Assessing human factors in AI adoption by employees: a composite questionnaire for subjective user evaluation

Antonia Azzini*, Ilaria Baroni and Irene Celino

Cefriel, viale Sarca, 226, 20126, Milan, Italy

Abstract

The adoption of AI is reshaping job roles, improving efficiency and introducing new opportunities. However, critical aspects are also emerging, such as increasing the stress levels, the implications for occupational safety and changing the cognitive load of employees. To ensure safe and ethical integration, regulatory frameworks such as the AI Trustworthy Guidelines and the AI Act safeguard employee rights, improve transparency and mitigate risks. These standards also support individuals in assessing their willingness and attitude towards these technologies.

Starting from these aspects, we investigate whether the adoption of AI in the workplace can be evaluated through a principle-based, multi-factor user evaluation framework in compliance with legal and ethical principles. To ensure a comprehensive assessment of the impact of AI on employees, we propose a composite solution based on both European regulations and standard scales used to capture employees' subjective perceptions across different dimensions.

Keywords

AI User Evaluation, Cognitive and Ethical Aspects, Human-AI Interaction, AI Guidelines and Regulations

1. Introduction

The adoption of human-centered technologically advanced systems in industry is increasingly transforming the role of employees, who interact with complex tools, including those based on Artificial Intelligence (AI). While these innovations aim to support the worker and improve the efficiency of processes, they can also lead to an increase in cognitive workload, mental fatigue and perceived stress, possibly affecting well-being and performance.

These aspects are particularly marked in the use of AI-based technologies, for which, in addition, other human factors also play a role. Among them, the main factors refer, for instance, to the perceived usefulness through the use of AI, to the awareness of new capabilities gained from its learning, but, at the same time, also to a user's lack of trust in AI, the fear of becoming dependent on the technology, or of being replaced in his or her role. Moreover, the integration of AI in a workplace suggests, in this direction, a careful analysis concerning legal regulations and social implications to evaluate, and then guarantee, a safe, transparent and ethical use.

At the EU level, the AI Trustworthy Guidelines [1], published by the European Commission (EC), outlined the fundamental ethical principles for the development and use of AI in a trustworthy manner. Subsequently, as reported in [2], the AI Act [3] translated these principles into specific guidance, creating a binding regulatory framework to ensure that AI is developed and used following EU fundamental rights. The AI Act and the AI Trustworthy Guidelines provide a set of rules and directives to promote trustworthy AI, mitigating the risks of human rights violations, biased or discriminatory decisions, and potential negative impacts on employees. Such standards support the effective and sustainable adoption of AI in the job context. Focusing on them allows to safeguard the security and fundamental rights of individuals, increases the transparency and reliability of AI and reduces the risks associated with it while promoting its adoption by an employee. Through them, the employee can assess his/her

HHAI-WS 2025: Workshops at the Fourth International Conference on Hybrid Human-Artificial Intelligence (HHAI), June 9–13, 2025, Pisa, Italy

*Antonia Azzini

antonia.azzini@cefriel.com (A. Azzini); ilaria.baroni@cefriel.com (I. Baroni); irene.celino@cefriel.com (I. Celino)

0000-0002-0877-706 (A. Azzini); 0000-0002-0877-706 (I. Baroni); 0000-0001-9962-7193 (I. Celino)

willingness and the way he/she approach an AI-based solution. These aspects involve the analysis of multiple factors, which not only refer to the results obtained from the application of the solution to the work environment, but are strictly subjective, linked to an evaluation based on personal aspects.

Starting from this point of view, we are therefore interested to investigate the following research question: *in the case of adoption of AI-based solutions in the workplace, can we run a sound, principled and multi-factor user evaluation assessment, taking also into account the relevant regulations and ethical principles?* In this article we present our work to answer such questions: we started from a survey of the state of the art in quantitative scales and models for user assessment and, following the guidelines established by the EC through the AI Act and the Guidelines for Trustworthy AI, we propose a composite questionnaire that allows to assess users' perceptions related to the adoption of AI along different relevant investigation dimensions.

The paper is structured as follows: after introducing the state of the art about user assessment in Section 2, we present our approach in Section 3; then we discuss the qualitative findings of our early evaluation in Section 4, and we conclude the paper in Section 5 with some future work.

