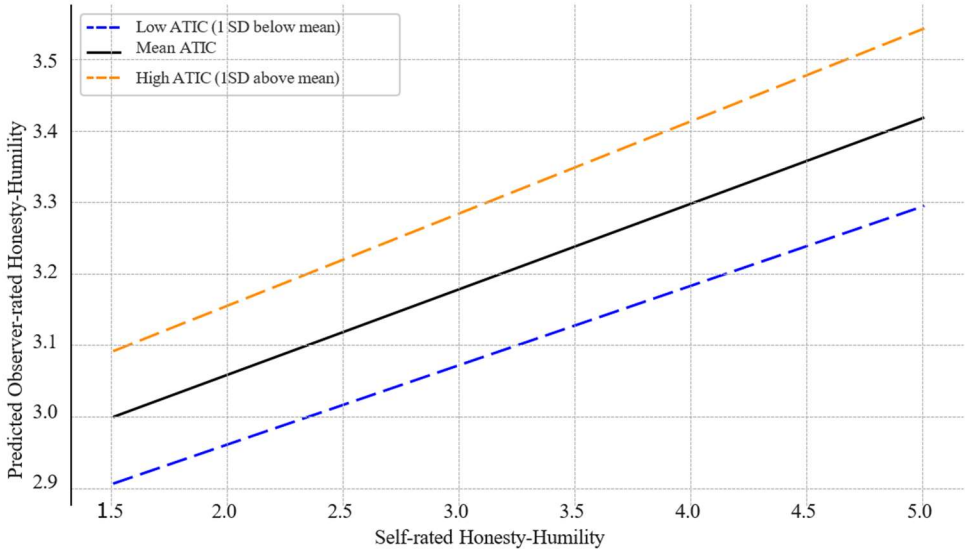
The overall model was significant, $F(3, 642) = 27.46$, $p < .001$, explaining approximately 11% of the variance in observer-reported Honesty-Humility ($R^2 = .11$; see Table 3). Higher self-reported Honesty-Humility was associated with higher observer ratings ($b = .12$, $p < .001$), suggesting that individuals who viewed themselves as more honest and modest were also perceived that way by external observers. Similarly, higher ATIC scores predicted higher observer-rated Honesty-Humility ($b = .14$, $p < .001$), indicating that individuals who better

identified the evaluative criteria in the assessment were also rated as more honest and humble by others.

However, the interaction effect between self-reported Honesty-Humility and ATIC on observer ratings was insignificant (b = .01, p = .71), indicating no moderation effect. As illustrated in Figure 2, the slopes are roughly parallel across ATIC levels, suggesting that ATIC did not influence the strength of self–observer agreement for Honesty-Humility. In other words, these results indicate no evidence for a moderation effect of ATIC on agreement in Honesty-Humility ratings. Therefore, hypothesis 1 was not supported.

Figure 2

Interaction Between Self-reported Honesty-Humility and ATIC on Observer-reported Honesty-Humility



Note. Predicted values are plotted at low (−1 SD), mean, and high (+1 SD) levels of ATIC. The roughly parallel lines indicate that ATIC did not significantly moderate the self–observer relationship for Honesty-Humility, consistent with the non-significant interaction found in the analysis.

Self–Observer Agreement for Extraversion and the Role of ATIC (Hypothesis 2)

Similar steps were also conducted for Extraversion, such as mean-centering self-reported Extraversion and ATIC scores and computing their interaction term. A linear regression analysis was performed to investigate whether ATIC moderates the agreement of self-other Extraversion.

The findings indicated that the overall model was significant, $F(3, 642) = 48.88$, $p < .001$, accounting for 19% of the variance in observer-reported Extraversion ($R^2 = .19$; see Table 4). Self-reported Extraversion was positively correlated with observer-rating ($b = .27$, $p < .001$), indicating that individuals who described themselves as more social and energetic were also perceived as such by observers. Additionally, ATIC showed a smaller but significant positive effect ($b = .07$, $p = .03$), suggesting that those with greater insight into the evaluative context were more likely to be rated as extraverted by others.

