Society of Consulting Psychology
Consulting Psychology Journal

© 2024 American Psychological Association
ISSN: 1065-9293
https://doi.org/10.1037/cpb0000289

SOME ETHICAL AND LEGAL ISSUES IN USING ARTIFICIAL INTELLIGENCE IN PERSONNEL SELECTION

Olaf Ringelband[1] and Christian Warneke[2]
[1] *Department of Psychology, University of Hamburg*
[2] *Department of Public Management, Hamburg University of Applied Sciences*

The present article identifies four types of data that can be analyzed using artificial intelligence (AI)-based assessments in various stages of the selection process: (a) data from traditional psychometric instruments, (b) data from AI-based assessment tools, (c) data from the "digital footprint" of given individuals, and (d) direct and indirect observations. In all methods, AI makes predictions and supports hiring/promotion decisions using historic data ("training data"). Understanding both the nature of these data and the underlying functionality of the AI is crucial when it comes to evaluating ethical and legal issues that might arise when employing AI in assessments, such as discrimination and inequities. Questions addressed here include the following: (1) Does a bias already exist in the data on which the AI system has been trained? (2) Does the algorithm have parameters based on psychological theories and evidence? (3) Does the AI system allow the "black box" to be opened? The article also notes that AI-driven assessments have the potential to deliver more precise evaluations of individuals. Additionally, because these assessments rely on a wealth of data and on numerous interactions drawn from real-life circumstances, they can address typical rather than merely maximal behavior, with the typical case capable of being measured via psychometric tests. The article furthermore discusses a number of ethical issues that arise in personnel selection using AI.

Theodore L. Hayes served as action editor.
Olaf Ringelband https://orcid.org/0009-0005-6783-6325
The authors have no known conflicts of interest to disclose.
Olaf Ringelband played a lead role in conceptualization, methodology, resources, supervision, writing–original draft, and writing–review and editing. Christian Warneke played a supporting role in writing–original draft and writing–review and editing.
Correspondence concerning this article should be addressed to Olaf Ringelband, Department of Psychology, University of Hamburg, Germany. Email: ringelband@management-diagnostik.de

What's It Mean? Implications for Consulting Psychology
Consulting psychologists need to understand the value and limitations of using artificial intelligence in personnel selection. They need to consider both the potential benefits and the ethical and legal concerns in such applications.

Keywords: artificial intelligence, assessment, big data, ethical tensions, personnel selection

The potential of generative artificial intelligence (AI) became clear throughout the world with the publication of ChatGPT 3.0 in 2023. Indeed, the rapid increase in the number of users and the impressive range of possibilities in the use of large language models (LLMs) have quickly demonstrated the disruptive effect of the new technology in virtually all industries and fields of work. However, in addition to the potential of this technology, it also goes hand in hand with serious risks of misuse.

In the present article, we aim to translate research into practice by highlighting the points of contact between assessment and AI and by focusing on the data available for assessment AI tools. As the American Psychological Association's (APA) *Guidelines for Psychological Assessment and Evaluation* (2020) state, "Inasmuch as computer technology, test instrument usage, and new instrument design are constantly evolving, the responsibilities and challenges to the psychologist practitioner using these modalities are likewise substantial." (p. 22).

The term "artificial intelligence" is not very consistently defined and is commonly used to cover various different technologies. In fact, the Stanford AI report/Stone et al. (2016) noted that the research field may have become so large, dynamic, and complex specifically due to the lack of a precise, generally accepted definition of AI. In this context, the definition formulated by Nilsson (2010) is frequently adopted: "Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment" (p. xiii). In contrast, Klingbeil-Döring (2023) suggested that AI is:

> More of a computer science concept, or more precisely the information technology project of enabling machines to use intelligence and learn when performing tasks. It is not precisely defined what intelligence means here. The various definitions essentially emphasize the ability to solve problems and achieve certain goals on the basis of knowledge and experience. (n.p.)

In order to foster a basic understanding of the concept of AI, we first define a few terms as used in the present article:

1. An "algorithm" can be quickly explained as "… a clear set of instructions for solving a problem. An input is converted into an output in precisely defined steps" (Berberich, 2019, S. 11).

2. "Machine learning" (ML) is a key aspect of AI. ML algorithms analyze data, recognize patterns and relationships using statistical techniques, and learn and improve automatically based on their own experience; thus, no explicit programming of individual rules is necessary.

3. "Large language model" is a type of AI trained on vast amounts of text data to understand and generate humanlike language. It uses deep learning techniques, particularly neural networks, to process and predict language patterns, enabling it to perform tasks such as text generation, translation, and question answering.

4. "Generative AI" is used to describe a subarea of AI that is concerned with the generation of new content, which can take the form of text, images, music, and so forth. The underlying principle is the learning of patterns from large volumes of data such that new data similar to the training data can be generated.

AI in the Context of Assessment and Selection

Currently, AI can clearly be seen to have opened up technical options in assessment that were previously not possible. Indeed, despite some concerns and limitations, AI is already being widely used in the context of assessment and selection. This includes but is not limited to the following:

1. Scoring: AI-based scoring has been shown to be equivalent in quality to that of human experts (Koenig et al., 2023; Thompson et al., 2023).

2. Creating content for assessment tools: LLMs are well-suited for creating items for questionnaires, scenarios for situational judgment tests, behavioral anchored rating scales, and structured interview questions (Götz et al., 2024).

3. Scoring applicants' biodata: AI tools are able to extract relevant information from CVs and to use biographical data to predict work outcomes (Sajjadiani et al., 2019).

4. Job analysis: In order to cluster similar job families, LLMs can be utilized to extract the relevant KSAOs from job descriptions (Lewis et al., 2023).

5. Identifying patterns in large data sets: LLMs can help in identifying patterns in large numbers of text files, such as interview transcripts.

