

Härtel, Tobias M.; Schuler, Benedikt A.; Back, Mitja D.

Article — Published Version

'LinkedIn, LinkedIn on the screen, who is the greatest and smartest ever seen?': A machine learning approach using valid LinkedIn cues to predict narcissism and intelligence

Journal of Occupational and Organizational Psychology

Suggested Citation: Härtel, Tobias M.; Schuler, Benedikt A.; Back, Mitja D. (2024) : 'LinkedIn, LinkedIn on the screen, who is the greatest and smartest ever seen?': A machine learning approach using valid LinkedIn cues to predict narcissism and intelligence, Journal of Occupational and Organizational Psychology, ISSN 2044-8325, Vol. 97, Iss. 4, pp. 1572-1602, <https://doi.org/10.1111/joop.12531>

This Version is available at:

<https://hdl.handle.net/10419/313767>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



<http://creativecommons.org/licenses/by-nc/4.0/>

RESEARCH ARTICLE

‘LinkedIn, LinkedIn on the screen, who is the greatest and smartest ever seen?’: A machine learning approach using valid LinkedIn cues to predict narcissism and intelligence

Tobias M. Härtel¹  | Benedikt A. Schuler²  | Mitja D. Back³ 

¹School of Business Administration and Economics, Osnabrück University, Osnabrück, Germany

²Institute of Management and Strategy, University of St. Gallen, St. Gallen, Switzerland

³Department of Psychology, University of Münster, Münster, Germany

Correspondence

Tobias M. Härtel, School of Business Administration and Economics, Osnabrück University, Rolandstr. 8, 49069 Osnabrück, Germany.
Email: tobias.haertel@uos.de

Abstract

Recruiters routinely use LinkedIn profiles to infer applicants' individual traits like narcissism and intelligence, two key traits in online network and organizational contexts. However, little is known about LinkedIn profiles' predictive potential to accurately infer individual traits. According to Brunswik's lens model, accurate trait inferences depend on (a) the presence of valid cues in LinkedIn profiles containing information about users' individual traits and (b) the sensitive and consistent utilization of valid cues. We assessed narcissism (self-report) and intelligence (aptitude tests) in a sample of 406 LinkedIn users along with 64 LinkedIn cues (coded by three trained coders) that we derived from trait theory and previous empirical findings. We used a transparent, easy-to-interpret machine learning algorithm leveraging practical application potentials (elastic net) and applied state-of-the-art resampling techniques (nested cross-validation) to ensure robust results. Thereby, we uncover LinkedIn profiles' predictive potential: (a) LinkedIn profiles contain valid information about narcissism (e.g. uploading a background picture) and intelligence (e.g. listing many accomplishments), and (b) the elastic nets sensitively and consistently using these valid cues attain prediction accuracy ($r = .35/.41$ for narcissism/intelligence). The results have practical implications for improving recruiters' accuracy and foreshadow

Selected results from an earlier version of this manuscript were presented at a conference: Härtel, T. M., Schuler, B. A., & Back, M. D. (2023, August 4–8). *Using valid cues to predict narcissism and intelligence from LinkedIn profiles* [Conference presentation]. The 83rd Annual Meeting of the Academy of Management (AOM), Boston, USA. <https://journals.aom.org/doi/abs/10.5465/AMPROC.2023.12425abstract>.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial](https://creativecommons.org/licenses/by-nc/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2024 The Author(s). *Journal of Occupational and Organizational Psychology* published by John Wiley & Sons Ltd on behalf of The British Psychological Society.

potentials and limitations of automated LinkedIn-based assessments for selection purposes.

KEYWORDS

Brunswikian lens model, cybervetting, LinkedIn, machine learning, trait assessment

Practitioner points

- LinkedIn profiles hold predictive potential to infer applicants' individual traits, such as narcissism and intelligence.
- Information about which LinkedIn cues provide valid trait signals and which do not is useful for improving recruiters' cybervetting accuracy.
- The use of machine learning algorithms to automate LinkedIn-based trait assessment may offer a non-intrusive approach to complement traditional trait testing in selection.
- Before automated LinkedIn-based trait assessments may be applied, more research is needed on the psychometric properties, adverse impacts, and applicant reactions.

INTRODUCTION

With the growth of online networks, practices like 'cybervetting' and 'social media assessments' (Cubrich et al., 2021; Hartwell & Campion, 2020) inform recruiters' selection decisions. LinkedIn is the most popular online professional network with 950 million users (LinkedIn, 2023). Recruiters routinely use applicants' LinkedIn profiles to infer traits (e.g. Roulin & Levashina, 2019), such as the presence of an About section as a signal of trait self-presentation (Van de Ven et al., 2017). LinkedIn profiles do not only provide information available in resumés (e.g. educational/professional experiences) or online social networks, such as Facebook (e.g. number of connections), but add information like followed interests and recommendations (Fernandez et al., 2021).

Despite calls for empirical insights (Roth et al., 2016; Van Iddekinge et al., 2016), LinkedIn's predictive potential to infer individual traits remains unclear. There are few robust and mixed findings on the LinkedIn information (*cues*) signalling traits, leading to contradictory conclusions as to whether LinkedIn profiles *allow* for accurate inferences (Fernandez et al., 2021) *or not* (Roulin & Stronach, 2022). Also, recruiters were shown to achieve only modest accuracy when inferring traits on LinkedIn (e.g. Roulin & Levashina, 2019; Van de Ven et al., 2017). This may be due to LinkedIn's inherent lacking capacity to signal valid trait information or perceivers' lack of sensitive (weighting valid cues more strongly than non-valid cues) and/or consistent (weighting cues the same way across targets) use of such information.

This study aims to illuminate LinkedIn's predictive potential for trait inferences using a twofold lens model (Brunswik, 1956) approach (Figure 1). First, we identify robust LinkedIn cues conveying valid trait information. Second, we examine machine learning algorithms' accuracy, acting as automated perceivers sensitively and consistently using these valid cues (Bleidorn & Hopwood, 2019; Tay et al., 2020). If valid information is present in LinkedIn profiles, automated perceivers will detect and consistently weight these cues across targets, thereby revealing LinkedIn's predictive potential.

We examine LinkedIn profiles ($N=406$) coded by three trained coders, alongside users' narcissism and intelligence test scores. Narcissism contributes to understanding workplace outcomes (Judge et al., 2006) and is expressed in online self-presentation (Gnambs & Appel, 2018). Intelligence is a

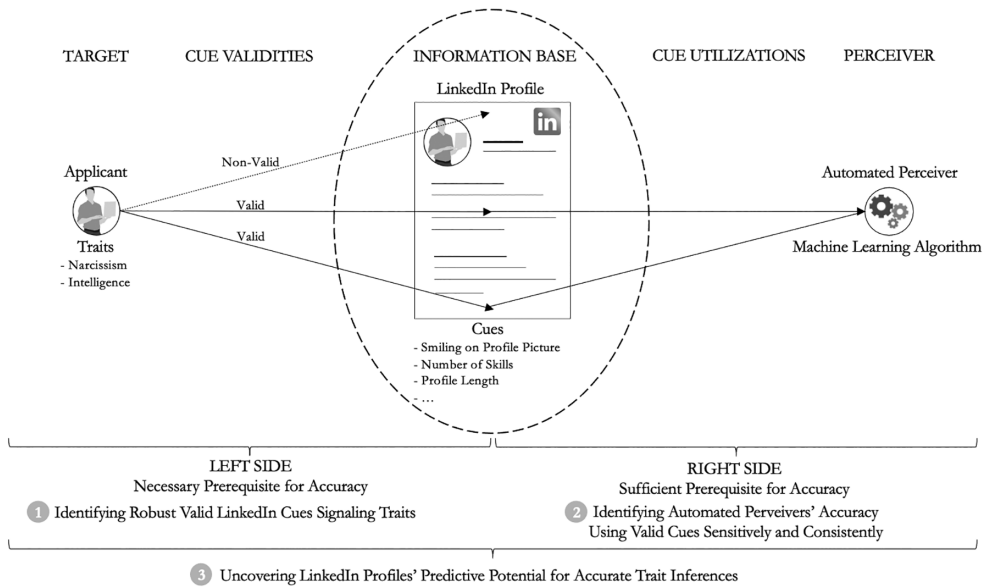


FIGURE 1 Brunswikian lens model in the context of automated cybervetting based on LinkedIn profiles.

key predictor of job performance (Sackett et al., 2022), and recruiters believe they can assess it from LinkedIn (Hartwell & Campion, 2020). We derived 64 LinkedIn cues based on theoretical/empirical evidence to signal these traits. This ensured relevance of cues, addressing content validity concerns in automated approaches. State-of-the-art resampling (nested cross-validation) provided robust results and an easy-to-interpret machine learning algorithm (elastic net) maximized practical applicability.

We enrich the literature on trait expression in selection contexts by (1) adding robust LinkedIn cues signalling narcissism (e.g. uploading background pictures) and intelligence (e.g. listing schools with many followers) to the information bases providing valid trait signals (e.g. resumés, job interviews; Härtel et al., 2024; DeGroot & Gooty, 2009). Also, we enrich the literature on automated trait assessments for selection purposes by (2) showing that automated perceivers can accurately infer traits not only from online social networks like Facebook (Azucar et al., 2018) but also from professional networks like LinkedIn (prediction accuracy $r = .35/.41$ for narcissism/intelligence). Combining these contributions (3) clarifies mixed findings on LinkedIn's potential for accurate trait inferences: LinkedIn offers the possibility to make accurate trait inferences when sensitively and consistently incorporating valid cues (as identified in this study) into trait inferences (like automated perceivers do).

Applying the lens model to (automated) trait inferences based on LinkedIn profiles

This study examines the accuracy of inferring individual traits of an unknown target person from their observable LinkedIn profile information. Following the lens model (Brunswik, 1956; see Back & Nestler, 2016; Hammond, 1996; Karelaia & Hogarth, 2008; Nestler & Back, 2013), latent traits are inferred indirectly in this situation by drawing on observable cues (Figure 1).

The necessary prerequisite for accuracy is the presence of valid cues – LinkedIn information associated with users' traits (left side of the lens model). LinkedIn may contain such valid information (Fernandez et al., 2021), but prior findings were instable (Roulin & Levashina, 2019; Roulin & Stronach, 2022). This raises concerns of whether LinkedIn profiles actually hold the capacity for accurate trait inferences.

The sufficient prerequisite for accurate LinkedIn-based trait inferences is that valid cues are used sensitively and consistently when forming judgements (right side of the lens model). Sensitive cue utilization means that perceivers prioritize valid cues over non-valid cues – being sensitive for what relevant and irrelevant cues are. Consistent cue utilization means weighting the cue information in the same way across targets – applying consistent judgement rules to each target (Back & Nestler, 2016). Although LinkedIn profiles likely contain some valid trait information, recruiters' inferences have been lacking accuracy (Roulin & Levashina, 2019; Roulin & Stronach, 2022; Van de Ven et al., 2017). This may stem from the lack of sensitivity and/or consistency in using valid cues. For instance, recruiters use the presence of a profile picture as a cue for agreeableness (Van de Ven et al., 2017), despite no actual association (Roulin & Levashina, 2019; Van de Ven et al., 2017), indicating a lack of sensitivity. Also, human perceivers are typically inconsistent in applying judgement rules across targets, indicating a lack of consistency (Karelaia & Hogarth, 2008).

The lens model's value lies in its scope covering human and automated approaches (Cannata et al., 2022; Tay et al., 2020) – the human recruiter is replaced by a machine learning algorithm in the automated approach. Algorithms can be construed as automated perceivers constructed to sensitively and consistently use valid cues to make inferences. They autonomously learn which cues are (not) valid based on training data (sensitivity) and apply consistent weights across targets when inferring traits in new data (consistency). Focusing on the automated perceiver thus allows us to quantify LinkedIn's predictive potential purified from the imperfections of human judgement: If LinkedIn profiles contain a valid cue base signalling narcissism/intelligence, the automated perceiver should achieve prediction accuracy because it fulfils the sufficient prerequisite for accuracy.

Previous research on valid cues signalling traits on LinkedIn

Initial research examined cue validities using small explorative LinkedIn cue sets ($9 \leq N_{\text{Cues}} \leq 22$) in relatively small samples ($97 \leq N \leq 154$; Roulin & Levashina, 2019; Roulin & Stronach, 2022; Van de Ven et al., 2017). This exploratory approach offers initial insights; for example, users high on self-presentation often included an About section, providing opportunity for self-promotion (Van de Ven et al., 2017). Yet, some findings were puzzling – self-presenters showing higher education (Van de Ven et al., 2017) – and instable – none of the valid cues in Roulin and Levashina (2019) was confirmed by Roulin and Stronach (2022) and vice versa. While these findings suggest limited predictive potential of LinkedIn, the exploratory nature warrants caution. In contrast, using a deductive approach recently enabled to demonstrate the existence of a valid LinkedIn cue base for Big Five traits. Fernandez et al. (2021) addressed prior limitations by reporting cue validities of (a) a wider range of LinkedIn cues ($N_{\text{Cues}} = 33$) (b) based on theoretical underpinnings to signal Big Five traits (c) in a large sample of 607 users. They suggest that LinkedIn could offer a sufficient cue base for accurate trait inferences, at least for selected Big Five traits.

We address mixed prior findings by expanding on valid LinkedIn cues signalling individual traits in terms of narcissism/intelligence. We therefore (a) examine an extensive cue set ($N_{\text{Cues}} = 64$) (b) derived from trait theory and empirical findings ensuring relevance and interpretability (c) in a sample of 406 users. We prioritize a high ratio of potential valid cues per trait to avoid underestimating predictive potentials. Also, while single trait-cue links typically show small effect sizes, combining subtle cues can reveal meaningful information. Further, we identify valid cues not only from correlations but based on feature importance in nested cross-validated models. This addresses concerns of capitalizing on chance and overfitting by prioritizing robust cues based on their ability to predict novel data.