Notably, a significant interaction effect between self-Extraversion and ATIC on the observer-reported Extraversion ($b = .11$, $p = .01$) was found, indicating that the relationship between self- and observer-reported Extraversion depends on individuals' ATIC scores. In other words, the alignment between self- and observer ratings became stronger as ATIC increased. This pattern is illustrated in Figure 3, which shows steeper regression slopes at higher levels of ATIC.

To further explore this interaction effect, a simple slopes analysis was conducted at low (−1 SD), mean, and high (+1 SD) levels of ATIC. The results indicated that the relationship between self- and observer-reported Extraversion was increasingly stronger as ATIC scores increased, consistent with a significant moderation effect. Specifically, at low ATIC (−1 SD = 1.31), the positive self- and observer-reported Extraversion alignment was significant ($b = .21$, $p < .001$); at the mean level of ATIC (M=1.84), this positive relationship was stronger ($b = .27$, $p <$

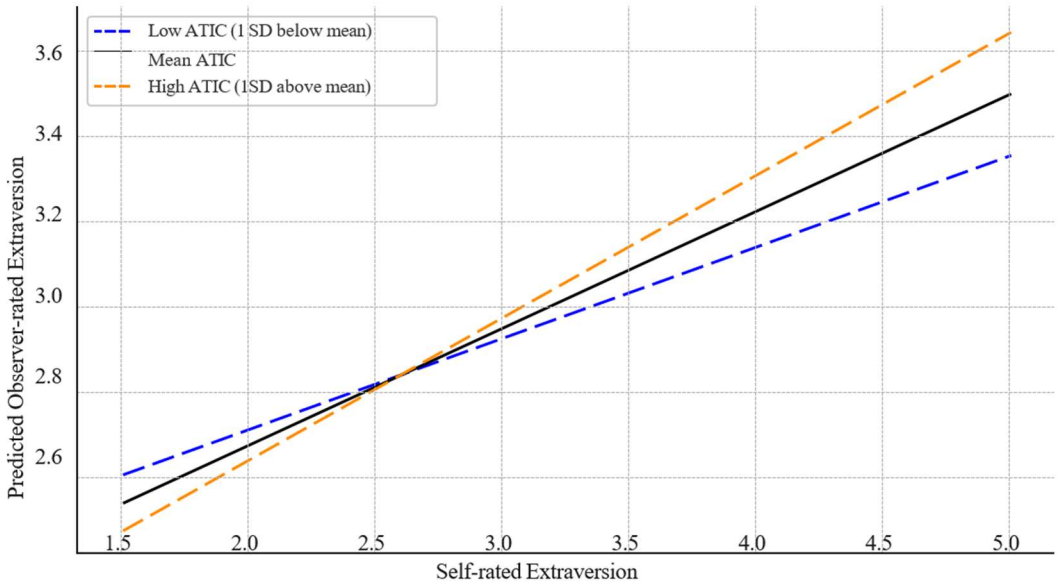
.001); and at high ATIC (+1 SD = 2.37), the association became even stronger (b = .33, $p < .001$; see Table 5).

Moreover, a Johnson-Neyman analysis also supported these findings, showing that the self–observer agreement was statistically significant at ATIC values above .56, a threshold exceeded by approximately 98.30% of participants (see Table 5).

Taken together, these findings highlight ATIC as a meaningful moderator that enhances self–observer agreement for Extraversion, supporting Hypothesis 2.

Figure 3

Interaction Between Self-rated Extraversion and ATIC on Observer-rated Extraversion



Note. Predicted values are shown at low (−1 SD), mean, and high (+1 SD) levels of ATIC. Greater slopes at higher ATIC levels indicate stronger alignment between self- and observer-reported Extraversion.

Discussion

The present study examined the alignment between self- and observer-reported personality traits in AVIs, with a specific focus on Extraversion and Honesty-Humility from the