6. Assessing individuals based on behavioral big data: AI tools can be utilized to evaluate employees' skills, personalities, competencies, social interactions, and decision-making abilities as well as to identify areas for development, potential career paths, and even person–job mismatches.

7. Generating working hypotheses on individuals: These hypotheses are based on the input of test data, such as personality, general mental ability, interview transcripts, and transcripts from simulation exercises, such as roleplays and case studies.

8. Generating narrative feedback reports: LLMs can support psychologists in writing individual reports based on the results of an assessment. Test providers use LLMs to generate individualized narrative reports based on psychometric test results (Koenig et al., 2023).

Potential Advantages of Using AI for Personnel Assessment

The potential benefits of using AI for personnel assessments include efficiency, speed, lower costs, the reduction of routine and administrative tasks, and an improvement in selection quality (Will et al., 2023). The use of AI to support personnel selection decisions dates back to the 1980s, when AI in the form of expert systems was utilized (Hooper et al., 1990). These expert systems translated the implicit and explicit knowledge of human resource (HR) experts into if–then rules (e.g., "If the candidate has relevant job experience, increase the job-fit score"). However, this rule-based approach ultimately failed because it did not reflect the complexity of human decision making, including hiring decisions (Davis & King, 1984, p. 29).

The complexity of judgments in assessment is illustrated by Schuler's (2008) trimodal concept of personnel assessment, which combines the three functional logics of (a) "explaining career success based on relatively stable traits," (b) "demonstrating success-relevant career-related behavior in the assessment," and (c) "predicting future behavior on the basis of past behavior" (pp. 129–130). Moreover, Schuler (2008) emphasized that high-quality decisions are made by combining these data sources. In order to utilize the possibilities of AI in assessment and to achieve the highest possible quality, it is important to look closely at how assessment decisions are made.

Use of Assessments in Personnel Selection Decision Making

"Assessment-based decisions" refer to a wide range of possibilities in the analysis of individuals' competencies, abilities, personality, and skills. These range from the consideration of individual performance scores as part of an exercise in an assessment center to decisions on hiring, promotion, or rejection. The basic premise is initially simple: The same rules apply whether decisions are made by an amateur, by an experienced consulting or industrial–organizational (CIO) psychologist, or by an AI tool. Every decision in the assessment is based on assessment data, regardless as to whether how these data have been collected. It is therefore worth taking a closer look at these assessment data.

Typically, a variety of types of data are used in selection activities. These include numerical data (test scores, structured observation sheets), written data (e.g., in a CV), and verbalized data (e.g., interviews). It is also possible to evaluate video/audio recordings (e.g., roleplays, interviews, presentations) and data from AI-based assessment tools (e.g., AI-supported gamification tools; see Pérez et al., 2023, for the use of AI in serious games) and chatbots (Chen, 2023; for possibilities of text analysis in organizational behavior, see Short et al., 2018) as well as data from the "digital footprint" of individuals on the internet or from social networks, with such data typically being digitally written. For example, the University of Cambridge's Psychometrics Centre offers an online application that estimates personality data and biographical data from social media activities. All of these data formats can be processed using ML methods, which in many cases are closely related to classical statistical methods, such as regression analysis. In contrast, modern generative AI tools can use a large variety of data (e.g., text, speech, visual) in parallel since these tools are "multimodal" by design. However, their underlying transformer architecture makes the data even less transparent than in ML approaches. Indeed, predictions and hiring/promotion decisions are made based on such data. Future AI applications in the field of assessments will also be based on such assessment data because these data will be used to train the AI applications.

How AI Tools Work

Understanding both the nature of assessment data and the algorithms underpinning the AI is crucial when it comes to assessing ethical and legal issues, such as discrimination and unfairness, which could arise when using AI in assessments. In most cases, AI systems based on ML algorithms permit an examination of the system's internal functioning. For example, it is possible to examine the parameters that influence an outcome, such as a hiring recommendation. However, AI systems based on deep learning algorithms (e.g., generative AIs) are characterized by a lack of transparency. In other words, the algorithms can be seen as a "black box," which makes it challenging to comprehend their internal workings, even for their developers. The only observable output is based on complex input. When employing AI for personnel selection, it is imperative to possess a fundamental understanding of the system's internal workings in order to ascertain the feasibility of "opening the black box" and elucidating the methodology underlying the system's recommendation.

Despite the terms "intelligence" and "learning," it is important to bear in mind that these methods are essentially mathematical and computer science methods of pattern recognition. In ML, a model is first created based on training data using a learning algorithm in the so-called "training phase." The model can then be applied to a problem in a second step, which is known as the "test phase" or "application phase." From a technical point of view, the more relevant data that are used, the better. Technically speaking, all data that are generated during the entire personnel selection process can be processed. This initially includes classic assessment data, such as test results, interview transcripts, behavior, and evaluations in assessment centers as well as biographical data from application documents.

However, a fully digitalized application process (for an overview, see Benhmama & Bennyni, 2023) generates far more data, which can include the following: (a) video and audio recordings of digital assessments, (b) extensive individual data of a technical nature (e.g., from communication

devices or operating systems) or of a socioeconomic nature (e.g., address data, educational background, income), (c) data on digital interactions (e.g., data on user behavior on the company's career website, dwell time, click behavior, and interaction patterns for predicting user competence and performance via mouse interaction parameters; Attig et al., 2019), and (d) communication data (e.g., speed, length, language style).

In addition, publicly available data on individuals from their digital footprint on the internet or on social networks can also be used to assess candidates. Such data may include (a) personal information (e.g., name, age, gender, date of birth, place of residence), (b) education/occupation details (e.g., school education, higher education, current/previous jobs), (c) interests/activities (e.g., likes, group memberships, event participation, hobbies), (d) communication patterns (e.g., manner of communication, language style, frequency of posts/interactions), (e) network structure (e.g., information about friends and followers/connections, including the type and strength of these connections), (f) media content (e.g., photos, videos, and other media shared by the person), (g) public posts/comments (e.g., the content of the posts as well as comments and their context), and (h) geographical data (e.g., location information from posts/check-ins).