Automated approach to trait inferences based on LinkedIn

Automated approaches applied to online social networks like Facebook based on various cue sets (e.g. posts/pictures/likes) achieved remarkable prediction accuracy for individual traits (Azucar et al., 2018;

Settanni et al., 2018), including narcissism (e.g. Garcia & Sikström, 2014; Sumner et al., 2012) and intelligence (e.g. Kosinski et al., 2013, 2014). Unlike social networks like Facebook, LinkedIn focuses on professional identity (Hartwell & Campion, 2020), making it advantageous for selection contexts due to fewer issues with adverse impact, acceptance, and legal matters (Stoughton et al., 2015). An initial attempt to transfer the automated approach to LinkedIn used automated language-based Big Five assessments, but found no prediction accuracy (Roulin & Stronach, 2022). This raised calls for less reliance on limited textual information, showing weak trait associations (Chen et al., 2020; Holtzman et al., 2019). Generally, research using automated approaches typically applies data-driven, explorative approaches without preselecting cues based on theory/empirical evidence (Settanni et al., 2018). While useful for uncovering novel associations and theory building, this lacks content validity and has limitations: (a) related constructs might be measured rather than trait content, (b) cues' interpretability may suffer reducing application potentials ('black box'), and (c) generalizability across online networks is constrained (Alexander et al., 2020; Bleidorn & Hopwood, 2019; Tay et al., 2020).

We quantify LinkedIn's predictive potential by applying the automated approach. We go beyond textual cues by targeting an extensive set of cues that depict qualitatively different types of information reflecting LinkedIn's broad information spectrum. This aligns with the 'good information' principle (Back & Nestler, 2016), positing that interpersonal judgements' accuracy depends on the quantity and quality of accessible information. Incremental valid information enhances accuracy (Borkenau et al., 2004; see Azucar et al., 2018; Settanni et al., 2018). Including a diverse array of cues is thus important to depict predictive potentials: If valid cues exist but are not coded, predictive potentials are underestimated. We base our cue set on theory and empirical evidence, establishing conceptual cue–trait links. We prioritize elastic nets over more complex algorithms, emphasizing transparency of how inferences are made over gaining some predictive power. Elastic nets are regularized regressions that avoid in-sample overfitting and optimize out-of-sample prediction accuracy by shrinking the regression coefficients of cues that contain less predictive information towards zero (see Hastie et al., 2009; James et al., 2013, for an introduction). This makes them parsimonious models, similar in interpretation to multiple linear regressions, but better at predicting new data (see [Analytical Approach](#)). Transparent selection decisions are essential for ensuring fair and legal procedures (Goretzko & Israel, 2022).

Relevance of narcissism and intelligence in LinkedIn profile assessment

We heed calls to prioritize alternative traits in cybervetting (Mönke & Schäpers, 2022; Settanni et al., 2018). Grandiose narcissism is a form of entitled self-importance with agentic and antagonistic components (Back, 2018) evoking organizational consequences (Campbell et al., 2011) beyond the Big Five (Judge et al., 2006). Due to assertive charm, individuals high on narcissism often propel into managerial roles (Wille et al., 2019). While some job roles benefit from moderate narcissism (Satornino et al., 2023), high levels pose risks due to overconfidence and disregard for others (O'Boyle et al., 2012). Research (Blair et al., 2008; De Fruyt et al., 2009; Judge & LePine, 2007) and the practitioner literature (Rotolo & Bracken, 2022; Schwarzingner, 2022) advocate assessing narcissism in hiring, despite challenges using traditional selection methods (e.g. self-reports, job interviews) due to socially desirable responding (Bensch et al., 2019; Kowalski et al., 2018) and impression management (Paulhus et al., 2013). Online networks serve individuals high on narcissism as platform to exercise their need for external validation by providing self-enhancement opportunities (Gnambs & Appel, 2018; McCain & Campbell, 2018). Thus, LinkedIn may offer an alternative to assess narcissism.

Intelligence – the ability to process complex information facilitating higher order thinking like reasoning/problem-solving (Gottfredson, 1997) – is the most potent trait to predict job performance across occupations (Sackett et al., 2022), particularly in complex jobs (Salgado et al., 2003) and high-level leadership positions (Ones & Dilchert, 2009). Whereas little is known about expressing intelligence online (Schroeder et al., 2020), recruiters use LinkedIn to infer intelligence (Roulin & Stronach, 2022; see Brown & Campion, 1994; Kluepfer & Rosen, 2009) and consider it an effective source to do so

(Hartwell & Campion, 2020). LinkedIn may offer a non-invasive alternative to established but less commonly applied intelligence tests (König et al., 2010; Zibarras & Woods, 2010).

METHOD

Sample/procedure

The sample comprised 406 (243 female) German-speaking LinkedIn users with substantial subscribers ($M = 297.15$, $SD = 468.32$) recruited through online postings, participant recruitment platforms, and university lecture announcements. The average age was 29.47 ($SD = 9.69$), 283 (69.70%) participants held a bachelor's degree or higher, 258 (63.55%) were studying (37.59% business/economics), and 215 (52.96%) were working ≥ 20 weekly hours in various sectors/professions. Participants completed an online survey assessing demographics and individual traits. At survey start, participants connected their LinkedIn profiles with a LinkedIn showcase page. We saved profiles in various formats, including pdf exports to assess profile length, profile pictures as separate images, and screenshots of other relevant information. For details on the survey and psychometric tests, see the Codebook, <https://osf.io/4ruqj/>.

Measures

Individual traits

Table 1 presents descriptive statistics and Cronbach's alphas for traits. The 18-item Narcissistic Admiration and Rivalry Questionnaire (NARQ; Back et al., 2013) measured narcissism with 6-point scales ranging from 1 (*do not agree at all*) to 6 (*agree completely*). The NARQ captures the subdimensions of narcissistic admiration (agentic self-presentation) and narcissistic rivalry (antagonistic self-defence), together constituting a theoretically informed assessment of grandiose narcissism.

To ensure a comprehensive intelligence measure, we included *fluid intelligence* (reasoning/problem-solving) and *crystallized intelligence* (accumulated knowledge) as general intelligence subdimensions (Cattell, 1963; Nisbett et al., 2012). To manage survey duration, we used short scales offered by a well-established open-access repository for measurement instruments in German (GESIS – Leibniz Institute for the Social Sciences, n.d.). The short version of the Hagen Matrices Test (HMT-S; Heydasch et al., 2012, 2013) measured fluid intelligence ($M = 4.31$, $SD = 1.45$). The test consists of six items where test-takers select the figure that logically completes a 3×3 matrix from eight alternatives. The short version of the Berlin Test for the Assessment of Fluid and Crystallized Intelligence (BEFKI GC-K; Schipolowski et al., 2013; Wilhelm et al., 2014) measured crystallized intelligence ($M = 8.78$, $SD = 2.04$). The test comprises 12 items presenting declarative knowledge questions from various fields (e.g. 'In a well-known painting by Dalí, "deliquescent clocks" are depicted. To which style can this painting be assigned?'), with test-takers selecting the correct answer from four options. Scores for both tests are based on the number of correctly completed items. We aggregated both measures' z -standardized scores ($r = .26$, $p < .001$) to obtain a comprehensive, accurate intelligence indicator by reducing measurement error and enhancing content validity (Breit et al., 2024). This also simplifies interpretation in personnel selection. The modest correlation of $r = .26$ between fluid and crystallized intelligence reflects their conceptual differences as related but distinct constructs (Ackerman et al., 2001; Ziegler et al., 2012).

LinkedIn cues

Cue derivation was based on trait theory and empirical findings ensuring content validity and interpretability. The cue set was developed before coding, without knowing the results, though not pre-registered.

The four-step approach yielded 64 LinkedIn cues, categorized into two types. Objective cues ($n = 55$) were straightforward to code by a single coder (e.g. counting skills). Subjective cues ($n = 9$) required judgement (e.g. rating physical attractiveness) and were rated by two coders on scales from 1 (*not at all*) to 6 (*completely*). The three coders were thesis candidates or research interns and received extensive training. Interrater agreement was good to excellent (Koo & Li, 2016; $ICC_{3,k} \geq .84$; see Table 1). We touch upon the four-step cue derivation process for illustrative purposes; details are provided in the Supplemental Results Section A in Appendix S1, <https://osf.io/4ruqj/>.

(1) We derived cues transferring theoretical implications/empirical findings on narcissism/intelligence from broader non-online contexts to LinkedIn. For example, individuals high on narcissism are motivated to attain leadership (Benson et al., 2016), serving as platform to earn admiration (Campbell & Campbell, 2009). Their assertiveness (Härtel et al., 2021) helps narcissistic individuals attain group leadership (Grijalva et al., 2015), managerial ranks (Wille et al., 2019), and prestigious leadership roles (Watts et al., 2013). On LinkedIn, this may manifest through leadership positions/skills in the Experience/Skills sections. The same may hold true for intelligence as a key leadership attribute and predictor of leadership emergence (Judge et al., 2004). (2) We transferred theoretical implications/empirical findings on associations between narcissism/intelligence and cues in related contexts like online social networks and resumés to LinkedIn. For instance, individuals high on narcissism often boost their grandiosity through online posting (McCain & Campbell, 2018). This self-promotion is effective with large audiences, providing attention and popularity (Marshall et al., 2020). Indeed, narcissism is associated with large online networks (Gnambs & Appel, 2018). On LinkedIn, this may manifest through frequent posting/liking/commenting and large network sizes. As another example, intelligence was shown to be expressed in listing scholastic awards on resumés (Cole et al., 2003), as it enhances learning and academic/occupational success (Kuncel et al., 2004; Ng et al., 2005). Accordingly, intelligence might be expressed through honours/awards in LinkedIn's Accomplishments. (3) We added cues linked to narcissism/intelligence-related traits in prior LinkedIn research. For instance, whereas self-presenters were found to exhibit less smiling on LinkedIn (Van de Ven et al., 2017), individuals high on narcissism smile in social situations (Back, Stopfer, et al., 2010). To clarify this association on LinkedIn, we examined genuine, wide smiles. Regarding intelligence, for example, we investigated whether it manifests in detailed descriptions of professional experiences, addressing conflicting results (Roulin & Levashina, 2019; Roulin & Stronach, 2022). (4) Research suggests that LinkedIn cues like the number of skills/interests on CEO profiles can indicate narcissism (Aabo & Eriksen, 2018; Cragun et al., 2020). We added such cues as potential narcissism indicators, though some cues' validity (e.g. self-referential language use) is questionable (Carey et al., 2015).

Analytical approach

To get an impression of cue validities, we computed bivariate correlations of narcissism/intelligence with the respective LinkedIn cues. We winsorized cues (Wilcox, 2011) by setting extreme values $|\tilde{z}| > 4.47$ to ± 4.47 . Winsorizing reduces extreme values' disproportionate influences while retaining their predictive value, reducing overfitting to the current sample, and ensuring generalizability to new samples. The threshold of 4.47 is based on Chebyshev's inequality (Saw et al., 1984), showing that at least 95% of the data fall within ± 4.47 standard deviations of the mean, regardless of the distribution. This limits winsorizing to extreme values representing no more than 5% of the data, usually much less.

We used a machine learning approach (see Hastie et al., 2009; James et al., 2013; Lantz, 2019, for an introduction) to illuminate LinkedIn's potential for trait inferences. The discipline of machine learning seeks to 'construct computer systems that automatically improve through experience' (Jordan & Mitchell, 2015, p. 255): Machine learning algorithms are programs consisting of rule sets enabling them to automatically improve their performance on a task by learning generalizable patterns from training

data. The learned patterns can then be used to perform the task on new data. For example, these algorithms can learn associations between LinkedIn profile cues and traits from training data and then predict traits for new users based on their profiles.

Multiple linear regression models perform poorly on prediction tasks as they typically overfit the training data (Zou & Hastie, 2005). Overfitting means that models learn spurious patterns from the training data that are not generalizable to new data. Multiple linear regressions optimize explained variation by minimizing the sum of the squared deviations of the predicted from the observed values of a target feature in the training data (i.e. *in-sample*). By construction, they adapt to the training data (and thus to its specificities and biases) to the point where the model predicts new data less well. This makes multiple linear regression unsuitable for prediction tasks. To address overfitting, we fitted elastic nets (Zou & Hastie, 2005) using the respective derived LinkedIn cues as independent variables (features)¹ to predict the dependent variables narcissism/intelligence (target features). Elastic nets are regularized regressions that optimize prediction performance on new data (i.e. *out-of-sample*) by automatically selecting a penalty λ that shrinks those overfitted regression coefficients that hurt out-of-sample prediction accuracy towards zero (James et al., 2013). This regularization prevents overfitting and enhances generalizability. Automatic selection of λ to maximize out-of-sample prediction accuracy characterizes elastic nets as machine learning algorithms.

The elastic net algorithm typically selects the optimal penalty λ using k -fold cross-validation (Hastie et al., 2009; James et al., 2013; Kohavi, 1995). The data are split into k (typically 10) equal-sized parts, with nine parts constituting the in-sample training data to select the optimal penalty λ and the remaining part serving as the out-of-sample test data to estimate the prediction performance of the elastic net given the penalty λ . Specifically, elastic nets fit regressions with a fixed number of penalties λ (typically 100) to the training data. Then, each regression is used to predict the target feature based on the features using the test data. Having computed the chosen prediction performance metric (e.g. mean squared error) for each regression with its specific penalty, the elastic net selects the penalty λ for which the cross-validation error (i.e. the chosen prediction performance metric on the test set) is smallest. This procedure is repeated k times so that each part of the data has served as test data once. This yields k elastic nets each with a different λ and out-of-sample prediction performance. Finally, the elastic net with the penalty λ that resulted in the smallest out-of-sample prediction error is fitted to all available data. This model can then be used to predict the target feature based on the features with which the elastic net was trained.