HEXACO model. The central research question asked if the relationship between self-reported and observer-reported personality traits depends on whether a candidate correctly identifies what assessors measure in asynchronous video interviews. Findings revealed that self-reported Honesty-Humility and Extraversion were each positively associated with observer-reported levels of those traits, suggesting that when participants rated themselves as high in Honesty-Humility or Extraversion, observers tended to give similar ratings. Additionally, greater ATIC scores predicted higher observer-reported ratings of Honesty-Humility and Extraversion, indicating that participants who more accurately identified the evaluation criteria were perceived more positively by observers. Consistent with the SOKA model, the study also found that self–observer agreement was significantly weaker for Honesty-Humility than Extraversion. Notably, a significant interaction effect between self-reported trait and ATIC was found only for Extraversion. Specifically, self-other agreement for Extraversion was significantly positive**,** and this correlation was stronger when participants correctly identified the criteria being assessed. This possibly means that the more participants identified the evaluated criteria, the more they used trait-relevant behavioral information in their interview responses, and this trait-relevant information contributed to higher self-other agreement. Accordingly, the findings partially supported the hypotheses: while ATIC moderated the relationship between self- and observer-reported Extraversion, it did not moderate the relationship for Honesty-Humility.

From a theoretical perspective, the observed moderation effect of ATIC on Extraversion further supported the Realistic Accuracy Model (RAM). According to RAM, accurate personality judgments require the target to express relevant behavioral cues, which must be available, detected, and correctly utilized by observers. Candidates with high ATIC for Extraversion, as a good target, are more likely to present such cues, thus facilitating more

accurate observer judgments. Consequently, the finding, the agreement of self-observer reported Extraversion was stronger with a higher ATIC score, implies that Extraversion is a good trait and ATIC is a moderator of the accuracy function.

In contrast, the absence of a similar moderation effect for Honesty-Humility contradicts the study's first hypothesis and prior theoretical expectations. Considering inter-rater reliability, the ATIC ratings for Honesty-Humility had better agreement between raters ($ICC_{1,2} = .61$), which suggests that the lack of a moderation effect there is more trustworthy and not just due to measurement problems. Some possible explanations can account for this finding. First, it may be attributable to the inherent nature of Honesty-Humility. According to the SOKA model, Honesty-Humility is a low-visibility trait, so it may not manifest overt behavioral cues during brief, structured interviews. As a result, observers, even being attentive or knowledgeable, may struggle to detect and accurately assess such traits. Second, RAM may be less applicable to such traits, as its core processes rely on observable behaviors (good traits) that are hard to see or understand for this kind of trait, like Honesty-Humility. In other words, Honesty- Humility is not a good trait that is easy to detect. Third, methodological factors may have played a role. The asynchronous video interview format may have further restricted participants' ability to display traits or behaviors like greed avoidance or modesty. Consequently, even participants with high ATIC may have lacked the means or opportunities to effectively signal Honesty-Humility, explaining the absence of a moderation effect.

Another important finding, comparing the self-other agreement strength between two traits of differing visibility, Extraversion and Honesty-Humility, also reinforces key assumptions of the SOKA model regarding trait visibility in personality judgment. The SOKA model predicts that highly visible traits, such as Extraversion, are more easily inferred by observers, which was

supported by the stronger correlation between self- and observer-reported Extraversion. This trait's behavioral cues, such as sociability and liveliness, are likely more readily observed in structured AVI responses. Conversely, the relatively lower self–observer agreement for Honesty-Humility aligns with its classification as a low-visibility trait. Facets of this trait, such as modesty or greed avoidance, may be difficult to infer from brief, structured responses, especially when social desirability or situational constraints limit the expression of relevant behaviors. This pronounced difference between the two traits emphasizes a critical limitation of AVIs, which is that low-visibility traits are inherently more difficult to judge accurately in such structured, brief settings.

From a practical perspective, these findings have clear implications for personnel selection using AVIs. The results suggest that AVIs can effectively assess certain personality traits, especially visible ones like Extraversion, when candidates understand what is being assessed. This has practical value for organizations seeking scalable, cost-efficient methods of evaluating candidate fit. However, the weaker agreement for Honesty-Humility implies that AVIs may not be sufficient on their own to evaluate traits that are harder to observe, regardless of candidates' awareness of evaluation criteria. In such cases, complementary assessment methods, such as integrity tests or structured behavioral simulations, may provide more reliable insights.

Moreover, these findings also imply that improving candidates' ATIC could enhance assessment validity, particularly for externally observable traits. Structured pre-interview briefings, clearer question framing, or feedback mechanisms may help candidates better understand what is being evaluated, leading to more targeted and interpretable responses. However, the limited success of ATIC in improving Honesty-Humility agreement also indicates

that better cue availability alone may not be sufficient for people to judge this trait accurately; traits with inherently low visibility might require richer, more dynamic assessment contexts like probing questions, extended response times, or role-plays that can reveal deeper moral character.