Some Likely Developments in the Near Future

A wide range of behavioral data will likely soon be used (see Chakrabarti et al., 2022) to assess individuals. These data will encompass not only written and spoken behavior (i.e., language) but also spatial behavior (i.e., physical movement). By using "wearables" (e.g., smart watches/virtual reality goggles), it will become possible to "tokenize" (i.e., to quantify and record) every aspect of an individual's behavior, thereby enabling future behaviors to be predicted based on a vast number of existing data. The predictive models based on physical movement may be designated as large spatial models and may be considered an extension of existing LLMs. As Oswald et al. (2020) stated:

> Clearly, gathering big data in the attempt to capture the multidimensionality and dynamics of organizations, teams, and employees is both a blessing and in some ways a curse, but we firmly believe that the expertise of IOP and HRM scholars will greatly increase the likelihood of blessings. The future of big data in organizations likely holds many exciting possibilities, at least some of which no algorithms or humans will predict. (p. 529)

AI in Personnel Selection: Ethical and Legal Tensions

As training data are of great importance, ethical tensions can easily arise, for that which is technically possible is not necessarily legally permissible, and that which is legally permissible can nevertheless be ethically and morally problematic. One example includes "parsing" (i.e., machine processing) career data from job platforms such as LinkedIn, which can help LLMs to "understand" exactly what constitutes successful careers but violates the job platform's terms. It is safe to assume that although illegal in the EU due to the General Data Protection Regulation, this process is nevertheless commonly undertaken. Another example of ethical tensions includes companies using their own data on employees' careers (e.g., for making selection decisions for future hires), which is ethically questionable since using existing data for future hires perpetuates existing biases.

In order for an AI system to be able to support hiring and promotion decisions, it must be trained with historical data. The data must be "labeled," that is, each data record representing an employee or candidate must be provided with a label indicating whether the employee is considered "successful" or not. This assessment can be made on the basis of quantitative data (e.g., sales figures or supervisor assessments) or past decisions (hiring or promotion of a candidate). Ethical concerns emerge in the process of selecting appropriate training data sets and in the labeling of individual data sets. Consequently, the requirements of AI models—particularly regarding the necessity of acquiring as many data as possible—frequently conflict with aspects of individual data protection, fairness, and antidiscrimination as well as with transparency requirements (see Table 1).

In choosing or labeling training data, all potential criteria as well as each single criterion refer to historical data and should therefore always be considered biased or outdated. As a basic

Table 1
Ethical Tensions in Selecting and Labeling Data Sets

	Ethical problem	Description of problem	Ethical field of tension
Selecting data sets	Inefficiency	The question is whether to use efficient, modern, data-driven models that may be biased or to prioritize fairness and nondiscrimination by using traditional methods.	The use of efficient data-driven models versus the need for fairness and nondiscrimination.
	Data privacy	Although the use of personal data for training AI models makes it possible to develop better models, there is a risk of breaching privacy.	The benefits of data use versus the right to informational self-determination.
	Responsibility/ liability	Determining responsibility for the incorrect or unfair decisions of an AI model is complex when decisions are partially or fully automated.	The distribution of responsibility between man and machine versus AI developers.
Labeling data sets	Subjectivity/ bias	Labeling data often requires human judgment. If the individuals who label the data have (un)conscious biases, these can be incorporated into the labeled data.	The need for the utilization of human knowledge versus the risk of transferring biases into the AI system.
	Definition of success	Determining what is considered "successful" or "suitable" in the training data can be subjective and may depend on the perspective of the decision makers.	The development of objective criteria for success that are not discriminatory.
	Problem	Description of problem	Ethical field of tension
Labeling data sets	Data privacy/ anonymity	Labeling data are often not completely anonymizable since specific information about individuals is required to assign relevant labels.	The privacy and protection of personal data versus the need for the largest number of data.
	Transparency/ explainability	Applicants and employees should know how and why certain data about them have been labeled, especially if these data affect the individuals' chances of getting a job.	Transparency in the labeling process.
	Opt-out	Applicants and employees should be able to object to the use of their data for training purposes, though opting out is detrimental to the number and balance of the data.	The possibility to object versus the demand for data quantity and data balance.

Note. AI = artificial intelligence.

example, the present authors analyzed a large database of psychometric, sociographic, and biographical data from more than 2,000 top German managers who had undergone individual psychological assessments for selection purposes by feeding the data into an LLM (in this case, Mistral7B). The LLM was asked to determine what constitutes a successful manager, and the system identified the following characteristics of successful candidates who had been hired: (a) a master's degree or PhD in economics or engineering, (b) age between 45 and 55, (c) male gender, (d) above-average cognitive skills (indicated by test results), (e) extraverted personality, and (f) high motivation to perform well. If this system were used to prescreen candidates, it would identify potentially successful candidates but would also exclude certain groups due to its reliance on existing data, which are the most likely data to contain biases (e.g., male candidates would be favored). It is worth noting that similar results and similar biases would also be found in a classic ML approach.

The Importance of Applicant Reactions and Fairness

As Schuler stated in 2006, the benefits of assessment procedures and processes are:

> No longer measured solely in economic terms, but also in terms of the applicants' reaction to selection situations and fairness towards applicant groups that may be disadvantaged by the use of certain types of procedures… . (p. 728)

However, even human and AI-free personnel selection is known to be replete with discrimination. For example, systematic errors are present in the assessment of individuals with regard to physical attractiveness and characteristics such as facial acne and body type as well as regarding an applicant's language or dialect (Kanning, 2017; Shields et al., 2024). A meta-analysis of field experiments on hiring discrimination against African Americans and Latinos by Quillian et al. (2017) additionally found no significant improvements in terms of racial discrimination.