2. User Assessment: an analysis of the State of the Art

We first carried out a reasoned, focused and timely review of the state of the art of the subjective aspects involved in user evaluation of digital systems, while keeping as a reference the indications given by the AI Act for the evaluation of AI-based systems. While we did not run a systematic literature review, our analysis focused on the assessment methods that are either widely used for their proven reliability and validity, and those that emerged in the last years with specific reference to AI technologies.

The analysis shows an interesting set of scales, constructs and standard models that assess aspects such as cognitive load, psychological stress, mental fatigue and usability of such technologies, as well as their acceptance by the employee, the confidence and the intention to use them. The analysis also considered ethical aspects that may influence the acceptance and interaction with such systems in work contexts. In the following, we summarise the scales and standards resulting from the analysis, grouped according to the aspects that we consider relevant to this work.

Cognitive overload, fatigue and stress The state of the art highlighted several aspects related to cognitive overload, mental fatigue and perceived stress in the work context, focusing on those where advanced technologies require high cognitive effort and mental fatigue. This is usually associated with reduced concentration and performance ability after long-term cognitively intensive activity. Perceived stress is related to environmental, organisational and technological factors, with a focus on the role of advanced digital interfaces and automated systems [4, 5, 6, 7].

The most relevant scales arising from the literature are: the Copenhagen Psychosocial Questionnaire II (COPSOQII) [8], which measures psychological well-being and work-related psychosocial factors, the Fatigue Assessment Scale (FAS) [9, 10], and the Individual Strength Checklist (CIS) [11], both focusing on mental fatigue. The General Health Questionnaire-12 (GHQ-12) [12, 13] is a well-known scale for general psychological health, while the Perceived Stress Scale-10 (PSS-10) [14] is used for subjective perception of stress, and the NASA Task Load Index (NASA-TLX) [15], is known to assess mental workload, based on multiple dimensions. For the aim of our work, however, we will not consider high-level scales that are too general, nor scales that are too focused on one or limited cognitive or psychological aspects.

Acceptance and Usability The evaluation of user acceptance and usability of technology in the workplace is essential for understanding how employees interact with new systems and tools. Various standardised models have been developed to assess these aspects, each serving a specific purpose. The usability assessment and technology acceptance are widely represented in the literature through standardised models such as the Technology Acceptance Model (TAM) [16], Unified Theory of Acceptance and Use of Technology (UTAUT) [17] and System Usability Scale (SUS) [18]. Each of them is defined by a representative scale of constructs. TAM analyses the factors that influence the adoption of new technologies, by defining a set of items, respectively, for the constructs: perceived usefulness (PU), ease of use (PEU), and behavioural intention (BI), key elements that determine user acceptance of a system. UTAUT defines four core constructs (performance expectancy, effort expectancy, social influence, and facilitating conditions) used as direct determinants of behavioural intention. Finally, SUS provides a quick and reliable method to assess the usability of a system through a standardised set of questions, measuring the degree of intuitiveness and accessibility perceived by the user.

AI human perception Assessing users' reactions to AI-based systems is essential to understand their acceptance and impact on the job tasks they have to perform. Several standardised scales have been developed to measure key factors such as trust, transparency, usability, anxiety, and perceived career implications. These instruments provide structured methodologies to assess how employees perceive AI, both as an opportunity and as a critical issue to be addressed.

The literature also proposed several scales to assess the perception of AI in work contexts, focusing on factors such as self-efficacy in AI learning [19], job stress, perceived AI-induced job insecurity [20], the level of human-machine interactivity, and perceived career achievement [21]. The literature used these scales to assess the employee's confidence in acquiring AI skills and to assess how the technology is perceived as an opportunity or a threat to his/her professional growth.

Other relevant scales analysed include: the AI Trust and Attitudes Readiness Index (ATTARI-12) [22], which measures trust and predisposition towards AI; the General Attitudes toward Artificial Intelligence (GAAI) [23, 24], which focuses on general perceptions of AI in society and at work; the Artificial Intelligence Anxiety Scale (AIAS) [25], which assesses the level of anxiety and concern related to the use of AI in the work context.

The Assessment List for Trustworthy AI (ALTAI) [26], presented by European Commission and defined by the high-level expert group on AI (AI HLEG) is a generic framework related to the previously mentioned Trustworthy Guidelines on AI; some other European Projects [1] leveraged such a framework as a tool to measure, in real-world contexts, the degree of reliability of AI-based systems according to both ethical and technical principles, specifically. However, we considered ALTAI as too qualitative and generic for our purposes, therefore, we preferred to operate a selection on the quantitative scales provided in the literature.