In addition to its valuable findings, this study also presents certain limitations. First, the inter-rater reliability for ATIC ratings was notably poor for Extraversion ($ICC_{1,2} = .07$), indicating substantial variance in how raters judged candidates' understanding of the criteria. This low reliability may have caused measurement errors, which could have either weakened or exaggerated the moderation effect. Although the effect was in the expected direction and statistically significant, the poor measurement makes it less reliable. Second, the study was carried out in a simulated hiring situation, which might have influenced how seriously participants took the interview or how honestly they responded. Third, most participants were White, well-educated, and fluent in English, which may limit the generalizability of the findings to more diverse populations. Fourth, the structured format of the interview questions may have made it difficult for participants to fully show certain traits, especially Honesty-Humility, which often shows up in more subtle or situational ways. Finally, the study focused only on two dimensions of the HEXACO model, Extraversion and Honesty-Humility, leaving out other important traits such as Emotionality, Agreeableness, Conscientiousness, and Openness to Experience. This narrow focus limits insights into how ATIC and self–observer agreement might operate across a broader range of personality characteristics.

Despite these limitations, the study has notable strengths. It combines observer ratings, self-report measures, and a theoretically grounded moderator (ATIC) in a large, demographically diverse online sample (N=646). The use of trained raters and standardized rating procedures strengthens the study's methodological quality. By integrating the RAM and SOKA frameworks,

the study also offers a theoretically rich lens to interpret self–observer agreement in a modern selection process.

Some directions for future research emerge from the findings of this current study. The first is to refine the measurement of ATIC, particularly for high-visibility traits. This may include involving more raters from diverse professional backgrounds, such as those with Bachelor's, Master's, and PhD degrees, offering more detailed and intensive training, and providing more specific rating guidelines. Moreover, if feasible, future studies should conduct assessments in a controlled environment with minimal distractions or disturbances, rather than free or informal settings (e.g., during commuting, in lecture halls, or at home), to reduce superficial or subjective judgements and enhance the accuracy of observer ratings. Another valuable direction for future research would be to investigate how AVIs might be adapted to better reveal lower-visibility traits, for instance, through probing questions or extending response opportunities. Moreover, future work could experimentally manipulate ATIC, through feedback or priming, to determine its causal role in shaping observer accuracy. Another promising avenue is to compare human rater judgments with machine-learning-based assessments in AVIs to evaluate how well AI systems detect personality traits relative to human observers. Finally, future work should extend to study other HEXACO dimensions, such as Emotionality, Agreeableness, and Conscientiousness, to explore whether visibility and judgment accuracy differ across a wider range of personality traits.

In sum, this study contributes to a deeper theoretical understanding of how personality traits are perceived in AVIs and highlights both the potential and constraints of this assessment format. While AVIs appear suitable for evaluating visible traits like Extraversion, they are less effective for low-visibility traits like Honesty-Humility. The candidate's ability to identify

criteria enhances observer accuracy for observable traits, but not for low-visibility traits. These findings suggest that AVIs can be a useful tool in hiring, but they need to be carefully designed and should be combined with other methods to make sure the assessment is fair and accurate.

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.

Chamorro-Premuzic, T., & Furnham, A. (2010). *The psychology of personnel selection*. Cambridge University Press.

Connelly, B. S., & Ones, D. S. (2010). An other perspective on personality: Meta-analytic integration of observers' accuracy and predictive validity. *Psychological Bulletin, 136*(6), 1092–1122. https://doi.org/10.1037/a0021212

De Vries, R. E., Lee, K., & Ashton, M. C. (2008). The Dutch HEXACO Personality Inventory: Psychometric properties, self-other agreement, and relations with psychopathy among low and high acquaintanceship dyads. *Journal of Personality Assessment*, *90*(2), 142–151. https://doi-org.tilburguniversity.idm.oclc.org/10.1080/00223890701845195

De Vries, R. E., & van Gelder, J.-L. (2015). Explaining workplace delinquency: The role of Honesty–Humility, ethical culture, and employee surveillance. *Personality and Individual Differences, 86*, 112–116. https://doi.org/10.1016/j.paid.2015.06.008

Dietl, E., & Kombeiz, O. (2025). Acquaintance ratings of honesty-humility and their relationship with extra-role behaviors. *Personality and Individual Differences, 233*, 112876. https://doi.org/10.1016/j.paid.2024.112876

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.tilburguniversity.idm.oclc.org/10.1111/ijsa.12372

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102*(4), 652–670. https://doi.org/10.1037/0033-295X.102.4.652.