AI-based discrimination is additionally known both within and outside of personnel selection. AI systems often contain biases and are in danger of reproducing these biases continuously, as has been noted by Schwartz et al. (2022). Mehrabi (2022, pp. 4–9) also included an overview of 19 different types of biases in AI systems and provided real-life examples of biased systems, such as an AI-based job advertisement system that was intended to promote STEM jobs in a gender-neutral way but that rarely presented the advertisements to women because women were a more expensive subgroup for targeting job ads (pp. 2–3). Furthermore, P. F. Langer and Weyerer (2020) pointed out the dangers of the AI-based reproduction of prejudices and discriminatory behavior, discriminatory effects based on programming, and biases in design and various data sets. They conclude:

> In general, this means that AI applications may not only violate national anti-discrimination laws enforced by a large number of countries, such as the USA … , the UK … or Germany … , but also the prohibition of discrimination enshrined in Article 2 of the Universal Declaration of Human Rights. (p. 221)

This conclusion underlines the responsibility of the assessor, who needs to identify and address biases in AI systems just as they would do with non-AI-based selection systems. Avgustin (2022) described several statistical methods for evaluating an AI system's discriminatory bias, including feeding algorithms with the same data set twice—once with information on gender, and once without this information—and then comparing the system's results. Similarly, Verma et al. (2021) proposed a novel black box methodology for enhancing fairness in ML models by identifying and removing biased training data. This approach significantly reduced individual discrimination (often to 0%) and improved both the accuracy and fairness of the models compared with models trained on a full data set.

Implications for Practice

According to Ihsan and Furnham (2018):

> Emerging technologies open a whole new world for organizations and will change the way people recruit … employees as we know it. Consulting psychologists in this enterprise need to constantly update themselves on what is available as well as [on] the evidence for reliability and validity. (p. 161)

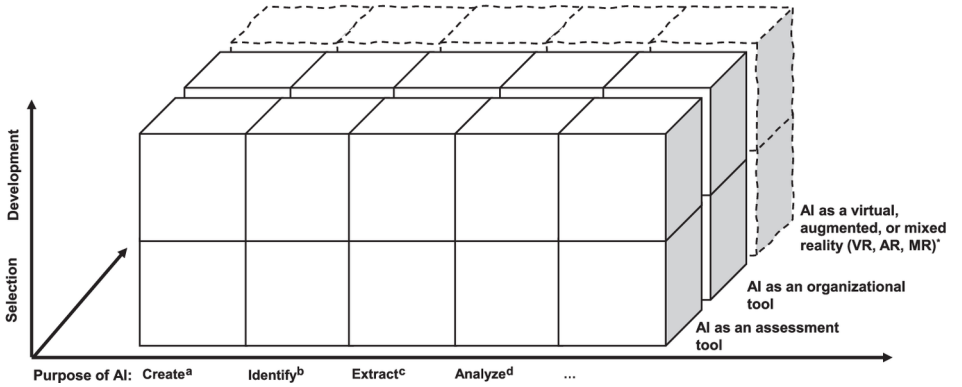
The authors of the present article make several recommendations regarding the use of AI in the field of assessment, as outlined below.

Acting Purposefully

When using or introducing AI tools in the context of assessment, many different professional, organizational, technical, legal, and/or ethical questions can arise in practice due to the complexity of the topic of AI. These questions need to be discussed with very different stakeholders, such as managers, HR employees, work councils, data protection officers, equal opportunity officers, and so forth, who in turn often have highly heterogeneous levels of knowledge about AI. In order to reduce the complexity of the field of action, the authors of the present article propose an orientation model (see Figure 1). Depending on (a) whether AI tools are to be used for assessment or administrative/organizational

Figure 1

AI in the Field of Selection and Development: An Orientation



[a] The possibilities offered by artificial intelligence (AI) in the areas of virtual reality (VR), augmented reality (AR), and mixed reality (MR) are not yet widely utilized in selection and assessment. [b] Items for questionnaires, case studies, and so forth. [c] Competencies for jobs, candidates with a certain skill set, and so forth. [d] Core competencies from job descriptions, patterns in large data, and so forth. [e] Behavioral data in BARS, interview answers, and so forth.

purposes, (b) whether the focus should be more on selection or development, and (c) what specific purpose the AI is to be used for, there should be differences with regard to (a) legal framework conditions and legal issues (e.g., a workers' council), (b) ethical tensions and challenges, (c) ethical opportunities/possibilities, (d) the objectives of the use of AI, (e) costs/budget, (f) the training needs of the users, (g) concrete AI tools/providers, and (h) the need for training data.

The outlined dimensions of selection and development lead to different target groups (usually—but not only—internal employees or external candidates to an organization). The dimension of use as an assessment or organizational tool also has a significant impact in terms of legal and ethical issues. Defining the intended purpose of the AI helps with focusing. The model does not aim to name all the abovementioned differences between the cubes or even to offer individual solutions, for example, because legal questions alone might be answered very differently in different countries. Instead, the model is intended to be used as a form of "help for self-help." CIO psychologists can use it to identify relevant questions for a specific use of AI more quickly and accurately together with the group of stakeholders or on their own. If the presented model is made available to all participants, it should help to focus, structure, and accelerate such discussions.