Another aspect that has been explored in literature is Explainable Artificial Intelligence (XAI), related to the impact of advanced technological solutions, including AI-based systems, to ensure a trade-off between technological complexity and effective usability [27, 28]. XAI refers to Artificial Intelligence techniques and models designed to make AI decision-making processes understandable and interpretable by users. In particular, the adoption of the trust construct defined by the XAI scale [29] is useful in assessing the level of employee trust in AI technologies, as greater transparency in AI systems can foster the acceptance and aware adoption of these solutions in work contexts, as well as improve human perceptions of them. Indeed, in our previous research, we extended the TAM model with XAI constructs defining the AI-TAM model [30] for the assessment of AI systems with human-in-the-loop.

Also for the assessment of AI human perception, for the purpose of this paper, scales that assess general attitudes of a user, or that do not assess specifically more subjective factors, are not considered.

AI biases The literature research also shows aspects related to possible biases that the use of AI systems may introduce with respect to a user.

Some literature contributions have also analysed the role of gender and age biases in work contexts where AI-based technologies are adopted, studying their potential impact on employees' perceptions of fairness and inclusion. The research did not reveal any standard scales or models specifically adopted for assessing such biases from the employees' perspective [31]. However, information such as gender and age were considered as factors in statistical analyses conducted on the evaluation questionnaires conducted in the experimental sessions. An example was reported in [22], where the

[1] Cf. http://ai4realnet.eu/

impact of participants' age and gender on their attitudes towards AI was studied. The analysis conducted in the study showed that the influence of gender was rated as insignificant, while a higher age may be significantly correlated with more aversive thoughts, feelings and behavioural intentions towards AI.

3. The Approach

The discussed state-of-the-art analysis (see Section 2) allow us to answer our research question and define a comprehensive user evaluation questionnaire which fulfils the following requirements:

- takes into account investigation dimensions suggested by the AI Act directives and the Trustworthy AI Guidelines defined by the European Commission,
- is mainly focused on a subjective evaluation of the user (opposed to a technical evaluation of how users interact with AI technology),
- is sound, principled and based on multiple relevant factors, to support the user in a careful and in-depth assessment on his/her own experience in the job context,
- is composed by a number of factors and item that, on the one hand, allows for the investigation of the different factors, but, on the other hand, has a limited set of questions to carry out the assessment in a few minutes.

To meet these requirements, we first selected the aspects of the AI Act and related guidelines to be taken into account in a person's subjective assessment of AI-based technology. Section 3.1 provides a description of the choices and corresponding rationale. Then, we selected, from the various scales and constructs analysed in the literature, those that allow us to perform a subjective assessment of the user in an effective and comprehensive way, taking into account the main human aspects resulting from the analysis, such as cognitive load and mental fatigue, acceptance and usability of the technology and the different factors that contribute to define a human AI perception. Section 3.2 describes all the selected scales, by motivating our choices. In doing this, we guided the choice of scales and constructs by also considering the number of items in their definition and any possible overlap.

Finally, we present a summary with the mapping between the AI Act indications and the Guidelines for Trustworthy AI on the one side, and the different scales and constructs that we select and propose as a comprehensive user assessment questionnaire on the other side.

3.1. AI Directives Analysis

The adoption of human-centred technologically advanced systems in industry is increasingly transforming the role of employees, who interact with complex tools, including those based on AI.

Nevertheless, the integration of AI in industry requires a careful analysis concerning legal regulations and social implications to guarantee its safe, transparent and ethical use. At the EU level, the AI Act [3] and the AI Trustworthy Guidelines [1] provide a framework of rules and directions to promote trustworthy AI, mitigating the risks of human rights violations, biased or discriminatory decisions, and potential negative impacts for employees. Such standards support the effective and sustainable adoption of AI in industry. It is therefore important to focus on them to safeguard the security and fundamental rights of individuals, promote innovation, increase the transparency and reliability of AI and reduce the risks associated with its usage.