Huffcutt, A. I., Culbertson, S. S., & Weyhrauch, W. S. (2013). Employment interview reliability: New meta-analytic estimates by structure and format. *International Journal of Selection and Assessment, 21*(3), 264–276. https://doi.org/10.1111/ijsa.12036

Koutsoumpis, A., Ghassemi, S., Oostrom, J.K., Holtrop, D.J., Van Breda, W., Zhang, T. & De Vries, R.E. (2024). Beyond traditional job interviews: Psychometric analysis of asynchronous job 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., Zhang, T., Oostrom, J. K., Holtrop, D., & De Vries, R.E. (2025). Open dataset AVI-6: Annotated asynchronous video interviews. https://doi.org/10.17605/OSF.IO/XTUYQ

Kleinmann, M., Ingold, P. V., Lievens, F., Jansen, A., Melchers, K. G., & König, C. J. (2011). A different look at why selection procedures work. *Organizational Psychology Review, 1*(2), 128–146. https://doi.org/10.1177/204138661038700

Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, *25*(5), 543-556. https://doi.org/10.1177/1073191116659134

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

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. https://doi.org/10.1093/oxfordhb/9780190912529.001.0001

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), Article 100789. https://doi.org/10.1016/j.hrmr.2020.100789

McCrae, R. R. (1994). Openness to experience: Expanding the boundaries of Factor V. *European Journal of Personality, 8*(4), 251–272. https://doi.org/10.1002/per.2410080404

Oh, I.-S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology, 96*(4), 762–773. https://doi.org/10.1037/a0021832

Oliphant, G. C., Hansen, K., & Oliphant, B. J. (2008). A Review of a Telephone-Administered Behavior-Based Interview Technique. *Business Communication Quarterly*, *71*(3), 383-386. https://doi.org/10.1177/1080569908321429

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology, 70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Pletzer, J. L., Bentvelzen, M., Oostrom, J. K., & de Vries, R. E. (2019). A meta-analysis of the relations between personality and workplace deviance: Big Five versus HEXACO. *Journal of Vocational Behavior, 112*, 369–383. https://doi.org/10.1016/j.jvb.2019.04.004

Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2021). Revisiting meta-analytic estimates of validity in personnel selection: Addressing systematic overcorrection for restriction of

range. *Journal of Applied Psychology, 107*(11), 2040–2068. https://doi.org/10.1037/apl0000994

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

Zettler, I., & Hilbig, B. E. (2010). Honesty-humility and a person-situation interaction at work. *European Journal of Personality, 24*(7), 569–582. https://doi-org.tilburguniversity.idm.oclc.org/10.1002/per.757

Appendix A

Analysis Results of The Study

Table 1

Means and Standard Deviations, Percentage of Key Variables (N = 646)

Variable	Mean (M)	SD	Min	Max	Percent
Demographics					
Age (years)	36.69	11.92	18.00	67.00	-
Education	3.61	1.03	1.00	7.00	-
White (%)	-	-	-	-	72.3
Black or African American	-	-	-	-	11.3
Self-reported Personality					
Honesty-Humility	3.65	0.64	1.56	5.00	-
Extraversion	3.23	0.73	1.13	5.00	-
Observer-reported Personality					
Honesty-Humility	3.26	0.43	2.01	4.68	-
Extraversion	3.02	0.51	1.51	4.68	-
ATIC Scores					
Honesty-Humility	1.48	0.79	0.00	3.00	-
Extraversion	1.84	0.53	0.25	3.00	-

Note. Descriptive statistics include mean (M), standard deviation (SD), and observed range for all key variables. ATIC = Ability to Identify Criteria., (-) indicate data not applicable

Table 2

Pearson Correlations and reliability estimates across all variables (N = 646)