Candidates' Consent

Mirowska (2020) and Lavanchy et al. (2023) reported that candidates prefer human evaluators over AI when it comes to making evaluations, whereas Li et al. (2021, p. 175) came to the more nuanced conclusion that "when considering the potential to adopt AI in hiring, the relative drawbacks and benefits of these AI systems differ greatly by the type of decisions and the type of candidate roles these systems are built to assist." Moreover, the use of AI in selection processes raises ethical questions about transparency for organizations and CIO psychologists. In this regard, it is important to distinguish between intentional and unintentional data collection. When answering items on a personality questionnaire, a candidate is assumed to have consented to having their answers evaluated by any technology, including an AI algorithm. However, online employee behavior—such as mouse clicks and the length of time spent interacting with others—is considered a form of unintentional data and requires in some countries as Germany (German Federal Ministry of Justice & the Federal Office of Justice, 2024, Section 87) employee consent in order to be evaluated. This situation is in line with ethical standards set by the APA (American Psychological Association, 2017, Section 9.03), even when the data are used for prosocial actions, such as offering individual development

suggestions. We suggest that candidates should always be informed when AI systems are involved in evaluating behavioral data. In particular, candidates should be informed about (a) what data are analyzed, (b) what results the AI tool produces, and (c) in what form these results are incorporated into a final decision (which, we contend, should always be done by a human). The candidates' consent should then be based on this information.

As outlined by Tippins et al. (2021), a professional and legal standard should be defined in order to inform individuals about when they are interacting with an AI app instead of with a human. It is important to note that LLMs can now easily pass the Turing Test (Mei et al., 2024), which means that participants in selection may be unable to distinguish between an AI app and a human, even after a long conversation. Therefore, it is crucial to establish clear guidelines for AI–human interactions.

Perceiving Environmental Changes

Changes in the work environment can have an impact on the usefulness and quality of data used in AI-based selection. Since AI systems are usually trained on "historical" data, that limits the scope of the systems to the past even when there has been change that would affect the skills and knowledge needed for job success. While humans usually only notice changes in the environment (e.g., when employees work mostly from home because of an epidemic), AI systems are not aware of those changes. Moreover, there is no consensus in the literature as to how often AI systems should be retrained with new data and it can be challenging to do this. As Zhang et al. (2020) noted, "Retraining recommender systems on both historical and new data is desirable for high model fidelity, but full model retraining can be time-consuming and memory-costly" (p. 1479). For instance, when ChatGPT Version 3.5 was released to the public in December 2022, OpenAI disclosed that the tool had been trained on internet data up until September 2021 (Aarthy, 2024). When training an AI system for assessment and selection, the system relies on existing data. AI models that are used for personnel selection may similarly not always be up to date due to the high costs and data requirements of retraining.

This issue is not unique to AI because the predictive power of certain psychological dimensions can change with shifts in the workplace environment, such as with the movement to remote work that occurred during the pandemic. Although CIO psychologists should be aware of changes in the work environment and are refining their assumptions about the validity of their predictors, an AI system can only be retrained when there is a sufficient number of data to resample. Practitioners should therefore be aware of fundamental changes in the working environment, which can be sudden or gradual. They should also be aware of the impact such changes can have on the data that are used to train AI tools. Of course, psychologists also need to be aware that test items and stimuli may need to be replaced or updated.

Caution When Making Fully Automated Decisions

AI is increasingly often taking over tasks in the recruitment and selection process, with the final decision typically being made by a human recruiter or manager (Fernández-Martínez & Fernández, 2020; Yarger et al., 2020). Some organizations have even turned to fully automated decision making due to the pressure to make the recruitment process as fast and cost-effective as possible (Lee, 2018; Vasconcelos et al., 2018). However, it is important to exercise caution in such circumstances. Without knowing what is inside the "black box" of AI decision-making tools and without understanding how an AI tool or an LLM reaches its conclusions—particularly when using generative AI—it may be difficult for organizations to defend the decisions being made.

Second, it is also necessary to verify whether the system's conclusions align with existing psychological models and empirical evidence. LLMs function similarly to a scientist analyzing a vast data set. Consequently, it is inevitable that a multitude of statistically significant correlations will emerge from such data sets, many of which will be the result of random chance. Cecil et al. (2023) noted that individuals (i.e., decision makers) tend to follow the advice of an AI system on personnel selection, even if the advice is incorrect. This process may result in worse decision making than if no predictive information were employed. As Kupfer et al. (2023) cautioned, simply making decision makers aware of their responsibility does not improve their decision quality. However, the way the data

are presented can have an impact. According to M. Langer et al. (2021), AI support for personnel selection should be provided after decision makers have evaluated the candidates themselves. Ultimately, as Hmoud and Laszlo (2019) argued, while AI will increasingly often impact HR recruitment and selection in the future, it is unlikely to completely replace human recruiters, which highlights the importance of thoughtfully integrating AI in order to support rather than replace human decision making in HR processes.

Utilizing Established Fairness Standards

Biases presented a challenge when it comes to assessing decision-making processes long before the technical possibilities of AI found their way into personnel selection. Established fairness benchmarks can therefore also provide a point of orientation when using AI. Görlich and Schuler (2014) distinguished between (a) statistical fairness (i.e., the consideration of the quality of the instruments in terms of objectivity, reliability, validity, and standardization), (b) biographical fairness (i.e., the exclusion of biographical criteria, e.g., parents' occupation, for which the applicant is not responsible), (c) requirement-related fairness (i.e., only relevant requirements should be used), (d) procedural fairness (i.e., transparency and fairness in the selection process), and (e) social fairness (i.e., fair selection and success criteria, e.g., performance instead of contacts; p. 1181). All of these fairness categorizations are influenced by the use of AI in selection and assessment.