Through cross-validation, elastic nets learn the penalty λ that minimizes out-of-sample prediction error, enabling them to learn more generalizable association patterns than multiple linear regression models. By shrinking coefficients of less predictive cues towards zero, they sensitively use cues. Furthermore, they use cues consistently by fitting a generalizable model that is consistently applied to all LinkedIn users. This makes elastic nets highly sensitive and consistent, and thus, a useful algorithm for unravelling LinkedIn's predictive potential based on the lens model principles. Importantly, elastic nets maintain interpretability as the coefficients can be obtained from the final model and interpreted similar to multiple linear regression (Alexander et al., 2020). Combining prediction performance and interpretability made elastic nets the ideal algorithm for this study.

Previous research has used k -fold cross-validation (Kohavi, 1995) to evaluate machine learning algorithms' prediction accuracy. However, this method leads to optimistically biased performance estimates as it combines hyperparameter tuning (e.g. optimize the penalty λ) and model selection based on the estimated prediction performance (Cawley & Talbot, 2010; Varma & Simon, 2006). By using the same data for hyperparameter tuning and estimating prediction accuracy, information from hyperparameter tuning leaks into the model selection (e.g. by selecting the penalty λ with the smallest out-of-sample prediction error as the final model). k -fold cross-validation thus allows algorithms to learn what model fits well to the test data, thereby inflating estimated prediction accuracy.

¹We included gender (0/1 = female/male) and age as controls, as traits show consistent differences for gender/age (e.g. Nisbett et al., 2012; Weidmann et al., 2023), which may also affect LinkedIn cues. For example, men often list leadership skills, while women smile more (Fernandez et al., 2021).

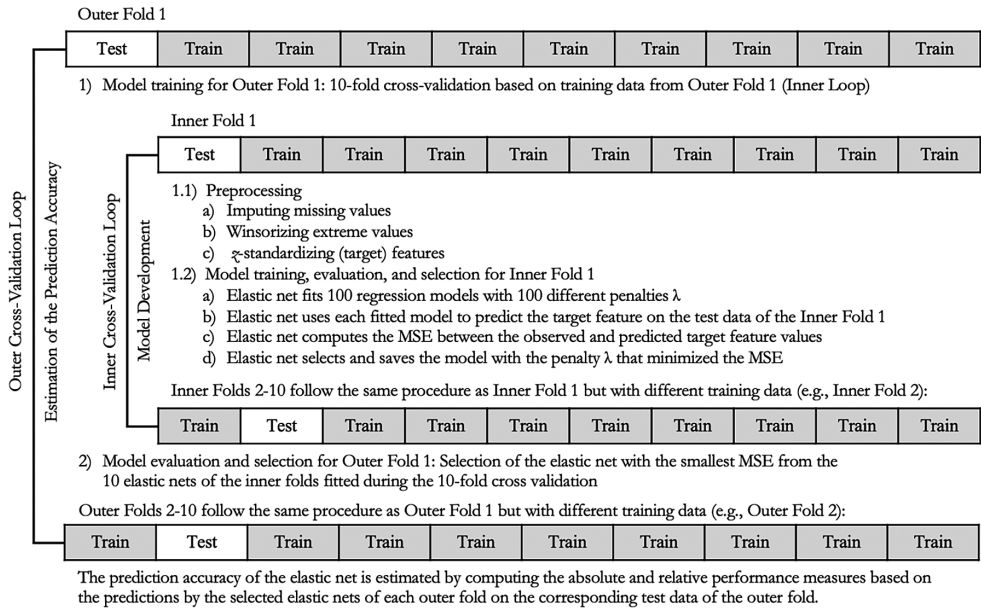


FIGURE 2 Overview of the nested cross-validation approach for the elastic nets applied in the present study.

Taking a *nested* cross-validation approach remedies this issue, efficiently utilizing all 406 LinkedIn profiles to both develop and validate the predictive model without the ‘double use’ of data that lead to optimistic bias in conventional k -fold cross-validation – making it the method of choice for extracting robust prediction results even from small datasets (Vabalas et al., 2019). Similar to k -fold cross-validation, the data are first partitioned into k (typically 10) equal parts. One part serves as the out-of-sample test data to estimate the prediction error of an elastic net (the outer loop). Instead of using the other nine parts as training data, these are each split into 10 parts again to conduct a tenfold cross-validation to optimize the penalty λ of the elastic net (the inner loop). This inner loop cross-validation is repeated, so that each part in the outer loop serves once as the out-of-sample test set for a model selected by a tenfold cross-validation in the inner loop. By separating hyperparameter tuning (inner loop) from out-of-sample prediction accuracy estimation for model selection (outer loop), elastic nets only learn from the training data (not the test data), ensuring accurate prediction accuracy estimates. Figure 2 summarizes this approach.

Machine learning algorithms, though capable of self-improvement within their rule sets, are human-designed artefacts that require human choices to optimize their learning in specific contexts. This is similar to running a computer program, where automatic numeric operations depend on user-defined settings that influence the program’s execution. In the following, we describe these settings, reflecting our decisions on model fitting.

The inner loop involved a ten fold cross-validation for preprocessing and hyperparameter tuning. Preprocessing comprised three steps: (1) imputing missing values using the missForest algorithm, effective for high-dimensional data with continuous and categorical features (Stekhoven & Bühlmann, 2012); (2) winsorizing² extreme values ($|z| > 4.47$) to mitigate disproportional impacts on prediction accuracy (Hastie et al., 2009); and (3) z -standardizing (target) features to ensure consistent application of λ to regression coefficients, making feature importance (i.e. the regression coefficients’ magnitude) comparable (Tibshirani, 1997). For hyperparameter tuning, the elastic net tested 100 λ

²Computing elastic nets based on raw instead of winsorized data led to slightly lower prediction performance and somewhat inflated regression coefficients (Supplemental Results Section D in Appendix S1, <https://osf.io/4ruqj/>).

and selected the λ minimizing the mean squared error between predicted and observed target feature values. We saved the best performing model of the tenfold cross-validation for evaluation on the test data in the outer loop.

The outer cross-validation loop repeated the inner loop 10 times, yielding 10 models representing the best fits from the inner tenfold cross-validation. We averaged (a) mean squared error (MSE), (b) root mean squared error (RMSE), and (c) mean absolute error (MAE) across the 10 outer folds for absolute prediction performance. For relative prediction performance, we averaged (d) correlations between predicted and observed target feature values r ; (e) the variance explained by the features R^2 , and (f) the adjusted variance explained by the features R^2_{Adj} across the 10 outer folds. We used a minimum-variance unbiased estimator to average correlations (Alexander, 1990; Olkin & Pratt, 1958). We computed the means, standard deviations, and cross-validation fold incidences (CV_{FF} , number of outer folds the regression coefficient of a feature was $\neq 0$) of the regression coefficients from the outer loop models as feature importance measures.

We compared the elastic nets' prediction performance with that of multiple linear regressions and intercept-only models using the same nested cross-validation approach. The comparison to multiple linear regressions shows whether elastic nets prevent overfitting, and the comparison to intercept-only models is standard practice to test whether algorithms effectively extract feature information beyond a simple mean-based prediction (James et al., 2013). Intercept-only models can make surprisingly accurate predictions (Lantz, 2019), serving as an important benchmark to gauge an algorithm's added predictive capability.

We used the statistical program R (version 4.1.0; R Core Team, 2018) and RStudio (version 1.4.1106; RStudio Team, 2016) for all analyses, including the *glmnet* (version 4.1-2; Friedman et al., 2010) and *missForest* packages (version 1.5; Stekhoven, 2022) for machine learning. The data and statistical code can be found at <https://osf.io/4ruqj/>. We offer Supplemental Results (Section E in Appendix S1, <https://osf.io/4ruqj/>) of the elastic nets with (a) the entire cue set as features, (b) narcissism subdimensions (Back et al., 2013) as target features, (c) narcissism measured with the German 13-item short form of the Narcissistic Personality Inventory (G-NPI-13; Brailovskaia et al., 2019; Gentile et al., 2013) as target feature, and (d) intelligence subdimensions (Cattell, 1963; Nisbett et al., 2012) as target features.

RESULTS

Bivariate correlations

Descriptive statistics and bivariate correlations of the LinkedIn cues with narcissism/intelligence are shown in Table 1. An extensive intercorrelation table is provided in the Supplemental Results Section C in Appendix S1, <https://osf.io/4ruqj/>. Ten of the 48 cues derived for narcissism showed significant correlations (individuals high on narcissism featured (a) less smiling on profile pictures; (b) background pictures; (c) leadership positions in the Experience section; (d) more skills; (e) public speaking skills; (f) more accomplishments related to organizations; (g) more received recommendations; (h) interests with more followers; (i) English profiles; (j) altruistic volunteering). Fifteen of the 39 cues derived for intelligence showed significant correlations (individuals high on intelligence featured (a) less dressed-up, trimmed appearances on profile pictures; (b) profile pictures above neckline; (c) more and (d) extensively described educational stations; (e) more and (f) longer durations of volunteer experiences; (g) more accomplishments and more accomplishments related to (h) honours/awards, (i) publications, (j) projects, (k) test scores; (l) schools with more followers; (m) more groups; (n) English profiles; (o) longer profiles). The findings offer initial evidence that LinkedIn cues signal narcissism/intelligence, which are to be confirmed by the elastic nets, and for a valid cue base that should allow accurate inferences by automated perceivers.

TABLE 1 Measurement and descriptive statistics for the LinkedIn cues and correlations with narcissism and intelligence.

Cue	Measurement	n^a	M	SD	r_N [CI]	r_I [CI]
Cues potentially signalling narcissism and intelligence						
Profile card						
Name with title	0 = not present; 1 = present (e.g. BSc, MA, PhD, Dr, Prof)	406	.17	.37	-.03 [-.12, .07]	.06 [-.04, .16]
Profile picture						
Profile picture	0 = not present; 1 = present (only pictures of an actual person coded as '1')	381	.94	.24	-.09 [-.19, .00]	.01 [-.09, .11]
Charming facial expression	1 = <i>not at all</i> ; 6 = <i>completely</i> (i.e. confident, friendly facial expression)	381	3.52	1.48	-.05 [-.15, .05]	-.09 [-.18, .01]
Professional/formal appearance	1 = <i>not at all</i> ; 6 = <i>completely</i> (e.g. suit/blazer/dress, formal accessories such as necktie, fly)	381	3.49	1.66	.05 [-.05, .15]	.00 [-.10, .11]
Stylish/flashy/fashionable appearance	1 = <i>not at all</i> ; 6 = <i>completely</i> (e.g. extravagant/dyed hair, hair ornaments (e.g. clip), colourful/extravagantly cut clothes, eye-catching jewellery (e.g. earrings, watch) and accessories (e.g. scarf), cleavage, tattoos, piercings)	381	3.48	1.55	.03 [-.07, .13]	-.06 [-.16, .04]
Dressed-up, trimmed appearance						
	1 = <i>not at all</i> ; 6 = <i>completely</i> (e.g. plucked eyebrows, make-up (e.g. lipstick, made-up eyelashes, eye shadow), manicured fingernails, styled/gelled hair, well-groomed/neatly shaved beard)	381	3.54	1.52	.09 [-.01, .19]	-.12 [-.21, -.01]
Professional shot						
	1 = <i>not at all</i> ; 6 = <i>completely</i> (i.e. looks like taken by a professional, good lighting, highly resolved, photo effects (e.g. blurring, editing software))	381	3.57	1.56	.06 [-.04, .16]	.03 [-.07, .13]
About						
About	0 = not present; 1 = present	406	.32	.47	.07 [-.02, .17]	.08 [-.02, .18]
Extensive About section	Number of words in About Section	406	22.77	50.22	.00 [-.10, .10]	.08 [-.02, .17]

TABLE 1 (Continued)

Cue	Measurement	n^a	M	SD	r_N [CI]	r_I [CI]
Experience						
Professional positions	Numerically counted	406	4.96	2.95	-.01 [-.11, .09]	.07 [-.03, .17]
Leadership positions	0 = not present; 1 = present (e.g. president, head, director, team lead)	406	.15	.35	.11 [.01, .21]	.07 [-.03, .17]
Education		406				
Business studies	0 = not present; 1 = present (i.e. studies of business administration or subdisciplines (e.g. finance, marketing))	406	.35	.48	.04 [-.06, .14]	-.08 [-.18, .02]
Volunteering						
Volunteer experiences	Numerically counted	406	.83	1.48	.04 [-.05, .14]	.13 [.04, .23]
Altruistic volunteering	0 = not present; 1 = present (i.e. volunteering with strong helping focus (e.g. food bank, red cross))	406	.10	.30	.11 [.01, .20]	.07 [-.03, .17]
Average duration	Total duration of volunteer experiences in months divided by total number of volunteer experiences	403	12.50	27.66	.04 [-.07, .13]	.14 [.05, .24]
Licences and certifications						
Licences and certifications	Numerically counted	406	1.00	2.13	.06 [-.03, .16]	-.01 [-.11, .09]
Skills						
Skills	Numerically counted	406	12.21	11.45	.14 [.05, .24]	.03 [-.07, .12]
Endorsements	Numerically counted	406	13.77	39.61	.05 [-.04, .15]	.08 [-.02, .18]
Leadership	0 = not present; 1 = present (e.g. team leadership, people management)	406	.17	.37	.08 [-.02, .18]	-.04 [-.13, .06]
Accomplishments						
Accomplishments	Numerically counted	406	3.58	5.28	.08 [-.02, .18]	.21 [.12, .30]
Recommendations						
Received recommendations	Numerically counted	406	.07	.41	.13 [.03, .22]	.05 [-.04, .15]

(Continues)

TABLE 1 (Continued)