Variable	1	2	3	4	5	6
1. Self-rated Honesty-Humility	**.84**					
2. Self-rated Extraversion	.10**	**.90**				
3. Observer-rated Honesty-Humility	.22**	.06	**.66**			
4. Observer-rated Extraversion	.08*	.42**	.15**	**.82**		
5. ATIC score for Honesty-Humility	.14**	.02	.29**	.09*	**.61**	
6. ATIC score for Extraversion	.01	.05	.12**	.10*	.31**	**.07**

Note: p < .05*, p < .01**. ATIC =Ability to Identify Criteria, boldfaced on the diagonal are Cronbach's alpha (for self-reports) or intra-class correlation coefficients $ICC_{1,4}$ (for observer reports) and $ICC_{1,2}$ (for ATIC score).

Table 3

Regression analysis of mean-centered self-reported and centered ATIC score for Honesty- Humility on observer ratings of Personality

	Self-reported Honesty-Humility, ATIC scores								
	Summary		ANOVA			Coefficients			
	R^2	ΔR^2	df	F	p	b	β	t	p
	.114	.110	3	27.46	<.001				
Constant						3.26		203.27	.000
Self-reported Honesty-Humility						.12	.18	4.74	<.001
ATIC scores of Honesty-Humility						.14	.26	6.98	<.001
Interaction term						**.01**	.01	.37	**.71**

Note. N= 646, dependent variable= observer-reported Honesty-Humility. This table shows the regression analysis to assess whether the self-observer reported agreement for Honesty-Humility will depend on ATIC, through the interaction effect.

Table 4

Regression analysis of mean-centered self-reported and centered ATIC scores for Extraversion on observer ratings of Personality

	Self-reported Extraversion, ATIC scores								
	Summary		ANOVA			Coefficients			
	R^2	ΔR^2	df	F	p	b	β	t	p
	.19	.18	3	48.88	<.001				
Constant						3.02		167.21	.000
Self-reported Extraversion						.27	.40	11.03	<.001
ATIC scores of Extraversion						.07	.08	2.14	.03
Interaction term						**.11**	.09	2.52	**.01**

Note. N= 646, dependent variable= observer-reported Extraversion. This table shows the regression analysis to assess whether the self-observer reported agreement for Extraversion will depend on ATIC, through the interaction effects.

Table 5

Insights of interaction effects between ATIC and self-reported Extraversion

Conditional effects of the focal predictor at values of the moderator(s):

ATIC_X	Effect	Se	t	P	LLCI	ULCI
1.3142	**.2145**	.0368	5.8237	**.0000**	.1422	2869
1.8433	**.2742**	.0249	11.0260	**.0000**	2254	2254
2.3724	**.3339**	.0317	10.5389	**.0000**	.2717	.3961

Moderator value(s) defining Johnson-Neyman significance region(s):

Value	% below	% above
.5605	1.7028	**98.2972**

Conditional effect of focal predictor at values of the moderator:

ATIC_X	Effect	Se	t	p	LLCI	ULCI
.2500	.0945	.0790	1.1955	.2323	-.0607	.2497
.3875	.1100	.0732	1.5026	.1334	-.0338	.2537
.5250	.1255	.0674	1.8612	.0632	-.0069	.2579
.5605	.1295	.0660	1.9637	**.0500**	.0000	.2590

.6625	.1410	.0617	2.2842	.0227	.0198	.2622
.8000	.1565	.0561	2.7883	.0055	.0463	.2668
.9375	.1720	.0507	3.3953	.0007	.0725	.2715
1.0750	.1876	.0454	4.1326	.0000	0984	.2767
1.2125	.2031	.0403	5.0333	.0000	.1238	.2823
1.3500	.2186	.0357	6.1297	.0000	.1486	.2886
1.4875	.2341	.0315	7.4348	.0000	.1723	.2959
1.6250	.2496	.0281	8.8972	.0000	.1945	.3047
1.7625	.2651	.0251	10.3304	.0000	.2147	.3155
1.9000	.2806	.0246	11.3989	.0000	.2323	.3290
2.0375	.2961	.0251	11.8035	.0000	.2469	.3454
2.1750	.3117	.0270	11.5447	.0000	.2586	.3647
2.3125	.3272	.0301	10.8819	.0000	.2681	.3862
2.4500	.3427	.0340	10.0832	.0000	.2759	.4094
2.5875	.3582	.0385	9.3046	.0000	.2826	.4338
2.7250	.3737	.0434	8.6078	.0000	.2885	.4590

2.8625	.3892	.0486	8.0058	.0000	.2938	.4847
3.0000	.4047	.0540	7.4921	.0000	.2987	.5108

Note. Value of ATIC: mean = 1.8433, -1SD =1.3142, +1SD=2.3724, Value of ATIC= 0.5605, a value that the effect of the self-reported Extraversion on the observer ratings is right at the threshold of statistical significance (p = .05), 98.2972% = ATIC values are above 0.5605 for about 98.30% of participants.