The APA's *Ethical Principles of Psychologists and Code of Conduct* (2017) and the SIOP Principles for the Validation and Use of Personnel Selection Procedures (Society for Industrial Organizational Psychology, 2018) provide a basic and overarching framework of fairness and ethical behavior. The works include guidelines for ethical behavior in psychology and are helpful when dealing with ethical dilemmas. For instance, if we take the second example from Table 1 on "Data Privacy," which addresses the dilemma that exists between using as many (personal) data as possible to make AI tools better on the one hand and the right to informational self-determination on the other hand, the APA ethical standards give advice. For instance, Standard 4.07 stipulates that the use of confidential information for other purposes (e.g., as training data) is only possible with the written consent of the individuals concerned (Standard 4.07). Moreover, Standard 6.02b recommends anonymizing or coding data for this purpose (Standard 6.02). The SIOP Principles for the Validation and Use of Personnel Selection Procedures (Society for Industrial Organizational Psychology, 2018) underline the notion that:

> Confidentiality is an ethical responsibility of the testing professional… . The testing professional should carefully decide what level of anonymity or confidentiality is required by relevant privacy laws and can be established, communicate it to participants, and maintain it thereafter. (p. 30)

Therefore, the use or development of even data-based AI tools for assessment purposes should only be possible for psychologists if they fulfill certain conditions, which include the requirement for these psychologists to constantly endeavor to develop and maintain their competence when using such assessment tools (Standard 2.03). APA Standard 2.01e is also relevant to the emerging field of AI and states that:

> In those emerging areas in which generally recognized standards for preparatory training do not yet exist, psychologists nevertheless take reasonable steps to ensure the competence of their work and to protect clients … and organizational clients, and others from harm Standard (p. 5).

Utilizing Established Models in the Area of Selection and Assessment

The technological disruption caused by AI technology—which is also changing recruitment and assessment strategies—is leading to revolutionary innovations and interesting opportunities. However, it still remains worthwhile to use proven psychological theories and models. Schuler's (2008) trimodal concept of personnel assessment can also be applied to new selection developments. For example, the automated processing of digital footprint data falls into the classification area of "biography," and AI-supported gamification tools can potentially be used—among other things—to analyze a candidate's behavior. It should be noted, however, that the selection of predictors must be carried out

with extreme caution and that predictive validity must first be demonstrated. The above examples are intended not to be a recommendation of predictors but to illustrate that known theoretical models can also be helpful in classifying new approaches.

Skepticism

Previous findings indicate that applicants and recruiters are both enthusiastic and skeptical about the possibilities of AI in personnel selection (Lee, 2018; Will et al., 2023), as is evident in the usefulness and risks of such selection procedures (Suen et al., 2019). Initial indications reveal that applicants' perception of fairness toward AI tools in personnel selection can improve with approaches that provide information and transparency regarding how the AI tools work and how the use of AI is beneficial to the applicant (Köchling & Wehner, 2023).

Active Role of Consulting and CIO Psychologists

AI systems are likely to take over many tasks that have traditionally been performed by consulting and CIO psychologists. Therefore, psychologists will have to be responsible for developing and implementing policies regarding the use of AI systems. CIO and consulting psychologists are experts in assessing and selecting individuals, but they should also have a better understanding of the nature of the algorithms underlying AI. Psychologists' training usually provides them with extensive knowledge of statistical methods and of the physiological foundations of the brain (i.e., neural networks) and should thus give them a solid basis for understanding the essential foundations of AI. CIO psychologists should familiarize themselves with concepts such as neural networks, predictive analytics, and stochastic methods and should be open to cooperating with computer engineers in designing and auditing AI systems.

Well-trained CIO psychologists may be able to take an active role in shaping governmental and legal regulations regarding AI in selection and assessment. They may assist by serving as advocates for the establishment of norms for the development, auditing, and application of AI systems in the hiring, promotion, and assessment process.

It is critical to establish clear regulations for preventing such outcomes. On the organizational level, it is important to implement standards for the ethical use of AI (see Norton, in press). These may include complying with existing privacy laws and providing transparency to applicants and managers regarding the assessment process and the role of AI. For example, at the beginning of the recruitment process, it could transparently be explained exactly which applicant data will be analyzed and how as well as exactly which AI technology will be used. In the European Union, data may also only be used for previously transparently defined purposes. It is especially important that there be clear rules related to human oversight, which means that decisions may ultimately need to be made by humans rather than by algorithms working independently. Psychologists can also advocate for a diverse range of voices to be included in the establishment of AI standards within organizations.

Conclusion

The remarkable pace at which AI is growing in sophistication has already had a significant impact on the process of personnel assessment and selection. However, as the saying goes, "with great power comes great responsibility." This is particularly true when it comes to decisions that have an impact on human careers and lives, not to mention organizations themselves. It would be naïve to assume that appealing to organizations' and individuals' sense of ethics is sufficient to prevent the misuse—be it deliberate or inadvertent—of AI when assessing individuals. A robust regulatory framework that encompasses companies, industry organizations, and nationwide legislation will be essential to preventing the unethical and unprofessional application of AI.

On a positive note, thanks to their studies and the ethical guidelines of existing professional associations, CIO psychologists have a good foundation for improving the ethics and validity of personnel assessment using the new possibilities of AI. In the past, behavioral data were collected

within a specific assessment environment (e.g., within an AC); however, the use of big data (or even spatial data) now enables psychologists to base their assessments on individuals' everyday behavior. This will open up new directions for future research, specifically to increase the predictive validity of psychological assessments by using this kind of big data from individuals or groups/interactions or organizations. And this will give rise to new legal and ethical questions in different countries and cultures when dealing with AI in personnel selection, which will probably be answered differently internationally but will certainly require the expertise of AI-interested CIO psychologists.