Cue	Measurement	n^a	M	SD	r_N [CI]	r_I [CI]
Other						
Profile in English	0 = not present; 1 = present (i.e. profile at least partly in English)	406	.47	.50	.10 [.01, .20]	.13 [.03, .22]
Profile length	Number of words counted based on pdf export from LinkedIn profile	406	171.57	112.65	.04 [-.06, .13]	.14 [.03, .22]
Cues potentially signalling narcissism						
Profile card		406	2.52	.60	.84 [.81, .86]	
Extensive profile card	Number of words to briefly describe the own person	406	7.03	4.21	.04 [-.05, .14]	-.03 [-.12, .07]
Twitter	0 = not present; 1 = present (i.e. link to user's Twitter profile provided)	406	.03	.18	.02 [-.08, .11]	.05 [-.05, .14]
Profile picture						
Glasses	0 = not present; 1 = present (i.e. glasses as vision aid, no sunglasses)	381	.24	.43	.01 [-.09, .11]	.04 [-.06, .14]
Smiling	1 = <i>not at all</i> ; 6 = <i>completely</i> (i.e. 'laughing with the eyes', teeth shown)	381	3.51	1.63	-.18 [-.28, -.08]	.01 [-.09, .11]
Physical attractiveness	1 = <i>not at all</i> ; 6 = <i>completely</i>	381	3.51	1.54	-.06 [-.16, .04]	-.06 [-.16, .04]
Additional pictures						
Background picture	0 = not present; 1 = present	406	.33	.47	.15 [.05, .24]	.01 [-.09, .11]
Self-promotional background picture	1 = <i>not at all</i> ; 6 = <i>completely</i> (i.e. geared towards convincing others of user's positive qualities)	132	3.25	1.57	.00 [-.17, .18]	.02 [-.15, .19]
Additional pictures/videos	0 = not present; 1 = present (i.e. user provides additional pictures/videos in the permanent profile)	406	.13	.34	-.02 [-.12, .08]	.00 [-.10, .10]
About						
Self-promotional content	1 = <i>not at all</i> ; 6 = <i>completely</i> (i.e. geared towards convincing others of user's positive qualities)	125	3.49	1.67	-.02 [-.19, .16]	.02 [-.16, .19]
Self-related words	Number of self-related words (e.g. 'I', 'me', user's name) divided by total number of words in About section	406	1.00	2.80	.01 [-.09, .10]	.05 [-.05, .15]

TABLE 1 (Continued)

Cue	Measurement	n^a	M	SD	r_N [CI]	r_I [CI]
Featured						
Posts	Numerically counted	406	.27	.92	.02 [-.08, .11]	-.01 [-.11, .09]
Activities						
Average activities per day	Total number of activities (limited to latest 50 entries) divided by time period (days between newest and oldest activity)	406	.56	1.30	.06 [-.03, .16]	-.05 [-.15, .05]
Articles/posts/documents	Numerically counted (i.e. activities categorized as 'articles'/ 'posts'/ 'documents'; limited to latest 50 entries)	406	6.17	13.12	.04 [-.05, .14]	.00 [-.10, .10]
Subscribers	Numerically counted (i.e. number of subscribers displayed on profile)	406	297.15	468.32	.06 [-.04, .15]	.02 [-.07, .12]
Experience						
Average duration	Total duration of professional positions in months divided by total number of professional positions	406	26.71	45.94	.04 [-.05, .14]	.00 [-.09, .10]
Skills						
Public speaking	0 = not present; 1 = present (i.e. skills directly related to public speaking; no presentation software skills)	406	.10	.30	.14 [.04, .23]	.01 [-.09, .10]
Interpersonal	Number of skills categorized as 'interpersonal' skills divided by total number of skills	406	.07	.11	.04 [-.06, .14]	-.14 [-.24, -.05]
Teamwork	0 = not present; 1 = present (i.e. skills related to cooperatively working together towards shared goals)	406	.21	.41	.01 [-.09, .11]	-.07 [-.17, .03]
Accomplishments						
Organizations	Number of accomplishments categorized as 'organizations' divided by total number of accomplishments	406	.02	.08	.14 [.04, .23]	.06 [-.04, .16]
Courses	Number of accomplishments categorized as 'courses' divided by total number of accomplishments	406	.04	.14	.08 [-.02, .18]	.05 [-.04, .15]

(Continues)

TABLE 1 (Continued)

Cue	Measurement	α /ICC [CI]	n^a	M	SD	r_N [CI]	r_I [CI]
Interests							
Interests	Numerically counted (i.e. interests related to groups, influencers, companies, and schools)		406	40.92	65.94	-.01 [-.11, .09]	-.04 [-.14, .05]
Median of interests' followers	Median of number of followers of interests related to groups, influencers, companies, and schools (assessing numbers of followers was limited to the first 50 interests in each category)		406	56.074	126,762	.16 [.06, .25]	-.12 [-.21, -.02]
Influencers	Number of interests related to influencers divided by total number of interests		406	.03	.05	.07 [-.03, .16]	-.02 [-.12, .08]
Median of influencers' followers	Median of number of followers of interests related to influencers (assessing numbers of followers was limited to the first 50 influencer interests)		406	1,311,217	4,007,637	.05 [-.05, .14]	-.01 [-.11, .09]
Other							
Sports activities	0 = not present; 1 = present (i.e. activities involving physical exertion (e.g. soccer captain, marathon))		406	.03	.16	.09 [-.01, .19]	.02 [-.08, .11]
Cues potentially signalling intelligence (aggregate of fluid and crystallized intelligence)		.60 [.54, .66]/.54 [.47, .60]	406	.00	.79		
Profile picture							
Picture above neckline	0 = not present; 1 = present		381	.65	.48	-.01 [-.11, .09]	.17 [.07, .27]
Experience							
Extensive description	Total number of words to describe professional positions divided by total number of professional positions		406	8.59	14.29	.02 [-.08, .12]	.06 [-.04, .16]

TABLE 1 (Continued)

Cue	Measurement	n^a	M	SD	r_N [CI]	r_I [CI]
Education						
Marks	0 = not present; 1 = present (e.g. marks as numerical number, percentile ranks, graduation with honours)	406	.27	.45	.01 [-.08, .11]	.06 [-.03, .16]
Averaged marks	Mean of marks presented as numerical number in German grading system (1.0–4.0)	106	1.63	.45	-.14 [-.32, .06]	-.17 [-.35, .03]
Educational stations	Numerically counted	406	2.64	1.48	.07 [-.03, .16]	.14 [.05, .24]
Extensive description	Total number of words to describe educational stations divided by total number of educational stations	406	5.60	10.61	.00 [-.09, .10]	.11 [.01, .20]
Average duration	Total duration of educational stations in years divided by total number of educational positions	406	3.58	1.61	-.05 [-.15, .05]	-.03 [-.13, .06]
Skills						
Industry knowledge	Number of skills categorized as 'industry knowledge' skills divided by total number of skills	406	.17	.20	.09 [-.01, .18]	.00 [-.09, .10]
Tools and technologies	Number of skills categorized as 'tools and technologies' skills divided by total number of skills	406	.13	.17	.07 [-.03, .16]	.03 [-.07, .12]
Accomplishments						
Honours/awards	Number of accomplishments categorized as 'honours and awards' divided by total number of accomplishments	406	.06	.15	.05 [-.05, .14]	.18 [.08, .27]
Publications	Number of accomplishments categorized as 'publications' divided by total number of accomplishments	406	.03	.10	.03 [-.07, .12]	.15 [.05, .24]
Projects	Number of accomplishments categorized as 'projects' divided by total number of accomplishments	406	.02	.07	.05 [-.05, .14]	.10 [.01, .20]
Test scores	Number of accomplishments categorized as 'test scores' divided by total number of accomplishments	406	.01	.03	-.02 [-.12, .07]	.11 [.02, .21]

(Continues)

TABLE 1 (Continued)

Cue	Measurement	α /ICC [CI]	n^a	M	SD	r_N [CI]	r_I [CI]
Interests							
Schools	Number of interests related to schools divided by total number of interests		406	.11	.09	.00 [−.10, .09]	.06 [−.03, .16]
Median of schools' followers	Median of number of followers of interests related to schools (assessing numbers of followers was limited to the first 50 school interests)		406	50,800	64,121	.12 [.02, .21]	.19 [.09, .28]
Groups	Number of interests related to groups divided by total number of interests		406	.07	.10	−.03 [−.13, .07]	.13 [.03, .22]

Note: CI = 95% confidence interval. r_N = bivariate correlation coefficient between grandiose narcissism as measured by the NARQ (Back et al., 2013) and the winsorized raw scores of the LinkedIn cues. r_I = bivariate correlation coefficient between general intelligence, computed as the aggregation of z -scores of fluid intelligence measured with the HMT-S (Heydasch et al., 2013) and crystallized intelligence measured with the BEFKI GC-K (Schipolowski et al., 2013), and the winsorized raw scores of the LinkedIn cues.

^aSome cues have missing values because certain information was not available on the LinkedIn profile, especially due to missing profile pictures and missing or incomplete sections and entries. Correlations in bold are significant at the $p \leq .05$ level (two-sided).

TABLE 2 Performance and hyperparameter estimates of nested cross-validated elastic nets, multiple linear regressions, and intercept-only models predicting narcissism and intelligence.

	Narcissism						Intelligence					
	EN		MLR		IOM		EN		MLR		IOM	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
MSE	.905	.081	1.032	.145	.975	.001	.892	.092	.973	.179	.975	.001
RMSE	.950	.042	1.014	.070	.988	.000	.943	.049	.982	.091	.988	.000
MAE	.760	.054	.818	.062	.793	.041	.744	.057	.780	.082	.782	.035
<i>r</i>	.35	.18	.27	.16			.41	.18	.34	.19		
<i>R</i> ²	.07	.01	.28	.07			.10	.03	.25	.06		
<i>R</i> ² _{Adj.}	.02	.01	.15	.08			.05	.03	.14	.07		
λ_{Min}	.100	.016					.087	.017				

Note: All performance measures are computed based on the observed and predicted values on the corresponding target features (i.e. narcissism and intelligence). λ_{Min} = penalty selected by the elastic net. The means and standard deviations of the performance measures and λ_{Min} are estimated based on the best performing models from the inner folds that are tested on the test data from the outer folds of the tenfold nested cross-validation approach.

Abbreviations: EN, elastic net; IOM, intercept-only model; MAE, mean absolute error; MLR, multiple linear regression; MSE, mean squared error; *r*, bivariate correlation coefficient; *R*², explained variation; *R*²_{Adj.}, adjusted explained variation; RMSE, root mean squared error.

Machine learning approach

Table 2 shows the performance estimates of the nested cross-validated elastic nets predicting narcissism/intelligence based on the respective derived LinkedIn cues. The elastic nets outperformed multiple linear regressions and intercept-only models in absolute prediction performance, with lower MSE, RMSE, and MAE. This indicates that elastic nets made smaller errors in predicting narcissism/intelligence. Interestingly, intercept-only models sometimes outperformed multiple linear regressions, suggesting the latter may overfit the training data. The elastic nets also showed better relative prediction performance, with correlations between predicted and observed target feature values of $r = .35/.41$ for narcissism/intelligence. These correlations were higher than those of the multiple linear regressions ($r = .27/.34$ for narcissism/intelligence). Overall, the results indicate that narcissism/intelligence can be predicted from LinkedIn profiles with accuracy comparable to meta-analytical findings on automated trait inferences from online social network information ($.29 \leq r \leq .40$; Azucar et al., 2018; Settanni et al., 2018).

We found that the LinkedIn cues driving prediction in the elastic nets (Table 3) were generally similar to those identified in the correlation approach, indicating relatively stable cue validities when controlling for other cues and using resampling. All 10 cues correlated with narcissism were important elastic net predictors ($\beta_M \geq |.017|$, $CV_{FI} \geq 9$). The five most important elastic net predictors were all significantly correlated with narcissism and indicate that individuals high on narcissism display (a) interests with more followers, (b) less smiling on profile pictures, (c) more accomplishments related to organizations, (d) background pictures, and (e) public speaking skills. Eleven of the 15 LinkedIn cues correlated with intelligence were important elastic net predictors ($\beta_M \geq |.017|$, $CV_{FI} \geq 7$). Four cues (extensive description of educational stations, number of volunteer experiences, profile length, and English profile) correlated with intelligence but were not important elastic net predictors ($\beta_M \leq |.003|$, $CV_{FI} \leq 3$). The five most important elastic net predictors were all significantly correlated with intelligence and indicate that individuals high on intelligence display (a) schools with more followers, (b) profile pictures above neckline, (c) more accomplishments, (d) less dressed-up, trimmed appearances on profile pictures, and (e) more accomplishments related to honours/awards.

TABLE 3 Regression coefficients of nested cross-validated elastic nets predicting narcissism and intelligence.