Appendix B

Management Traineeship at Aurélien

Do you want to set yourself apart from the crowd? Adding a management traineeship at 'Aurélien' to your resume is the best way to do this. You can participate in our professional trainee programs for up to 18 months and payment is guaranteed! For all traineeships at 'Aurélien' you will receive a competitive salary.

What is a traineeship?

A traineeship gives you real work experience with real responsibilities for a specified period of time. You will follow an intensive training program in your field, so you must be a highly motivated and hardworking person in order to qualify. Our goal is to create an innovative and dynamic environment - be part of this journey!

We are looking for you, if …

- You have a strong desire to develop yourself into a competitive professional.
- You enjoy being around and working with people.
- You don't mind taking the lead and have strong leadership characteristics.
- You have high ethical standards and personal integrity.
- You are able to work and cooperate with others.

Duration and Compensation

Traineeships are usually full-time for a period of 12 to 18 months. However, shorter programs are also available from 5 months on. There are no specific start dates for a traineeship at 'Aurélien', you can start throughout the entire year. Upon acceptance, we will offer a traineeship with a duration of 5-18 months.

At Aurélien we pride ourselves on using modern personnel selection methods. Applying for our traineeships is easy. The process involves completing a video interview and some questionnaires.

APPLY NOW

Appendix C

Interview questions and order of presentation

Order	Interview question
1	What would you consider among your greatest strengths and weaknesses as an employee? [Generic question]
2	How would your best friend describe you? [Generic 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]

Appendix D

Behavioral Anchored Rating Scale (BARS) for observer reports of HEXACO

Domain Facet	General Description	Description low	Very low	Low	Average	High	Very high	Description high
Honesty-Humility								
Sincerity	Being genuine in interpersonal relations.	Will flatter others or pretend to like them in order to obtain favors.	1	2	3	4	5	Are unwilling to manipulate others and are genuine in interpersonal relations.
Fairness	Tendency to avoid fraud and corruption.	Is willing to gain by cheating or stealing.	1	2	3	4	5	Is unwilling to take advantage of other individuals or of society at large.
Greed avoidance	Interest, or lack thereof, in	Enjoys and displays wealth, privileges, and status.	1	2	3	4	5	Not motivated by monetary or social-status considerations.

	wealth, luxury, and status							
Modesty	Being modest and unassuming.	Considers oneself as superior and entitled to privileges that others do not have.	1	2	3	4	5	Considers oneself as ordinary person without any claim to special treatment.
Emotionality								
Fearfulness	Experiencing fear.	Feel little fear of injury and are relatively tough, brave, and insensitive to physical pain.	1	2	3	4	5	Is sensitive to pain, 'soft', and strongly inclined to avoid physical harm.
Anxiety	General worrisomeness and stress,	Feels little stress in response to difficulties.	1	2	3	4	5	Becomes preoccupied even by relatively minor problems.

Dependence	Need for emotional support from others.	Feels self-assured and is able to deal with problems without any help or advice.	1	2	3	4	5	Wants to share their difficulties with those who will provide encouragement and comfort.
Sentimentality	Feeling strong emotional bonds with others.	Feels little emotion when saying good-bye or in reaction to the concerns of others.	1	2	3	4	5	Feels strong emotional attachments and empathic sensitivity to the feelings of others.
Extraversion								
Social self-esteem	Positive self-image in social situations.	Feels worthless and unpopular.	1	2	3	4	5	Is content with self and inclined to attribute positive qualities to self.