References

Aarthy, R. (2024, May 22). A detailed comparison between ChatGPT versions. *Medium*. https://aarthy-botcampusai.medium.com/a-detailed-comparison-between-chatgpt-versions-aa9dce022f51

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct (2002, amended effective June 1, 2010, and January 1, 2017)*. https://www.APA.org/ethics/code/

American Psychological Association. (2020). *Guidelines for psychological assessment and evaluation*. https://www.APA.org/about/policy/guidelines-psychological-assessment-evaluation.pdf

Attig, C., Then, E., & Krems, J. F. (2019). Show me how you click, and I'll tell you what you can: Predicting user competence and performance by mouse interaction parameters. In J. Kacprzyk (Ed.), *Advances in intelligent systems and computing* (pp. 801–806). Springer. https://doi.org/10.1007/978-3-030-11051-2_122

Avgustin, V. (2022). *Evaluating discrimination bias in AI decision making systems for personnel selection* [Master's thesis, National College of Ireland]. https://norma.ncirl.ie/6074/1/viktoravgustin.pdf

Berberich, N. (2019). Algorithmen. In K. Kersting, C. Lampert, & C. Rothkopf (Eds.), *Wie Maschinen lernen* (pp. 11–20). Springer. https://doi.org/10.1007/978-3-658-26763-6_2

Cecil, J., Lermer, E., Hudecek, M., Sauer, J., & Gaube, S. (2023, March 20). *The effect of AI-generated advice on decision-making in personnel selection*. https://doi.org/10.31219/osf.io/349xe

Chakrabarti, S., Biswas, N., Jones, L. D., Kesari, S., & Ashili, S. (2022). Smart consumer wearables as digital diagnostic tools: A review. *Diagnostics*, *12*(9), Article 2110. https://doi.org/10.3390/diagnostics12092110

Chen, Z. (2023). Collaboration among recruiters and artificial intelligence: Removing human prejudices in employment. *Cognition Technology and Work*, *25*(1), 135–149. https://doi.org/10.1007/s10111-022-00716-0

Davis, R., & King, J. J. (1984). The origin of rule-based systems in AI. In B. G. Buchanan & E. Shortliffe (Eds.), *Rule-based expert systems: The MYCIN experiments of the Stanford heuristic programming project* (pp. 20–52). Addison-Wesley.

Fernández-Martínez, C., & Fernández, A. (2020). AI and recruiting software: Ethical and legal implications. *Paladyn: Journal of Behavioral Robotics*, *11*(1), 199–216. https://doi.org/10.1515/pjbr-2020-0030

German Federal Ministry of Justice and the Federal Office of Justice. (2024). *Works constitution act (Betriebsverfassungsgesetz—BetrVG)*. https://www.gesetze-im-internet.de/englisch_betrvg/englisch_betrvg.pdf

Görlich, Y., & Schuler, H. (2014). Personalentscheidungen, nutzen und fairness. In H. Schuler & U. P. Kanning (Eds.), *Lehrbuch der personalpsychologie* (pp. 1137–1199). Hogrefe.

Götz, F. M., Maertens, R., Loomba, S., & van der Linden, S. (2024). Let the algorithm speak: How to use neural networks for automatic item generation in psychological scale development. *Psychological Methods*, *29*(3), 494–518. https://doi.org/10.1037/met0000540

Hmoud, B., & Laszlo, V. (2019). Will artificial intelligence take over human resources recruitment and selection? *Network Intelligence Studies*, *7*(13), 21–30.

Hooper, R. S., Galvin, T. P., Kilmer, R. A., & Liebowitz, J. (1998). Use of an expert system in a personnel selection process. *Expert Systems With Applications*, *14*(4), 425–432.

Ihsan, Z., & Furnham, A. (2018). The new technologies in personality assessment: A review. *Consulting Psychology Journal*, *70*(2), 147–166. https://doi.org/10.1037/cpb0000106

Kanning, U. P. (2017). Fairness und Akzeptanz von Personalauswahlmethoden. In D. Krause (Ed.), *Personalauswahl* (pp. 271–299). Springer. https://doi.org/10.1007/978-3-658-14567-5_12

Klingbeil-Döring, W. (2023). Die Auswirkungen von Künstlicher Intelligenz auf den Arbeitsmarkt. In *Dossier Arbeitsmarktpolitik der Bundeszentrale für politische Bildung*. bpb:Medienzentrum Berlin.

Köchling, A., & Wehner, M. C. (2023). Better explaining the benefits why AI? Analyzing the impact of explaining the benefits of AI-supported selection on applicant responses. *International Journal of Selection and Assessment*, *31*(1), 45–62. https://doi.org/10.1111/ijsa.12412

Koenig, N., Tonidandel, S., Thompson, I., Albritton, B., Koohifar, F., Yankov, G., Speer, A., Hardy, J. H., III, Gibson, C., Frost, C., Liu, M., McNeney, D., Capman, J., Lowery, S., Kitching, M., Nimbkar, A., Boyce, A., Sun, T., Guo, F. … Newtown, C. (2023). Improving measurement and prediction in personnel selection through the application of machine learning. *Personnel Psychology*, *76*(4), 1061–1123. https://doi.org/10.1111/peps.12608

Kupfer, C., Prassl, R., Fleiß, J., Malin, C., Thalmann, S., & Kubicek, B. (2023). Check the box! How to deal with automation bias in AI-based personnel selection. *Frontiers in Psychology*, *14*, Article 1118723. https://doi.org/10.3389/fpsyg.2023.1118723

Langer, M., König, C. J., & Busch, V. (2021). Changing the means of managerial work: Effects of automated decision support systems on personnel selection tasks. *Journal of Business and Psychology*, *36*(5), 751–769. https://doi.org/10.1007/s10869-020-09711-6

Langer, P. F., & Weyerer, J. C. (2020). Diskriminierungen und Verzerrungen durch Künstliche Intelligenz. Entstehung und Wirkung im gesellschaftlichen Kontext. In M. Oswald & I. Borucki (Eds.), *Demokratietheorie im Zeitalter der Frühdigitalisierung* (pp. 219–240). Springer. https://doi.org/10.1007/978-3-658-30997-8_11

Lavanchy, M., Reichert, P., Narayanan, J., & Savani, K. (2023). Applicants' fairness perceptions of algorithm-driven hiring procedures. *Journal of Business Ethics*, *188*(1), 125–150. https://doi.org/10.1007/s10551-022-05320-w

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, *5*(1). https://doi.org/10.1177/2053951718756684

Lewis, J. A., Morris, J., & Lewis, P. (2023). *Developing career word clouds for the O*NET system*. https://www.onetcenter.org/dl_files/WordClouds.pdf

Li, L., Lassiter, T., Oh, J., & Lee, M. K. (2021, July). Algorithmic hiring in practice: Recruiter and HR Professional's perspectives on AI use in hiring. *Proceedings of the 2021 AAAI/ACM conference on AI, ethics, and society* (pp. 166–176).