Cue	CV _{FI}	β_M	β_{SD}	β_{Full}
Narcissism				
Interests: Median of interests' followers	10	.098	.016	.106
Profile picture: Smiling	10	−.096	.014	−.105
Gender (0/1 = female/male)	10	.089	.016	.094
Accomplishments: Organizations	10	.077	.014	.085
Additional pictures: Background picture	10	.069	.014	.074
Skills: Public speaking	9	.067	.027	.074
Volunteering: Altruistic volunteering	10	.045	.018	.051
Profile picture	10	−.042	.021	−.051
Additional pictures: Additional pictures/videos	10	−.037	.020	−.056
Profile picture: Dressed-up, trimmed appearance	9	.033	.016	.039
Other: Sport activities	10	.031	.013	.040
Recommendations: Received recommendations	9	.030	.023	.038
Experience: Leadership positions	10	.021	.014	.029
Other: Profile in English	9	.019	.014	.026
Skills	9	.017	.014	.019
Profile picture: Professional shot	6	.009	.013	.011
Education: Business studies	5	.008	.011	.003
Accomplishments: Courses	9	.007	.009	.013
Interests	3	−.007	.014	−.015
Profile picture: Stylish/flashy/fashionable appearance	2	.003	.008	.007
About: Extensive About section	2	−.002	.006	.000
Skills: Endorsements	1	−.001	.004	.000
Profile card: Name with title	1	.000	.001	.000
Intercept	10	.000	.000	.000
Intelligence				
Interests: Median of schools' followers	10	.105	.012	.103
Profile picture: Picture above neckline	10	.091	.020	.093
Gender (0/1 = female/male)	10	.090	.024	.090
Accomplishments	10	.080	.020	.084
Profile picture: Dressed-up, trimmed appearance	10	−.075	.015	−.077
Accomplishments: Honours/awards	10	.058	.015	.059
Interests: Groups	10	.056	.019	.057
Accomplishments: Publications	10	.054	.022	.054
Accomplishments: Test scores	10	.045	.010	.045
Education: Averaged marks	9	−.043	.024	−.040
Volunteering: Average duration	10	.040	.013	.042
Skills: Leadership	9	−.033	.020	−.032
Education: Educational stations	7	.019	.021	.019
Profile picture: Charming facial expression	5	−.019	.026	−.009
Accomplishments: Projects	8	.017	.015	.019
Licences and certifications	6	−.013	.015	−.004
Age	5	.011	.015	.000

TABLE 3 (Continued)

Cue	CV _{FI}	β_M	β_{SD}	β_{Full}
Profile picture: Professional shot	4	.010	.021	.000
Profile picture	3	.009	.016	.000
Interests: Schools	3	.007	.011	.000
Education: Marks	3	.004	.009	.000
Volunteering: Volunteer experiences	1	.003	.011	.000
Skills: Industry knowledge	1	-.003	.009	.000
Experience: Leadership positions	3	.003	.005	.000
Other: Profile length	3	.003	.005	.000
Education: Extensive description	1	.002	.006	.000
Education: Business studies	2	-.002	.003	.000
Education: Average duration	1	-.001	.002	.000
Recommendations: Received recommendations	1	.001	.002	.000
Other: Profile in English	1	.000	.000	.000
Intercept	10	.000	.000	.000

Note: CV_{FI} = cross-validation fold incidence, that is, the number of outer folds the regression coefficient of a feature was $\neq 0$; β_M = regression coefficients averaged across outer folds; β_{SD} = standard deviation of regression coefficients across outer folds; β_{Full} = regression coefficients of elastic net trained on full data. Only cues are shown for that CV_{FI} > 0. Cues sorted by $|\beta_M|$. All values on z -scale.

DISCUSSION

This study uncovers LinkedIn's predictive potential to infer traits. Building on the lens model (Brunswik, 1956), we show that LinkedIn profiles contain a valid cue base signalling users' narcissism (e.g. background pictures, public speaking skills) and intelligence (e.g. more accomplishments, schools with more followers), meeting the necessary condition for accurate inferences. Sensitively and consistently utilizing these valid cues, the automated perceiver attained prediction accuracy ($r = .35/.41$ for narcissism/intelligence). This shows that when the sufficient condition for accurate trait inferences is ensured to be met, LinkedIn-based inferences can reach accuracy levels comparable to automated trait inferences from online social networks (Azucar et al., 2018; Settanni et al., 2018) and at the upper bound of recruiters' LinkedIn-based trait inferences (e.g. Roulin & Levashina, 2019; Van de Ven et al., 2017).

Applying the lens model to (automated) trait inferences based on LinkedIn profiles

This study underlines the utility of extending the lens model to automated perceivers to systematically examine trait inference processes (Cannata et al., 2022; Tay et al., 2020): Machine learning algorithms are trained to use valid cues sensitively and consistently, and thus turned out as a useful vehicle to demonstrate that LinkedIn profiles possess predictive potential. Machine learning thus offers a promising alternative methodological approach to lens model studies, avoiding overfitting and prioritizing prediction over explanation, which suits the nature of trait inferences (Yarkoni & Westfall, 2017). Resampling ensures robust results, reducing contradictory conclusions on cue validities. Algorithms like elastic nets effectively handle many intercorrelated cues common in most information bases (James et al., 2013) and feature importance provides practical insights for cue weighting.

The automated approach also helps understand the modest accuracy of human perceivers' LinkedIn-based trait inferences (Roulin & Levashina, 2019; Roulin & Stronach, 2022; Van de Ven et al., 2017). According to the lens model, this may result from (a) a lack of a valid cue base or (b) a lack of a sensitive and consistent use of valid cues. While Roulin and Stronach (2022) challenged the existence of a valid

LinkedIn cue base based on a small exploratory cue set, Fernandez et al. (2021) found support for it based on deductively derived cues. Deriving an extensive cue set from theory and empirical findings, we identified a rich set of valid LinkedIn cues for narcissism/intelligence as robust predictors in nested cross-validated models. As such, we found that the automated perceiver, designed to sensitively and consistently use these valid cues, achieved notable prediction accuracy. This suggests that recruiters' lack of accuracy may stem from insensitive and inconsistent use of valid cues, rather than LinkedIn's limited predictive capability.

Indeed, comparing the cues used by recruiters in prior research with those used by our automated perceiver model reveals human perceivers' lack of sensitivity. For example, recruiters were found to use the number of subscribers to infer trait self-presentation (Van de Ven et al., 2017), but this did not signal narcissism in our study. Thus, the sensitive automated perceiver did not use this information. Vice versa, valid cues like 'profile in English' used by the automated perceiver were not used accordingly by humans (Van de Ven et al., 2017). Human perceivers are also inconsistent in applying judgement rules, whereas statistical models remove these inconsistencies, typically leading to more accurate inferences (Karelaia & Hogarth, 2008).

Extending previous research on valid cues signalling traits on LinkedIn

This study advances the field of trait expression in online professional networks by identifying 10/11 easy-to-interpret LinkedIn cues correlated with narcissism/intelligence tested to be robust in nested cross-validated models. This addresses calls to examine online network information signalling important applicant characteristics (Schroeder et al., 2020). We also enrich the broader literature on trait expression in selection contexts; adding LinkedIn cues to the information bases recruiters can use to infer traits, like resumés (Burns et al., 2014; Härtel et al., 2024), application photos (Fernandez et al., 2017), online social networks (Back, Stopfer, et al., 2010; Stopfer et al., 2014), job interviews (DeGroot & Gooty, 2009), or even handshakes (Stewart et al., 2008). *LinkedIn-specific* cues, like listing interests with many followers for narcissism or schools with many followers for intelligence, offer unique insights unavailable in other sources.

Our findings also reveal that some cues related to trait expression in other contexts lack predictive power on LinkedIn. For example, narcissism often manifests in physical appearance cues (e.g. professional/formal attire; Back, Schmukle, et al., 2010; Vazire et al., 2008) and indicators of online socializing behaviour (e.g. more contacts/frequent posting; Gnambs & Appel, 2018; McCain & Campbell, 2018), but less so on LinkedIn. LinkedIn's professional norms – maintaining a formal appearance, building a professional network, and posting business content – may render these cues less discriminative (Hartwell & Campion, 2020). Similarly, whereas narcissism often involves smiling in everyday situations (Back, Schmukle, et al., 2010), consistent with Van de Ven et al. (2017), we observed *less* smiling to be a narcissism predictor. LinkedIn's norms may prompt most users to smile, masking narcissists' agentic smiling tendencies. This could create space for the expression of antagonistic narcissism, resulting in arrogant expressions with fewer genuine smiles. Indeed, less smiling was a strong predictor of narcissistic rivalry ($\beta_M = -.106$), not admiration ($\beta_M = -.011$; Supplemental Results Section E in Appendix S1, <https://osf.io/4ruqj/>). Overall, this highlights the role of situational strength in trait expression (Meyer et al., 2010).

Automated approach to trait inferences based on LinkedIn

This study advances the literature on automated trait assessments based on online networks by applying the automated approach to LinkedIn using a broad set of theoretically and empirically grounded cues (Bleidorn & Hopwood, 2019; Tay et al., 2020). This way, we obtained high predictivity and the capability to provide explanations for our findings (Yarkoni & Westfall, 2017), both critical for practical applications. The machine learning models' accuracies predicting narcissism/

intelligence were $r = .35/.41$, translating to hit rates of 67.5%/70.5%, indicated by the binomial effect size display (Rosenthal, 2014; Rosenthal & Rubin, 1982). This means individuals testing above average narcissistic/intelligent are correctly identified about 67.5%/70.5% of instances. In *absolute* terms, these effect sizes are *medium* by conventional standards (Cohen, 1988) and (*very*) *large* by modern standards (Funder & Ozer, 2019; Gignac & Szodorai, 2016). On a *relative* basis, these accuracy levels exceed work colleagues' accuracy of trait inferences ($r = .27$; Connelly & Ones, 2010; Youyou et al., 2015) and are comparable to the agreement between self- and supervisor ratings of job performance ($r = .35$; Harris & Schaubroeck, 1988). However, we emphasize not to overstate the validity of automated LinkedIn-based inferences. LinkedIn profiles provide limited insight into a snippet of individuals' traits – they reflect an external view of the publicly revealed professional identity and cannot replace well-established psychometric measures.

Figure 3 compares the prediction accuracies from our study with those of human and automated perceivers drawing from similar sources like resumés, social media profiles, and LinkedIn profiles. Our accuracies ($r_{\text{Narcissism}} = .35/r_{\text{Intelligence}} = .41$) are comparable to meta-analytical results of automated approaches predicting similar traits from social media information like Facebook ($r = .40/.29$; Azucar et al., 2018; Settanni et al., 2018). These meta-analyses describe the relationship between digital traces and traits as 'quite strong' (Azucar et al., 2018, p. 157) and the accuracy level as 'remarkable' (Settanni et al., 2018, p. 217). Given LinkedIn's contextually constrained nature focusing on professional identity, the relative accuracy of automated LinkedIn-based trait inferences is notable. Comparing our results with human perceivers' accuracy using similar information, the automated perceiver appears to perform better.³ Also note that reported accuracy levels for human perceivers often represent averaged inferences from *groups* of perceivers (*aggregated perceiver approach*; Back & Nestler, 2016), with 10 perceivers in Van de Ven et al. (2017), two-to-four in Roulin and Levashina (2019), and six or more in Roulin and Stronach (2022). Single perceivers yield substantially lower accuracy (Back & Nestler, 2016; Van Iddekinge et al., 2016) due to poorer reliability (De Vet et al., 2017). Therefore, automated LinkedIn-based trait inferences appear to outperform single recruiters' cybervetting accuracy as commonly practiced in many organizations.

Practical implications

Recruiters believe they can accurately infer applicants' traits through online profiles (Van Iddekinge et al., 2016). Our study shows that there is some substance to this assumption – LinkedIn profiles possess predictive potential if the contained valid cues are used sensitively and consistently. Yet, although recruiters routinely infer traits based on LinkedIn, they lack accuracy, which can lead to erroneous decisions, like falsely rejecting applicants due to their 'not suitable' personality. Training recruiters to use valid LinkedIn cues sensitively and consistently could improve accuracy (Cole et al., 2005; Karelaia & Hogarth, 2008; Powell & Bourdage, 2016). The machine learning approach identifies which LinkedIn information robustly signal traits, providing a foundation for such training. Going further, a hybrid scoring approach combining automated/human perceivers could offer additional value (Cannata et al., 2022). Automated perceivers could continuously scan LinkedIn users to identify promising candidates not actively job searching and screen large applicant pools quickly, allowing recruiters to focus on candidates with desired traits. Human recruiters may then consider subtle nuances and cue interactions (Cole et al., 2007; Knouse, 1994) missed by automated systems. A hybrid system may be more acceptable to applicants, often reacting negatively to decisions made solely by algorithms (Gonzalez et al., 2022).

³Comparing the automated perceiver's accuracy with previous findings on human perceivers might be seen as favouring the former, as our automated perceiver was exposed only to pre-selected, potentially valid cues. Note that providing the automated perceiver with potentially valid cues was a pragmatic decision to give it access to meaningful LinkedIn information that was time-consuming to code manually. In contrast, human perceivers with their rich perception have access to much more LinkedIn information. Even if all this additional information was invalid, it would not substantially harm automated perceivers' accuracy due to their sensitivity. Accordingly, pre-selection of cues likely underestimates rather than overestimates automated perceivers' accuracy.

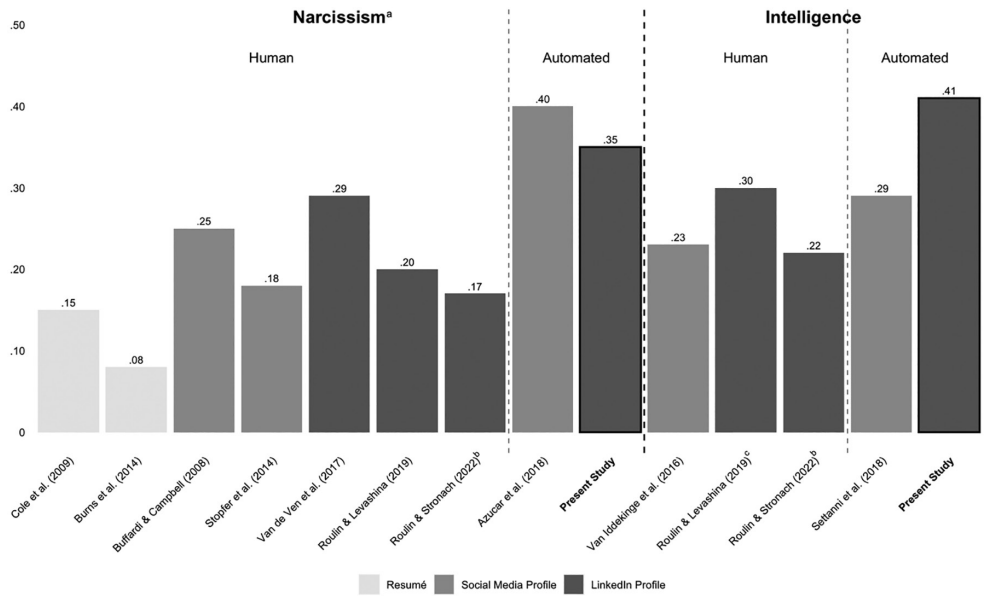


FIGURE 3 Prediction accuracy in the present study compared to human and automated perceivers' accuracy reported in prior research. *Note.* Prediction accuracy is defined as the correlation (r) between observed trait test scores and inferred trait scores based on the respective information base. ^aWhenever available, we consulted results directly referring to narcissism and otherwise drew on related constructs, namely need for popularity, trait self-presentation, honesty-humility, and extraversion. ^bAveraged correlation for I-O psychology students and hiring professionals as raters. ^cThe correlation between cognitive ability test scores and aggregated LinkedIn-based cognitive ability ratings by the four perceivers across the two time points was much lower ($r = -.02$, see Table S2).