In particular, concerning the AI Act, we conducted a preliminary analysis, based on the excellent summary provided by VAIR and AIRO. VAIR (Vocabulary of AI Risks) [32] is an open vocabulary, describing concepts for the risks defined in the AI Act. It identifies, documents and classifies AI risks, using the core concepts of the AI Risk Ontology (AIRO) [33], which formalises the main systems and activities described in the AI Act. Our analysis extracted a selection of the most relevant concepts for the scope of our project.

For example, we focused our attention on risks related to: evaluating employees' behaviour, performance and learning outcomes; decision-making and task allocation; cognitive computing and personal impacts such as the impact on well-being, psychological health, non-discrimination rights and measures of supervision and involvement of people.

Complementing the AI Act, the European Union has established the already mentioned Trustworthy AI Guidelines [1], which outline key requirements for trustworthy AI systems. These guidelines emphasise principles such as human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being, and accountability. By adhering to these principles, organisations can foster trust and acceptance of AI technologies among users and stakeholders.

Our analysis of the Trustworthy AI Guidelines highlighted the following requirements, among the seven described in the document, as the most relevant to our objective: ”human agency and oversights”, by posing attention to aspects like fundamental rights, human agency and oversight; "transparency", by considering explainability and communication, "diversity, non-discrimination and fairness", by considering accessibility and universal design and stakeholder participation, and finally "societal and environmental well-being", by mainly considering the social impact.

3.2. Scale and Construct Selection

While complex tools aim to support the employee and improve the efficiency of processes, they can also lead to an augmented cognitive workload, mental fatigue and perceived stress, possibly affecting well-being and performance. Therefore, for all these aspects, after a first analysis of the main models and standards from the state of the art (see Section 2), we select the scales and constructs that best meet the AI Act and Trustworthy guidelines illustrated before (see Section 3.1) and we motivate our choice.

Concerning the assessment of *cognitive load*, *fatigue* and *stress*, we chose NASA-TLX [15] since it allows us to meet some of the requirements from the AI-Act, such as the impact on well-being, with particular attention to the physical and psychological impacts. The NASA-TLX provides a detailed assessment of cognitive load, by considering several aspects, such as mental, physical and temporal demand, effort, frustration and performance. Moreover, its task-oriented nature makes this scale particularly suitable for analysing tasks performed in highly interactive work environments with advanced technologies.

For what concerns *acceptance* and *usability*, we chose to consider both TAM [16] and SUS [18] scales for the definition of the evaluation questionnaire. Through the respective constructs, indeed, we intend to assess aspects such as employees’ learning of new solutions for performing tasks, and the tool evaluation w.r.t. human perceptions like ease of use, usefulness, clarity and behavioural intention to adopt such solutions.

Finally, regarding the AI human perception, the choice of scales and constructs was guided by the requirements and key aspects that were found to be the most relevant for assessing the reliability of AI. In detail, the *job replacement* construct of the AIAS scale [25] assesses aspects related to human oversight and social well-being; this construct allows us to evaluate the impact on the employee, together with distortions in human behaviour, related to anxiety generated from the AI adoption. The *self-efficacy* in AI learning scale [19], on the other hand, allows us to assess aspects mainly concerning the explainability and the employee’s ability to self-learn and develop skills, while the perception of *career achievement* scale [21] can be used to consider aspects related to human fundamental rights and overall supervision, and the social impact perceived by the employee him/herself. Finally, we selected the *Trust in AI* construct from the XAI scale [29], which allows us to meet the requirements relating to the human perception of safety and trust on one hand, and the stability and reliability of the solution on the other.

3.3. Proposed Questionnaire and Mapping

Based on the considerations offered in the previous section, in Table 1, we present the final list of the selected scales and constructs, mapped towards the relevant concepts from the AI-related regulations and guidelines. In particular, in Column 2 of Table 1 we report the relevant factors which, among several reported in the literature, are selected because they support subjective assessments of a user. These factors are represented in the questionnaire through the selection of scales specified in Section 3.2, that we also report in Column 1. Then, in the following columns of Table 1, we report, for such scales and factors, the corresponding dimensions suggested by the Trustworthy AI Guidelines (Column 3) and the AI Act directives (Column 4) defined by the European Commission and described in Section 3.1. The user evaluation questionnaire we propose for AI assessment in the workplace is therefore constituted by the items defined by each selected construct and scale.