Social boldness	Ease and confidence in different social situations.	Feels shy and uncomfortable in a leadership position or when required to speak in public.	1	2	3	4	5	Has no problem addressing strangers and speaking in a group.
Sociability	Fun in conversations, social interactions and at parties.	Prefers solitary activities and does not seek out conversations with others.	1	2	3	4	5	Enjoys conversations, (unexpected) visitors, festivities and the company of others.
Liveliness	Enthusiasm and energy.	Feels less cheerful and lively.	1	2	3	4	5	Is optimistic and positive.
Agreeableness								

Forgivingness	Feeling trust and liking towards those who may have caused harm.	Tends to 'hold a grudge' against those who have offended them.	1	2	3	4	5	Is usually ready to trust others again and to re-establish friendly relations after having been treated badly.
Gentleness	Being mild and lenient with others.	Tends to be critical when evaluating others.	1	2	3	4	5	Is reluctant to judge others harshly.
Flexibility	Compromise and cooperate with others.	Is stubborn, unaccommodating, and willing to argue with others.	1	2	3	4	5	Avoids arguments and accommodates others' suggestions, even when these may be unreasonable.
Patience	Remain calm.	Quickly loses temper and feels anger.	1	2	3	4	5	Has a high threshold for feeling or experiencing anger.

Conscientiousness								
Organization	Have an organized and tidy approach towards the physical environment.	Handles things and tasks in a sloppy and chaotic way.	1	2	3	4	5	Tends to organize things and perform tasks in a structured way.
Diligence	Willingness to work hard.	Has little self-discipline in performing a task and is not highly motivated to perform well.	1	2	3	4	5	Has a strong work ethic and is inclined to make an optimal effort to bring tasks to a good result.
Perfectionism	Thoroughness and rigor in	Has a higher tolerance for errors and less attention to detail.	1	2	3	4	5	Tends to check work for errors and attempts

	dealing with details.							possible improvements.
Prudence	Makes informed decisions and is not guided by sudden urges.	Responds impulsively and considers the consequences of his/her actions to a lesser extent.	1	2	3	4	5	Has a high degree of self-control and tends to consider several options carefully.
Openness to Experience								
Aesthetic appreciation	Enjoyment of beauty in art and nature.	Does not become absorbed in works of art or in natural wonders.	1	2	3	4	5	Has a strong appreciation of various art forms and of natural wonders.

Inquisitiveness	Tendency to seek information and knowledge.	Has little curiosity about the natural or social sciences.	1	2	3	4	5	Is curious, reads widely and is interested in discovering the world through traveling.
Creativity	Innovation and experimentation.	Has little inclination for original thought, innovation, and artistic expression.	1	2	3	4	5	Actively seeks new solutions to problems and expresses themselves in art.
Unconventionality	Acceptance of the unusual.	Avoids eccentric or non-conforming persons.	1	2	3	4	5	Is receptive to ideas that might seem strange or radical.

Appendix E

Ability to identify criteria (ATIC) instructions to participants

Instructions to participants	You just participated in a one-way interview. You might have wondered what characteristics the researchers want to assess with this interview. We would like to know your best guess. Below, you see the interview questions again. For each question, please write down: 1) any specific characteristics you think that the researchers are interested in measuring, and 2) a brief rationale to your response. As an example, in the question: "*What would you consider among your greatest strengths and weaknesses as an employee?*" Some example **characteristics** could include: '*self-knowledge'*, '*honesty'*, '*self-image'*, etc. The **rationale** could be: '*Tried to balance between strengths and weaknesses so that I don't look arrogant or self-effacing'*. Please, include at least 10 characters in your response.

Appendix F

Ability to identify criteria (ATIC) rating form

Numeric value	Brief description	Elaborate description	Percentage representation
0	No or poor understanding	Off-topic; The participant identified an incorrect personality trait.	0-24% correct criterion
1	Moderate understanding	The participant partially identified the correct personality trait, but also identified incorrect trait(s).	25-49% correct criterion
2	Good understanding	The participant identified the correct trait. However, either did not identify the correct facet of did not identify the correct direction of the trait (e.g., high end-point of a trait or facet).	50-74% correct criterion
3	Excellent understanding	The participant correctly identified the personality trait or facet, as well as the correct (high) end-point of the trait or facet.	75-100% correct criterion

Note. rating form with a 4-point scale (0-3) with half-point increments allowed (e.g., “2.5” is possible).