Besides, some may be inclined to interpret our findings suggesting that LinkedIn-based trait inferences should be delegated to automated perceivers due to their accuracy and resource savings. Indeed, automated trait inferences *could* offer a less intrusive, cost-effective, and scalable supplement to traditional psychological assessments (Morgeson et al., 2007). Yet, this is too early to call. Although our study provides robust results by applying resampling, it remains a single study with a medium-sized sample of 406 LinkedIn users. Also, automated LinkedIn-based assessments pose practical and fairness challenges, so further research is needed before endorsing them.

As such, LinkedIn use is dynamic, and cues may evolve over time, for instance, as users become acquainted with the platform, they may upload background pictures more frequently. Also, applicants already tailor their resumés for favourable algorithmic scoring (Caprino, 2019; Weed, 2021) and will likely adjust their LinkedIn profiles once automated assessments become common. These dynamic processes demand continuous cross-validation and algorithmic recalibration, necessitating efficient, automated cue collection processes. Yet, we coded LinkedIn cues manually, which was time-consuming and labour-intensive. The imminent shift towards automated cue extraction – initially demonstrated by studies leveraging web scraping for textual data (Landers et al., 2016; Youyou et al., 2015) and deep learning for image data (Wei & Stillwell, 2017) – is gaining momentum in rapid developments of large language and computer vision models. This will likely diminish manual coding. Generative pre-trained transformers can parse textual information to evaluate self-promotional content in job descriptions or categorize skills, and convolutional neural networks can scrutinize profile images to discern facial expressions, poses, or attire choices.

Another concern involves subgroups (dis-)favoured by the algorithm. The algorithm is programmed to minimize error when predicting test scores, which may show subgroup differences for attributes like gender (Kheloui et al., 2023) and ethnicity (Roth et al., 2001). This carries the risk that the algorithm learns the ‘predictive value’ of subgroup differences. In our study, for example, men

scored higher than women on intelligence ($\Delta M = .27$, $t(345.32) = 3.36$, $p < .001$), leading the algorithm to use gender to predict intelligence. Using the elastic net, an interpretable algorithm revealing the cues it uses, discrimination can be addressed by removing biased cues from the model or adjusting their regression weights. Alternatively, subgroup differences can be mitigated in the training data through fair test construction or statistical corrections (Hough et al., 2001). Implementing such measures may then even hold potentials to *reduce* discrimination, as automated systems are principally immune to (unconscious) human biases (Mönke et al., 2024; Zhang et al., 2020). It is also crucial to test whether the algorithm exhibits varying accuracy across subgroups, constituting another form of discrimination.

Limitations and directions for future research

Generalizing our findings to the LinkedIn population depends on the representativeness of our sample, which was not restricted to subpopulations like students/recent graduates (cf. Fernandez et al., 2021; Roulin & Levashina, 2019), but meant to reflect LinkedIn's mixed professional population. Indeed, a comparison of age distributions between the LinkedIn population and our sample indicates reasonable alignment.⁴ Yet, our sample is not representative in several respects. For example, we examined German-speaking LinkedIn users. Cultural differences shape motivations and self-presentation tactics related to using online networks (Jackson & Wang, 2013; Rui & Stefanone, 2013) and may thus affect trait expression and LinkedIn's predictive potential. For instance, whereas uploading background pictures may be a valid narcissism indicator in Germany, it may be a more commonly used feature in other countries, providing less insight into narcissism levels. Also, in creative industries like web design or advertising, it may merely serve as a portfolio. Future research should confirm prediction accuracies and cue validities in larger samples mirroring the LinkedIn population across dimensions like age, gender, education, job roles, and industries.

We optimized the automated perceiver to predict trait test scores. The 'gold-standard' would be aggregating various trait indicators as prediction criteria, like self-reports, informant reports, and behavioural assessments (Cannata et al., 2022). Further steps preceding practical applications are then examining the extent automated trait inferences predict work outcomes, particularly job performance (criterion-related validity; Cubrich et al., 2021; Klumper et al., 2012), and explain incremental variance beyond (i) recruiters' LinkedIn inferred traits and (ii) established tests. More research is needed on adverse impact (Hunkenschroer & Luetge, 2022) and applicant reactions (Ostrom et al., 2023). Until then, organizations should avoid strongly relying on automated LinkedIn-based trait inferences.

Our findings suggest that automated perceivers infer traits on LinkedIn more accurately than single recruiters, who appear to be less sensitive and/or consistent in using valid cues. Yet, evidence on recruiters' cue utilizations is limited (Roulin & Stronach, 2022; Van de Ven et al., 2017), and our study lacks a direct comparison, warranting caution in drawing final conclusions. Initial evidence suggests that recruiters may indeed use non-valid cues and overlook valid ones, indicating a lack of sensitivity. Also, human perceivers are typically inconsistent in applying judgement rules. Future research should compare automated and human perceivers directly, disentangling whether recruiters can distinguish valid from non-valid LinkedIn cues and apply consistent judgement rules. Such research could inform training programs educating about rarely used valid cues and how to use them consistently. It would be fascinating to see if the accuracy of human inferences could be improved through such measures, and whether they could even surpass automated perceivers by capturing subtle idiosyncratic information.

⁴Percentage of age groups in the LinkedIn population (Dixon, 2023) versus our sample: 18–24 years, 22% versus 34%; 25–34 years, 60% versus 48%; 35–54 years, both 15%; ≥ 55 years, both 3%. Although our sample skews slightly younger on average ($M_{\text{age}} = 29$), it seems to align more closely with the LinkedIn population compared to prior studies (e.g. $M_{\text{age}} = 21/41$ in Roulin & Levashina, 2019/Roulin & Stronach, 2022).

CONCLUSION

We illuminate LinkedIn's predictive potential to infer individual traits adopting a lens model approach. This involved holistically mapping the cue environment according to the principle of good information by focusing on relevant cues with high content validity. We fed these cues to a transparent machine learning algorithm sensitively and consistently using them to make predictions. We identified robust valid LinkedIn cues holding notable predictive potential to infer narcissism/intelligence. We hope our approach inspires further research on the potentials and limitations of (automated) trait inferences in selection contexts.

AUTHOR CONTRIBUTIONS

Tobias M. Härtel: Methodology; data curation; visualization; writing – original draft; writing – review and editing; formal analysis; project administration; investigation; conceptualization. **Benedikt A. Schuler:** Formal analysis; conceptualization; methodology; writing – review and editing; visualization. **Mitja D. Back:** Supervision; resources; writing – review and editing.

FUNDING INFORMATION

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTEREST STATEMENT

We have no information to disclose about any interests or activities that might be viewed as influencing the present research.

DATA AVAILABILITY STATEMENT

Supplemental materials including a codebook, data, analysis code, and supplemental results are provided on the Open Science Framework and can be retrieved from the following link: <https://osf.io/4ruqj/>.

ORCID

Tobias M. Härtel  <https://orcid.org/0000-0002-6015-4525>

Benedikt A. Schuler  <https://orcid.org/0000-0002-3030-2881>

Mitja D. Back  <https://orcid.org/0000-0003-2186-1558>

REFERENCES

- Aabo, T., & Eriksen, N. B. (2018). Corporate risk and the humpback of CEO narcissism. *Review of Behavioral Finance*, 10(3), 252–273. <https://doi.org/10.1108/RBF-07-2017-0070>
- Ackerman, P. L., Bowen, K. R., Beier, M. E., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology*, 93(4), 797–825. <https://doi.org/10.1037/0022-0663.93.4.797>
- Alexander, L., Mulfinger, E., & Oswald, F. L. (2020). Using big data and machine learning in personality measurement: Opportunities and challenges. *European Journal of Personality*, 34(5), 632–648. <https://doi.org/10.1002/per.2305>
- Alexander, R. A. (1990). A note on averaging correlations. *Bulletin of the Psychonomic Society*, 28(4), 335–336. <https://doi.org/10.3758/BF03334037>
- Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*, 124, 150–159. <https://doi.org/10.1016/j.paid.2017.12.018>
- Back, M. D. (2018). The narcissistic admiration and rivalry concept. In A. D. Hermann, A. B. Brunell, & J. D. Foster (Eds.), *Handbook of trait narcissism: Key advances, research methods, and controversies* (pp. 57–67). Springer. https://doi.org/10.1007/978-3-319-92171-6_6
- Back, M. D., Küfner, A. C. P., Dufner, M., Gerlach, T. M., Rauthmann, J. F., & Denissen, J. J. A. (2013). Narcissistic admiration and rivalry: Disentangling the bright and dark sides of narcissism. *Journal of Personality and Social Psychology*, 105(6), 1013–1037. <https://doi.org/10.1037/a0034431>
- Back, M. D., & Nestler, S. (2016). Accuracy of judging personality. In J. A. Hall, M. Schmid Mast, & T. V. West (Eds.), *The social psychology of perceiving others accurately* (pp. 98–124). Cambridge University Press. <https://doi.org/10.1017/CBO9781316181959.005>

- Back, M. D., Schmukle, S. C., & Egloff, B. (2010). Why are narcissists so charming at first sight? Decoding the narcissism–popularity link at zero acquaintance. *Journal of Personality and Social Psychology*, 98(1), 132–145. <https://doi.org/10.1037/a0016338>
- Back, M. D., Stopfer, J. M., Vazire, S., Gaddis, S., Schmukle, S. C., Egloff, B., & Gosling, S. D. (2010). Facebook profiles reflect actual personality, not self-idealization. *Psychological Science*, 21(3), 372–374. <https://doi.org/10.1177/0956797609360756>
- Bensch, D., Paulhus, D. L., Stankov, L., & Ziegler, M. (2019). Teasing apart overclaiming, overconfidence, and socially desirable responding. *Assessment*, 26(3), 351–363. <https://doi.org/10.1177/1073191117700268>
- Benson, A. J., Jordan, C. H., & Christie, A. M. (2016). Narcissistic reactions to subordinate role assignment: The case of the narcissistic follower. *Personality and Social Psychology Bulletin*, 42(7), 985–999. <https://doi.org/10.1177/0146167216649608>
- Blair, C. A., Hoffman, B. J., & Helland, K. R. (2008). Narcissism in organizations: A multisource appraisal reflects different perspectives. *Human Performance*, 21(3), 254–276. <https://doi.org/10.1080/08959280802137705>
- Bleidorn, W., & Hopwood, C. J. (2019). Using machine learning to advance personality assessment and theory. *Personality and Social Psychology Review*, 23(2), 190–203. <https://doi.org/10.1177/1088868318772990>
- Borkenau, P., Mauer, N., Riemann, R., Spinath, F. M., & Angleitner, A. (2004). Thin slices of behavior as cues of personality and intelligence. *Journal of Personality and Social Psychology*, 86(4), 599–614. <https://doi.org/10.1037/0022-3514.86.4.599>
- Brailovskaia, J., Bierhoff, H.-W., & Margraf, J. (2019). How to identify narcissism with 13 items? Validation of the German Narcissistic Personality Inventory–13 (G-NPI-13). *Assessment*, 26(4), 630–644. <https://doi.org/10.1177/1073191117740625>
- Breit, M., Scherrer, V., Tucker-Drob, E. M., & Preckel, F. (2024). The stability of cognitive abilities: A meta-analytic review of longitudinal studies. *Psychological Bulletin*, 150, 399–439. <https://doi.org/10.1037/bul0000425>
- Brown, B. K., & Campion, M. A. (1994). Biodata phenomenology: Recruiters' perceptions and use of biographical information in resume screening. *Journal of Applied Psychology*, 79(6), 897–908. <https://doi.org/10.1037/0021-9010.79.6.897>
- Brunswick, E. (1956). *Perception and the representative design of psychological experiments*. University of California Press.
- Buffardi, L. E., & Campbell, W. K. (2008). Narcissism and social networking web sites. *Personality and Social Psychology Bulletin*, 34(10), 1303–1314. <https://doi.org/10.1177/0146167208320061>
- Burns, G. N., Christiansen, N. D., Morris, M. B., Periard, D. A., & Coaster, J. A. (2014). Effects of applicant personality on resume evaluations. *Journal of Business and Psychology*, 29(4), 573–591. <https://doi.org/10.1007/s10869-014-9349-6>
- Campbell, W. K., & Campbell, S. M. (2009). On the self-regulatory dynamics created by the peculiar benefits and costs of narcissism: A contextual reinforcement model and examination of leadership. *Self and Identity*, 8(2–3), 214–232. <https://doi.org/10.1080/15298860802505129>
- Campbell, W. K., Hoffman, B. J., Campbell, S. M., & Marchisio, G. (2011). Narcissism in organizational contexts. *Human Resource Management Review*, 21(4), 268–284. <https://doi.org/10.1016/j.hrmr.2010.10.007>
- Cannata, D., Breil, S. M., Lepri, B., Back, M. D., & O'Hora, D. (2022). Toward an integrative approach to nonverbal personality detection: Connecting psychological and artificial intelligence research. *Technology, Mind, and Behavior*, 3(2), 1–16. <https://doi.org/10.1037/tmb0000054>
- Caprino, K. (2019). *How to write a resume that passes the artificial intelligence test*. Forbes. <https://www.forbes.com/sites/kathycaprino/2019/03/23/how-to-write-a-resume-that-passes-the-artificial-intelligence-test/>
- Carey, A. L., Brucks, M. S., Küfner, A. C. P., Holtzman, N. S., Große Deters, F., Back, M. D., Donnellan, M. B., Pennebaker, J. W., & Mehl, M. R. (2015). Narcissism and the use of personal pronouns revisited. *Journal of Personality and Social Psychology*, 109(3), e1–e15. <https://doi.org/10.1037/pspp0000029>
- Cattell, R. B. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54(1), 1–22. <https://doi.org/10.1037/h0046743>
- Cawley, G. C., & Talbot, N. L. C. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11, 2079–2107. <https://doi.org/10.48550/arXiv.1809.09446>
- Chen, J., Qiu, L., & Ho, M.-H. R. (2020). A meta-analysis of linguistic markers of extraversion: Positive emotion and social process words. *Journal of Research in Personality*, 89, 104035. <https://doi.org/10.1016/j.jrp.2020.104035>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Erlbaum.
- Cole, M. S., Feild, H. S., & Giles, W. F. (2003). What can we uncover about applicants based on their resumes? A field study. *Applied HRM Research*, 8(2), 51–62.
- Cole, M. S., Feild, H. S., Giles, W. F., & Harris, S. G. (2009). Recruiters' inferences of applicant personality based on resume screening: Do paper people have a personality? *Journal of Business and Psychology*, 24(1), 5–18. <https://doi.org/10.1007/s10869-008-9086-9>
- Cole, M. S., Feild, H. S., & Stafford, J. O. (2005). Validity of resumé reviewers' inferences concerning applicant personality based on resumé evaluation. *International Journal of Selection and Assessment*, 13(4), 321–324. <https://doi.org/10.1111/j.1468-2389.2005.00329.x>
- Cole, M. S., Rubin, R. S., Feild, H. S., & Giles, W. F. (2007). Recruiters' perceptions and use of applicant resumé information: Screening the recent graduate. *Applied Psychology*, 56(2), 319–343. <https://doi.org/10.1111/j.1464-0597.2007.00288.x>
- 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>