Finally, we report the details of the overall evaluation questionnaire, by indicating the items that define the constructs of each considered scale and the corresponding rating ranges. More specifically, in Table 2 we report the items for assessing aspects such as cognitive load, psychological stress, mental fatigue and usability of such technologies, as well as their acceptance by the employee, confidence and intention to use them. In Table 3 we report the items for assessing specific aspects concerning human perceptions of AI adoption. Each item corresponds to a single question in the evaluation questionnaire.

Scale	**Construct**	**Trustworthy AI Reference [1]**	**AI-Act Reference [3]**
NASA-TLX [15]	cognitive overload, mental fatigue	human agency and oversight, societal well-being (social impact)	Well-being and psychological health, cognitive effort, fatigue, social impact
TAM [16]	ease of use, perceived usefulness	human agency and oversight, fundamental rights, non-discrimination and fairness	behaviour, performance and outcomes evaluation, task allocation perception, involvement, non-discrimination
SUS [18]	system usability	transparency (explainability and communication)	behaviour, decision making, performance and outcomes evaluation, involvement, task-allocation perception, learning outcomes
XAI [29]	trust on AI	human agency and oversight, robustness and safety (reliability), societal well-being (social impact)	well-being (social impact), behaviour, decision making, measures of supervision and involvement
AIAS [25]	job replacement	human agency and oversight societal well-being(social impact)	psychological health, behaviour, task allocation perception
self-efficacy in AI learning [19]	self-efficacy	Human agency and oversight, transparency (explainability)	learning outcomes and performance evaluation, involvement, behaviour
AI perception on career achievement [21]	job performance evaluation	Human agency and oversight (fundamental rights), societal well-being (social impact)	fundamental rights, non-discrimination, task-allocation perception, performance evaluation

Table 1
Mapping between selected constructs and AI dimensions

4. Preliminary Application and Discussion

We performed a preliminary application of the questionnaire within the PERKS project[2]. Co-funded by the Horizon Europe Programme, PERKS adopts AI to provide digital support to industrial practitioners in creating, using and governing procedural knowledge, i.e. the 'know-how' in industrial processes [34]. By prioritizing human-centered, trustworthy AI and following a human-in-the-loop paradigm, PERKS puts industry workers at the center, in line with the Industry 5.0 vision, to satisfy their concrete needs, to provide AI-powered digital tools to perform their tasks better and more easily, following a human-in-the-loop paradigm to enhance the technologies and the solutions.

[2]Cf. https://perks-project.eu/

Scale	Construct	Item	Answer Type
NASA-TLX	N-TLX1	How mentally demanding was the activity?	(1-7)
	N-TLX2	How hard did you have to work to accomplish your level of performance?	
	N-TLX3	How successful were you in accomplishing what you were asked to do?	
	N-TLX4	How physically demanding was the task?	
	N-TLX5	How hurried or rushed was the pace of the task?	
	N-TLX6	How insecure, discouraged, stressed and annoyed were you?	
TAM	PU-1	Using this tool in my job would enable me to accomplish tasks more quickly	(1-7)
	PU-2	Using this tool would improve my job performance	
	PU-3	Using this tool in my job would increase my productivity	
	PU-4	Using this tool would enhance my effectiveness on the job	
	PU-5	Using this tool would make it easier to do my job	
	PU-6	I would find this tool useful in my job	
	PEU-1	Learning to operate this tool would be easy for me	
	PEU-2	I would find it easy to get this tool to do what I want it to do	
	PEU-3	My interaction with this tool would be clear and understandable	
	PEU-4	I would find this tool would be clear and understandable	
	PEU-5	It would be easy for me to become skilful at using this tool	
	PEU-6	I would find this tool easy to use	
	BI	I think I would use this tool regularly at work	
SUS	SUS-1	I think that I would like to use this system frequently	(1-5)
	SUS-2	I found the system unnecessarily complex	
	SUS-3	I thought the system was easy to use	
	SUS-4	I think that I would need the support of a technical person to be able to use this system	
	SUS-5	I found the various functions in this system were well integrated	
	SUS-6	I thought there was too much inconsistency in this system	
	SUS-7	I would imagine that most people would learn to use this system very quickly	
	SUS-8	I found the system very cumbersome to use	
	SUS-9	I felt very confident using the system	
	SUS-10	I needed to learn a lot of things before I could get going with this system	
Career achievement	Achiev-1	I can still solve work-related problems effectively with my existing knowledge and skills	(1-7)
	Achiev-2	I feel that my work is still important to the business and customers	
	Achiev-3	In my opinion, I can still use the knowledge and skills I am good at	
	Achiev-4	I can still fully demonstrate the value of my work	
	Achiev-5	I can still easily understand and deal with problems at work	
	Achiev-6	I feel that my work can still have an important impact on the business and customers	