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., & Galstyan, A. (2019). *A survey on bias and fairness in machine learning*. arXiv. https://doi.org/10.48550/arXiv.1908.09635

Mei, Q., Xie, Y., Yuan, W., & Jackson, M. O. (2024). A Turing test of whether AI chatbots are behaviorally similar to humans. *Proceedings of the National Academy of Sciences of the United States of America*, *121*(9), Article e2313925121. https://doi.org/10.1073/pnas.2313925121

Mirowska, A. (2020). AI evaluation in selection. *Journal of Personnel Psychology*, *19*(3), 142–149. https://doi.org/10.1027/1866-5888/a000258

Nilsson, N. J. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Cambridge University Press.

Norton, L. W. (in press). Ethical and legal issues in ai for organizational consultants and managers. *Consulting Psychology Journal*.

Oswald, F. L., Behrend, T. S., Putka, D. J., & Sinar, E. (2020). Big data in industrial-organizational psychology and human resources management: Forward progress for organizational research and practice. *Annual Review of Organizational Psychology and Organizational Behavior*, *7*(1), 505–533. https://doi.org/10.1146/annurev-orgpsych-032117-104553

Pérez, J., Castro, M., & López, G. (2023). Serious games and AI: Challenges and opportunities for computational social science. *IEEE Access: Practical Innovations, Open Solutions*, *11*, 62051–62061. https://doi.org/10.1109/ACCESS.2023.3286695

Quillian, L., Pager, D., Hexel, O., & Midtbøen, A. H. (2017). Meta-analysis of field experiments shows no change in racial discrimination in hiring over time. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(41), 10870–10875. https://doi.org/10.1073/pnas.1706255114

Sajjadiani, S., Sojourner, A. J., Kammeyer-Mueller, J. D., & Mykerezi, E. (2019). Using machine learning to translate applicant work history into predictors of performance and turnover. *Journal of Applied Psychology*, *104*(10), 1207–1225. https://doi.org/10.1037/apl0000405

Schuler, H. (2006). Berufseignungsdiagnostik. In F. Petermann & M. Eid (Eds.), *Handbuch der Psych. Diagnostik* (pp. 717–729). Hogrefe.

Schuler, H. (2008). Improving assessment centers by the trimodal concept of personnel assessment. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, *1*(1), 128–130. https://doi.org/10.1111/j.1754-9434.2007.00024.x

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., & Hall, P. (2022). *Towards a standard for identifying and managing bias in artificial intelligence*. National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.1270

Shields, A., Nock, M. R., Ly, S., Manjaly, P., Mostaghimi, A., & Barbieri, J. S. (2024). Evaluation of stigma toward individuals with acne. *JAMA Dermatology*, *160*(1), 93–98. https://doi.org/10.1001/jamadermatol.2023.4487

Short, J. C., McKenny, A. F., & Reid, S. W. (2018). More than words? Computer-aided text analysis in organizational behavior and psychology research. *Annual Review of Organizational Psychology and Organizational Behavior*, *5*(1), 415–435. https://doi.org/10.1146/annurev-orgpsych-032117-104622

Society for Industrial Organizational Psychology. (2018). Principles for the validation and use of personnel selection procedures. *Society*.

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., & Teller, A. (2016). *Artificial intelligence and life in 2030: One hundred year study on artificial intelligence: Report of the 2015–2016 study panel*. https://ai100.stanford.edu/2016-report

Suen, H. Y., Chen, M. Y. C., & Lu, S. H. (2019). Does the use of synchrony and artificial intelligence in video interviews affect interview ratings and applicant attitudes? *Computers in Human Behavior*, *98*, 93–101. https://doi.org/10.1016/j.chb.2019.04.012

Thompson, I., Koenig, N., Mracek, D. L., & Tonidandel, S. (2023). Deep learning in employee selection: Evaluation of algorithms to automate the scoring of open-ended assessments. *Journal of Business and Psychology*, *38*(3), 509–527. https://doi.org/10.1007/s10869-023-09874-y

Tippins, N. T., Oswald, F. L., & McPhail, S. M. (2021). Scientific, legal, and ethical concerns about AI-based personnel selection tools: A call to action. *Personnel Assessment and Decisions*, *7*(2), Article 1. https://doi.org/10.25035/pad.2021.02.001

Verma, S., Ernst, M., & Just, R. (2021). *Removing biased data to improve fairness and accuracy*. arXiv preprint arXiv: 2102.03054.

Will, P., Krpan, D., & Lordan, G. (2023). People versus machines: Introducing the HIRE framework. *Artificial Intelligence Review*, *56*(2), 1071–1100. https://doi.org/10.1007/s10462-022-10193-6

Yarger, L., Cobb Payton, F., & Neupane, B. (2020). Algorithmic equity in the hiring of underrepresented IT job candidates. *Online Information Review*, *44*(2), 383–395. https://doi.org/10.1108/OIR-10-2018-0334

Zhang, Y., Feng, F., Wang, C., He, X., Wang, M., Li, Y., & Zhang, Y. (2020). How to retrain recommender system? A sequential meta-learning method. *Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval* (pp. 1479–1488). Association for Computing Machinery. https://doi.org/10.1145/3397271.3401167

Received May 23, 2024
Revision received September 27, 2024
Accepted September 30, 2024 ■