- Cragun, O. R., Olsen, K. J., & Wright, P. M. (2020). Making CEO narcissism research great: A review and meta-analysis of CEO narcissism. *Journal of Management*, 46(6), 908–936. <https://doi.org/10.1177/0149206319892678>
- Cubrich, M., King, R. T., Mracek, D. L., Strong, J. M. G., Hassenkamp, K., Vaughn, D., & Dudley, N. M. (2021). Examining the criterion-related validity evidence of LinkedIn profile elements in an applied sample. *Computers in Human Behavior*, 120, 106742. <https://doi.org/10.1016/j.chb.2021.106742>
- De Fruyt, F., De Clercq, B. J., Miller, J., Rolland, J., Jung, S., Taris, R., Furnham, A., & Van Hiel, A. (2009). Assessing personality at risk in personnel selection and development. *European Journal of Personality*, 23(1), 51–69. <https://doi.org/10.1002/per.703>
- De Vet, H. C. W. d., Mokkink, L. B., Mosmuller, D. G., & Terwee, C. B. (2017). Spearman–Brown prophecy formula and Cronbach's alpha: Different faces of reliability and opportunities for new applications. *Journal of Clinical Epidemiology*, 85, 45–49. <https://doi.org/10.1016/j.jclinepi.2017.01.013>
- DeGroot, T., & Gooty, J. (2009). Can nonverbal cues be used to make meaningful personality attributions in employment interviews? *Journal of Business and Psychology*, 24(2), 179–192. <https://doi.org/10.1007/s10869-009-9098-0>
- Dixon, S. J. (2023). *LinkedIn: Distribution of global audiences 2023, by age group*. Statista. <https://www.statista.com/statistics/273505/global-linkedin-age-group/>
- Fernandez, S., Stöcklin, M., Terrier, L., & Kim, S. (2021). Using available signals on LinkedIn for personality assessment. *Journal of Research in Personality*, 93, 104122. <https://doi.org/10.1016/j.jrp.2021.104122>
- Fernandez, S., Stosic, G., & Terrier, L. (2017). Does your résumé photograph tell who you are? *Personality and Individual Differences*, 104, 186–189. <https://doi.org/10.1016/j.paid.2016.08.006>
- Friedman, J., Hastie, T., & Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software*, 33(1), 1–22. <https://doi.org/10.18637/jss.v033.i01>
- Funder, D. C., & Ozer, D. J. (2019). Evaluating effect size in psychological research: Sense and nonsense. *Advances in Methods and Practices in Psychological Science*, 2(2), 156–168. <https://doi.org/10.1177/2515245919847202>
- Garcia, D., & Sikström, S. (2014). The dark side of Facebook: Semantic representations of status updates predict the dark triad of personality. *Personality and Individual Differences*, 67, 92–96. <https://doi.org/10.1016/j.paid.2013.10.001>
- Gentile, B., Miller, J. D., Hoffman, B. J., Reidy, D. E., Zeichner, A., & Campbell, W. K. (2013). A test of two brief measures of grandiose narcissism: The Narcissistic Personality Inventory–13 and the Narcissistic Personality Inventory–16. *Psychological Assessment*, 25(4), 1120–1136. <https://doi.org/10.1037/a0033192>
- GESIS – Leibniz Institute for the Social Sciences. (n.d.). *Open Access Repository for Measurement Instruments*. <https://zis.gesis.org/en>
- Gignac, G. E., & Szodorai, E. T. (2016). Effect size guidelines for individual differences researchers. *Personality and Individual Differences*, 102, 74–78. <https://doi.org/10.1016/j.paid.2016.06.069>
- Gnambs, T., & Appel, M. (2018). Narcissism and social networking behavior: A meta-analysis. *Journal of Personality*, 86(2), 200–212. <https://doi.org/10.1111/jopy.12305>
- Gonzalez, M. F., Liu, W., Shirase, L., Tomczak, D. L., Lobbe, C. E., Justenhoven, R., & Martin, N. R. (2022). Allying with AI? Reactions toward human-based, AI/ML-based, and augmented hiring processes. *Computers in Human Behavior*, 130, 107179. <https://doi.org/10.1016/j.chb.2022.107179>
- Goretzko, D., & Israel, L. S. F. (2022). Pitfalls of machine learning-based personnel selection: Fairness, transparency, and data quality. *Journal of Personnel Psychology*, 21(1), 37–47. <https://doi.org/10.1027/1866-5888/a000287>
- Gottfredson, L. S. (1997). Why g matters: The complexity of everyday life. *Intelligence*, 24(1), 79–132. [https://doi.org/10.1016/S0160-2896\(97\)90014-3](https://doi.org/10.1016/S0160-2896(97)90014-3)
- Grijalva, E., Harms, P. D., Newman, D. A., Gaddis, B. H., & Fraley, R. C. (2015). Narcissism and leadership: A meta-analytic review of linear and nonlinear relationships. *Personnel Psychology*, 68(1), 1–47. <https://doi.org/10.1111/peps.12072>
- Hammond, K. R. (1996). *Human judgment and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice*. Oxford University Press.
- Harris, M. M., & Schaubroeck, J. (1988). A meta-analysis of self-supervisor, self-peer, and peer-supervisor ratings. *Personnel Psychology*, 41(1), 43–62. <https://doi.org/10.1111/j.1744-6570.1988.tb00631.x>
- Härtel, T. M., Breil, S. M., Grunenberg, E., & Back, M. D. (2024). Relationships between résumé cues and applicants' personality. *Applied Psychology*, 1–44. <https://doi.org/10.1111/apps.12522>
- Härtel, T. M., Leckelt, M., Grosz, M. P., Küfner, A. C. P., Geukes, K., & Back, M. D. (2021). Pathways from narcissism to leadership emergence in social groups. *European Journal of Personality*, 35(5), 1–23. <https://doi.org/10.1177/08902070211046266>
- Hartwell, C. J., & Campion, M. A. (2020). Getting social in selection: How social networking website content is perceived and used in hiring. *International Journal of Selection and Assessment*, 28(1), 1–16. <https://doi.org/10.1111/ijsa.12273>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Heydasch, T., Haubrich, J., & Renner, K.-H. (2013). The short version of the Hagen matrices test (HMT-S): 6-item induction intelligence test. *Methods, Data, Analyses*, 7(2), 2. <https://doi.org/10.12758/mda.2013.011>
- Heydasch, T., Renner, K.-H., & Haubrich, J. (2012). Der Hagener Matrizen-Test (HMT): Ein kostenloser, web-basierter Intelligenztest. 17. *Coping-Tagung*.

- Holtzman, N. S., Tackman, A. M., Carey, A. L., Brucks, M. S., Küfner, A. C. P., Deters, F. G., Back, M. D., Donnellan, M. B., Pennebaker, J. W., Sherman, R. A., & Mehl, M. R. (2019). Linguistic markers of grandiose narcissism: A LIWC analysis of 15 samples. *Journal of Language and Social Psychology*, 38(5–6), 773–786. <https://doi.org/10.1177/0261927X19871084>
- Hough, L. M., Oswald, F. L., & Ployhart, R. E. (2001). Determinants, detection and amelioration of adverse impact in personnel selection procedures: Issues, evidence and lessons learned. *International Journal of Selection and Assessment*, 9(1–2), 152–194. <https://doi.org/10.1111/1468-2389.00171>
- Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: A review and research agenda. *Journal of Business Ethics*, 178(4), 977–1007. <https://doi.org/10.1007/s10551-022-05049-6>
- Jackson, L. A., & Wang, J.-L. (2013). Cultural differences in social networking site use: A comparative study of China and the United States. *Computers in Human Behavior*, 29(3), 910–921. <https://doi.org/10.1016/j.chb.2012.11.024>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning*. Springer. <https://doi.org/10.1007/978-1-4614-7138-7>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- Judge, T. A., Colbert, A. E., & Ilies, R. (2004). Intelligence and leadership: A quantitative review and test of theoretical propositions. *Journal of Applied Psychology*, 89(3), 542–552. <https://doi.org/10.1037/0021-9010.89.3.542>
- Judge, T. A., & LePine, J. A. (2007). The bright and dark sides of personality: Implications for personnel selection in individual and team contexts. In J. Langan-Fox, C. L. Cooper, & R. J. Klimoski (Eds.), *Research companion to the dysfunctional workplace: Management challenges and symptoms* (pp. 332–355). Edward Elgar Publishing. <https://doi.org/10.4337/9781847207081.00028>
- Judge, T. A., LePine, J. A., & Rich, B. L. (2006). Loving yourself abundantly: Relationship of the narcissistic personality to self and other perceptions of workplace deviance, leadership, and task and contextual performance. *Journal of Applied Psychology*, 91(4), 762–776. <https://doi.org/10.1037/0021-9010.91.4.762>
- Karelaia, N., & Hogarth, R. M. (2008). Determinants of linear judgment: A meta-analysis of lens model studies. *Psychological Bulletin*, 134(3), 404–426. <https://doi.org/10.1037/0033-2909.134.3.404>
- Kheloui, S., Jacmin-Park, S., Larocque, O., Kerr, P., Rossi, M., Cartier, L., & Juster, R.-P. (2023). Sex/gender differences in cognitive abilities. *Neuroscience & Biobehavioral Reviews*, 152, 105333. <https://doi.org/10.1016/j.neubiorev.2023.105333>
- Kluemper, D. H., & Rosen, P. A. (2009). Future employment selection methods: Evaluating social networking web sites. *Journal of Managerial Psychology*, 24(6), 567–580. <https://doi.org/10.1108/02683940910974134>
- Kluemper, D. H., Rosen, P. A., & Mossholder, K. W. (2012). Social networking websites, personality ratings, and the organizational context: More than meets the eye? *Journal of Applied Social Psychology*, 42(5), 1143–1172. <https://doi.org/10.1111/j.1559-1816.2011.00881.x>
- Knouse, S. B. (1994). Impressions of the resume: The effects of applicant education, experience, and impression management. *Journal of Business and Psychology*, 9(1), 33–45. <https://doi.org/10.1007/BF02230985>
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 14(2), 1137–1143.
- König, C. J., Klehe, U.-C., Berchtold, M., & Kleinmann, M. (2010). Reasons for being selective when choosing personnel selection procedures. *International Journal of Selection and Assessment*, 18(1), 17–27. <https://doi.org/10.1111/j.1468-2389.2010.00485.x>
- 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>
- Kosinski, M., Bachrach, Y., Kohli, P., Stillwell, D., & Graepel, T. (2014). Manifestations of user personality in website choice and behaviour on online social networks. *Machine Learning*, 95(3), 357–380. <https://doi.org/10.1007/s10994-013-5415-y>
- Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the National Academy of Sciences of the United States of America*, 110(15), 5802–5805. <https://doi.org/10.1073/pnas.1218772110>
- Kowalski, C. M., Rogoza, R., Vernon, P. A., & Schermer, J. A. (2018). The dark triad and the self-presentation variables of socially desirable responding and self-monitoring. *Personality and Individual Differences*, 120, 234–237. <https://doi.org/10.1016/j.paid.2017.09.007>
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86(1), 148–161. <https://doi.org/10.1037/0022-3514.86.1.148>
- Landers, R. N., Brusso, R. C., Cavanaugh, K. J., & Collmus, A. B. (2016). A primer on theory-driven web scraping: Automatic extraction of big data from the internet for use in psychological research. *Psychological Methods*, 21(4), 475–492. <https://doi.org/10.1037/met0000081>
- Lantz, B. (2019). *Machine learning with R: Expert techniques for predictive modeling*. Packt Publishing Ltd.
- LinkedIn. (2023). *About Us*. <https://news.linkedin.com/about-us>
- Marshall, T. C., Ferenczi, N., Lefringhausen, K., Hill, S., & Deng, J. (2020). Intellectual, narcissistic, or Machiavellian? How twitter users differ from Facebook-only users, why they use twitter, and what they tweet about. *Psychology of Popular Media*, 9(1), 14–30. <https://doi.org/10.1037/ppm0000209>