Table 2
Evaluation Questionnaire: part 1

The AI-powered solution implemented in PERKS supports the user in (1) collecting procedural knowledge, also thanks to automatic extraction from documents, (2) managing procedural knowledge with ontologies and knowledge graphs, and (3) supporting the access and reuse of such knowledge at procedure execution time, also through a conversational AI chatbot.

Scale	Construct	Item	Answer Type
XAI	Trust on AI	I would be confident in the tool. I feel that it works well	(1-7)
	Trust on AI	I feel that, by relying on the tool, I will get the right answers	
	Trust on AI	I tend not to trust the tool	
	Trust on AI	It seems that the tool can perform the task better than a novice human user	
	Trust on AI	The tool is very reliable. I could count on it to be correct all the time	
AIAS	Job Rep-1	I am afraid that AI techniques/products will replace someone's job	(1-7)
	Job Rep-2	I am afraid that if I begin to use AI techniques/products I will become dependent upon them and lose some of my reasoning skills	
	Job Rep-3	I am afraid that an AI technique/product may make us dependent	
	Job Rep-4	I am afraid that an AI technique/product may make us even lazier	
	Job Rep-5	I am afraid that an AI technique/product may replace humans	
Self-efficacy in AI learning	Self-Eff-1	I am confident in my ability to learn artificial intelligence technology appropriately in my work	(1-7)
	Self-Eff-2	I'm able to learn artificial intelligence technology to perform my job well, even when the situation is challenging	
	Self-Eff-3	I can develop my competencies needed for my job through AI technology learning	
	Self-Eff-4	I will be able to learn important information and skills from my AI training	

Table 3
Evaluation Questionnaire: part 2

The AI solution provided by PERKS was deployed in three different industrial contexts for a piloting phase. The procedural knowledge use cases are diverse and complementary: LOTO safety procedures for shutting down a production line for maintenance intervention (in the home appliances company Beko Europe in Italy), commissioning processes CNC systems (in the machine automation manufacturing company Fagor in Spain), and microgrid optimization operations (in the microgrid tesbed of Siemens in Austria).

The piloting phase involved a limited number of testers from the three companies, with 13 employees of different ages, gender, and company roles. The employees followed an experimentation protocol that was shared as a guide, and used the AI application implemented in PERKS to perform tasks that are usually carried out without such support. At the end of the pilot session, each participant finally evaluated his/her experience by filling in the user evaluation questionnaire that we propose in this paper.

To give a measure to each construct of the questionnaire, we used the following approach. We calculated the Cognitive Load by applying the unweighted aggregated score of the NASA-TLX dimensions, the System Usability by applying the standard formula for System Usability score, while we calculated the remaining constructs by applying the average of the individual score given by a user for each of the construct-related items.

We do not have the ambition to claim that what we present here is a statistical validation of the questionnaire, as this will be subject of our future work. What we present are the results of a group of interesting preliminary applications conducted in different industrial areas. In fact, the collected results, even if not statistically significant given the small sample, still allowed us to make some considerations, about the participants' subjective evaluation of AI adoption in their workplace. Overall, the results obtained from this early evaluation enabled us to identify strengths, criticalities and, consequently, specific areas of improvement for the assessed AI solution.

On the one hand, the questionnaire enabled the identification of improvements that can be implemented over the tested application, to meet some criticalities highlighted by the evaluation questionnaire. Indeed, even if all the participants are, in general, quite willing to move towards a *regular adoption* of this technology at work, with an average 67% behavioural intention score, some employees are less willing than others. Additionally, all pilots emphasised that the perceived *cognitive workload* and mental fatigue are still medium-high, with a global 39% score (the lower the score, the lower the perceived load). Concerning the application itself, the results also indicated the need to improve confidence w.r.t. the perceived system *usability*, with an average score of 60%. We obtained similar results for the perceived *ease of use* (67%) and for the perceived *usefulness* (66%).