- McCain, J. L., & Campbell, W. K. (2018). Narcissism and social media use: A meta-analytic review. *Psychology of Popular Media Culture*, 7(3), 308–327. <https://doi.org/10.1037/ppm0000137>
- Meyer, R. D., Dalal, R. S., & Hermdia, R. (2010). A review and synthesis of situational strength in the organizational sciences. *Journal of Management*, 36(1), 121–140. <https://doi.org/10.1177/0149206309349309>
- Mönke, F. W., Lievens, F., Hess, U., & Schäpers, P. (2024). Politics speak louder than skills: Political similarity effects in hireability judgments in multiparty contexts and the role of political interest. *Journal of Applied Psychology*, 109(1), 1–12. <https://doi.org/10.1037/apl0001124>
- Mönke, F. W., & Schäpers, P. (2022). Too early to call: What we do (not) know about the validity of cybervetting. *Industrial and Organizational Psychology*, 15(3), 334–341. <https://doi.org/10.1017/iop.2022.51>
- Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel selection contexts. *Personnel Psychology*, 60(3), 683–729. <https://doi.org/10.1111/j.1744-6570.2007.00089.x>
- Nestler, S., & Back, M. D. (2013). Applications and extensions of the lens model to understand interpersonal judgments at zero acquaintance. *Current Directions in Psychological Science*, 22(5), 374–379. <https://doi.org/10.1177/0963721413486148>
- Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58(2), 367–408. <https://doi.org/10.1111/j.1744-6570.2005.00515.x>
- Nisbett, R. E., Aronson, J., Blair, C., Dickens, W., Flynn, J., Halpern, D. F., & Turkheimer, E. (2012). Intelligence: New findings and theoretical developments. *American Psychologist*, 67(2), 130–159. <https://doi.org/10.1037/a0026699>
- O'Boyle, E. H., Forsyth, D. R., Banks, G. C., & McDaniel, M. A. (2012). A meta-analysis of the dark triad and work behavior: A social exchange perspective. *Journal of Applied Psychology*, 97(3), 557–579. <https://doi.org/10.1037/a0025679>
- Olkin, I., & Pratt, J. W. (1958). Unbiased estimation of certain correlation coefficients. *The Annals of Mathematical Statistics*, 29(1), 201–211.
- Ones, D. S., & Dilchert, S. (2009). How special are executives? How special should executive selection be? Observations and recommendations. *Industrial and Organizational Psychology*, 2(2), 163–170. <https://doi.org/10.1111/j.1754-9434.2009.01127.x>
- Oostrom, J. K., Holtrop, D., Koutsoumpis, A., van Breda, W., Ghassemi, S., & de Vries, R. E. (2023). Applicant reactions to algorithm- versus recruiter-based evaluations of an asynchronous video interview and a personality inventory. *Journal of Occupational and Organizational Psychology*, 97, 160–189. <https://doi.org/10.1111/joop.12465>
- Paulhus, D. L., Westlake, B. G., Calvez, S. S., & Harms, P. D. (2013). Self-presentation style in job interviews: The role of personality and culture. *Journal of Applied Social Psychology*, 43(10), 2042–2059. <https://doi.org/10.1111/jasp.12157>
- Powell, D. M., & Bourdage, J. S. (2016). The detection of personality traits in employment interviews: Can “good judges” be trained? *Personality and Individual Differences*, 94, 194–199. <https://doi.org/10.1016/j.paid.2016.01.009>
- R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. <http://www.R-project.org/>
- Rosenthal, R. (2014). Binomial effect size display. In N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, & J. L. Teugels (Eds.), *Wiley StatsRef: Statistics reference online*. John Wiley & Sons. <https://doi.org/10.1002/9781118445112.stat06256>
- Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, 74(2), 166–169. <https://doi.org/10.1037/0022-0663.74.2.166>
- Roth, P. L., Bevier, C. A., Bobko, P., Switzer Iii, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, 54(2), 297–330. <https://doi.org/10.1111/j.1744-6570.2001.tb00094.x>
- Roth, P. L., Bobko, P., Van Iddekinge, C. H., & Thatcher, J. B. (2016). Social media in employee-selection-related decisions: A research agenda for uncharted territory. *Journal of Management*, 42(1), 269–298. <https://doi.org/10.1177/0149206313503018>
- Rotolo, C. T., & Bracken, D. W. (2022). Assessing the dark side: Making informed decisions throughout the leadership life-cycle. In D. Lusk & T. L. Hayes (Eds.), *Overcoming bad leadership in organizations: A handbook for leaders, talent management professionals, and psychologists* (pp. 277–324). Oxford University Press. <https://doi.org/10.1093/oso/9780197552759.003.0014>
- Roulin, N., & Levashina, J. (2019). LinkedIn as a new selection method: Psychometric properties and assessment approach. *Personnel Psychology*, 72(2), 187–211. <https://doi.org/10.1111/peps.12296>
- Roulin, N., & Stronach, R. (2022). LinkedIn-based assessments of applicant personality, cognitive ability, and likelihood of organizational citizenship behaviors: Comparing self-, other-, and language-based automated ratings. *International Journal of Selection and Assessment*, 30(4), 503–525. <https://doi.org/10.1111/ijsa.12396>
- RStudio Team. (2016). *RStudio: Integrated development for R*. RStudio, Inc. <https://www.Rproject.org/>
- Rui, J., & Stefanone, M. A. (2013). Strategic self-presentation online: A cross-cultural study. *Computers in Human Behavior*, 29(1), 110–118. <https://doi.org/10.1016/j.chb.2012.07.022>
- Sackett, P. R., Zhang, C., Berry, C. M., & Lievens, F. (2022). 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>
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., & Fruyt, F. D. (2003). International validity generalization of GMA and cognitive abilities: A European community meta-analysis. *Personnel Psychology*, 56(3), 573–605. <https://doi.org/10.1111/j.1744-6570.2003.tb00751.x>

- Satornino, C. B., Allen, A., Shi, H., & Bolander, W. (2023). Understanding the performance effects of “dark” salesperson traits: Machiavellianism, narcissism, and psychopathy. *Journal of Marketing*, 87(2), 298–318. <https://doi.org/10.1177/00222429221113254>
- Saw, J. G., Yang, M. C. K., & Mo, T. C. (1984). Chebyshev inequality with estimated mean and variance. *The American Statistician*, 38(2), 130–132. <https://doi.org/10.1080/00031305.1984.10483182>
- Schipolowski, S., Wilhelm, O., Schroeders, U., Kovaleva, A., Kemper, C. J., & Rammstedt, B. (2013). BEFKI GC-K: A short scale for the measurement of crystallized intelligence. *Methods, Data, Analyses*, 7(2), 153–181. <https://doi.org/10.12758/mda.2013.010>
- Schroeder, A. N., Odd, K. R., & Whitaker, J. H. (2020). Agree to disagree: Examining the psychometrics of cybervetting. *Journal of Managerial Psychology*, 35(5), 435–450. <https://doi.org/10.1108/JMP-09-2018-0420>
- Schwarzinger, D. (2022). *The dark triad of personality in personnel selection*. Hogrefe.
- Settanni, M., Azucar, D., & Marengo, D. (2018). Predicting individual characteristics from digital traces on social media: A meta-analysis. *Cyberpsychology, Behavior, and Social Networking*, 21(4), 217–228. <https://doi.org/10.1089/cyber.2017.0384>
- Stekhoven, D. J. (2022). missForest: Nonparametric missing value imputation using random forest. R Package Version 1.5.
- Stekhoven, D. J., & Bühlmann, P. (2012). MissForest—Non-parametric missing value imputation for mixed-type data. *Bioinformatics*, 28(1), 112–118. <https://doi.org/10.1093/bioinformatics/btr597>
- Stewart, G. L., Dustin, S. L., Barrick, M. R., & Darnold, T. C. (2008). Exploring the handshake in employment interviews. *Journal of Applied Psychology*, 93(5), 1139–1146. <https://doi.org/10.1037/0021-9010.93.5.1139>
- Stopfer, J. M., Egloff, B., Nestler, S., & Back, M. D. (2014). Personality expression and impression formation in online social networks: An integrative approach to understanding the processes of accuracy, impression management and meta-accuracy. *European Journal of Personality*, 28(1), 73–94. <https://doi.org/10.1002/per.1935>
- Stoughton, J. W., Thompson, L. F., & Meade, A. W. (2015). Examining applicant reactions to the use of social networking websites in pre-employment screening. *Journal of Business and Psychology*, 30(1), 73–88. <https://doi.org/10.1007/s10869-013-9333-6>
- Sumner, C., Byers, A., Booshever, R., & Park, G. J. (2012). Predicting dark triad personality traits from Twitter usage and a linguistic analysis of tweets, 2, 386–393. <https://doi.org/10.1109/ICMLA.2012.218>
- Tay, L., Woo, S. E., Hickman, L., & Saeef, R. M. (2020). Psychometric and validity issues in machine learning approaches to personality assessment: A focus on social media text mining. *European Journal of Personality*, 34(5), 826–844. <https://doi.org/10.1002/per.2290>
- Tibshirani, R. (1997). The lasso method for variable selection in the cox model. *Statistics in Medicine*, 16(4), 385–395. [https://doi.org/10.1002/\(SICI\)1097-0258\(19970228\)16:4<385::AID-SIM380>3.0.CO;2-3](https://doi.org/10.1002/(SICI)1097-0258(19970228)16:4<385::AID-SIM380>3.0.CO;2-3)
- Vabalas, A., Gowen, E., Poliakoff, E., & Casson, A. J. (2019). Machine learning algorithm validation with a limited sample size. *PLoS One*, 14(11), e0224365. <https://doi.org/10.1371/journal.pone.0224365>
- Van de Ven, N., Bogaert, A., Serlie, A., Brandt, M. J., & Denissen, J. J. A. (2017). Personality perception based on LinkedIn profiles. *Journal of Managerial Psychology*, 32(6), 418–429. <https://doi.org/10.1108/JMP-07-2016-0220>
- Van Iddekinge, C. H., Lanivich, S. E., Roth, P. L., & Junco, E. (2016). Social media for selection? Validity and adverse impact potential of a facebook-based assessment. *Journal of Management*, 42(7), 1811–1835. <https://doi.org/10.1177/0149206313515524>
- Varma, S., & Simon, R. (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics*, 7(1), 91. <https://doi.org/10.1186/1471-2105-7-91>
- Vazire, S., Naumann, L. P., Rentfrow, P. J., & Gosling, S. D. (2008). Portrait of a narcissist: Manifestations of narcissism in physical appearance. *Journal of Research in Personality*, 42(6), 1439–1447. <https://doi.org/10.1016/j.jrp.2008.06.007>
- Watts, A. L., Lilienfeld, S. O., Smith, S. F., Miller, J. D., Campbell, W. K., Waldman, I. D., Rubenzer, S. J., & Faschingbauer, T. J. (2013). The double-edged sword of grandiose narcissism: Implications for successful and unsuccessful leadership among U.S. presidents. *Psychological Science*, 24(12), 2379–2389. <https://doi.org/10.1177/0956797613491970>
- Weed, J. (2021). Résumé-writing tips to help you get past the a.I. Gatekeepers. *The New York Times*. <https://www.nytimes.com/2021/03/19/business/resume-filter-artificial-intelligence.html>
- Wei, X., & Stillwell, D. (2017). How smart does your profile image look? Estimating intelligence from social network profile images. Proceedings of the Tenth ACM International Conference on Web Search and Data Mining, 33–40. <https://doi.org/10.1145/3018661.3018663>
- Weidmann, R., Chopik, W. J., Ackerman, R. A., Allroggen, M., Bianchi, E. C., Brecheen, C., Campbell, W. K., Gerlach, T. M., Geukes, K., Grijalva, E., Grossmann, I., Hopwood, C. J., Hutteman, R., Konrath, S., Küfner, A. C. P., Leckelt, M., Miller, J. D., Penke, L., Pincus, A. L., ... Back, M. D. (2023). Age and gender differences in narcissism: A comprehensive study across eight measures and over 250,000 participants. *Journal of Personality and Social Psychology*, 124(6), 1277–1298. <https://doi.org/10.1037/pspp0000463>
- Wilcox, R. R. (2011). *Introduction to robust estimation and hypothesis testing*. Academic Press.
- Wilhelm, O., Schroeders, U., & Schipolowski, S. (2014). *Berliner Test zur Erfassung fluiden und kristallinen Intelligenz für die 8. bis 10. Jahrgangsstufe*. Hogrefe. <https://fis.uni-bamberg.de/handle/uniba/20955>
- Wille, B., Hofmans, J., Lievens, F., Back, M. D., & De Fruyt, F. (2019). Climbing the corporate ladder and within-person changes in narcissism: Reciprocal relationships over two decades. *Journal of Vocational Behavior*, 115, 103341. <https://doi.org/10.1016/j.jvb.2019.103341>
- Yarkoni, T., & Westfall, J. (2017). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspectives on Psychological Science*, 12(6), 1100–1122. <https://doi.org/10.1177/1745691617693393>

- Youyou, W., Kosinski, M., & Stillwell, D. (2015). Computer-based personality judgments are more accurate than those made by humans. *Proceedings of the National Academy of Sciences of the United States of America*, 112(4), 1036–1040. <https://doi.org/10.1073/pnas.1418680112>
- Zhang, L., Van Iddekinge, C. H., Arnold, J. D., Roth, P. L., Lievens, F., Lanivich, S. E., & Jordan, S. L. (2020). What's on job seekers' social media sites? A content analysis and effects of structure on recruiter judgments and predictive validity. *Journal of Applied Psychology*, 105(12), 1530–1546. <https://doi.org/10.1037/apl0000490>
- Zibarras, L. D., & Woods, S. A. (2010). A survey of UK selection practices across different organization sizes and industry sectors. *Journal of Occupational and Organizational Psychology*, 83(2), 499–511. <https://doi.org/10.1348/096317909X425203>
- Ziegler, M., Danay, E., Heene, M., Asendorpf, J., & Bühner, M. (2012). Openness, fluid intelligence, and crystallized intelligence: Toward an integrative model. *Journal of Research in Personality*, 46(2), 173–183. <https://doi.org/10.1016/j.jrp.2012.01.002>
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2), 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix S1

How to cite this article: Härtel, T. M., Schuler, B. A., & Back, M. D. (2024). ‘LinkedIn, LinkedIn on the screen, who is the greatest and smartest ever seen?’: A machine learning approach using valid LinkedIn cues to predict narcissism and intelligence. *Journal of Occupational and Organizational Psychology*, 97, 1572–1602. <https://doi.org/10.1111/joop.12531>