On the other hand, the questionnaire also revealed more specific aspects about the adoption of AI and its perception by the test participants. In particular, regarding the *self-efficacy* in AI perception learning, participants gave value to the knowledge they have gained, feel confident in their ability to learn AI technology properly in terms of their job, and understand the potential value of AI in improving their knowledge and performance. The average score of 77% confirms such insight. The participants are, in general, confident, w.r.t. the *career achievement*, with an average score of 83%, in their ability to carry out their work, and in their ability to effectively address and solve problems, thus significantly and positively contribute to their work. At the same time, however, for all the three pilots, the participants were concerned about the effect use of AI on their job: they fear that, by using such technologies, they may become dependent on them, lose some of their reasoning skills, or even be replaced by such a technology. This is proven by a high *job-replacement* score of 59% (the higher the score, the higher the fear of replacement). A further aspect that requires attention concerns the participants' *trust in AI* and its performance, which is still low to medium with an average score of 55%. These results mean that the adoption of AI in the examined industrial environments is still far from full acceptance.

With respect to our initial research question, the results obtained from the application of the questionnaire allow us state that, in the case of the adoption of AI-based solutions in the workplace, it is possible to run a sound, principled and multi-factor user evaluation, by taking also into account the relevant regulations and ethical principles. Furthermore, our testers took on average 7 minutes to fill in the questionnaire, satisfying our goal to conduct a multi-dimensional assessment in a reasonable time.

5. Conclusions

In this paper, we proposed a composite questionnaire to evaluate users' adoption of AI in workplaces, according to significant survey dimensions, and following the relevant AI Act regulations and European Trustworthy Guidelines. We performed an initial application of the questionnaire within a specific project implementing an AI-based solution evaluated within three case studies in different industrial contexts. The results obtained from an early assessment allow us to assess the questionnaire as a potentially valid tool for a multi-dimensional user evaluation.

Starting from this initial experience, we can outline some further steps to improve the proposed questionnaire. As a first step, we would like to apply the questionnaire on a large scale pilot to validate it from a statistical point of view. This would allow us to test its reliability and consistency, increasing its applicability in different real-world contexts. Large-scale tests would also enable us to analyse the relationships, and potential influencing links, among the constructs that define the questionnaire (e.g. through exploratory and confirmatory factor analysis).

Our early results confirmed us that the systematic and continuous integration of artificial intelligence in the workplace needs particular attention. For this reason, we would like to investigate new dimensions that might reveal additional aspects of user evaluation concerning the systematic use of AI. We consider this as a growing research area that should be better explored to understand and evaluate the impact of such emerging innovative technologies on individuals.

Therefore, we would like to explore further aspects not yet covered by the questionnaire that could have a significant impact on user evaluation. This would allow us to extend our analysis to a wider perspective, leading to even more detailed and complete user evaluations. This work in turn would then help to improve the design and development of AI-based solutions to properly meet the emerging workplace needs.

Acknowledgments

This work is partially supported by the PERKS project, co-funded by the European Commission (Grant id 101120323). The authors would like to thank all the participants who contributed to the results of this work.

Declaration on Generative AI

The author(s) have not employed any Generative AI tools.

References

[1] E. C., Ethics guidelines for trustworthy ai, 2019. URL: http://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai.

[2] A. Bringas Colmenarejo, L. Nannini, A. Rieger, K. M. Scott, X. Zhao, G. K. Patro, G. Kasneci, K. Kinder-Kurlanda, Fairness in agreement with european values: An interdisciplinary perspective on ai regulation, in: Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society, Association for Computing Machinery, New York, NY, USA, 2022, p. 107–118. doi:10.1145/3514094.3534158.

[3] E. Commission, Ai-act european commission, http://artificialintelligenceact.eu/the-act/, ???? Accessed: 2025-03-25.

[4] A. Beurskens, U. Bültmann, I. Kant, J. Vercoulen, G. Bleijenberg, G. Swaen, Fatigue among working people: Validity of a questionnaire measure, Occupational and environmental medicine 57 (2000) 353–7. doi:10.1136/oem.57.5.353.

[5] J. De Vries, H. Michielsen, G. Van Heck, Assessment of fatigue among working people: a comparison of six questionnaires, Occupational and Environmental Medicine 60 (2003) i10–i15.

[6] K. W. Kim, H. C. Lim, J. H. Park, S. G. Park, Y. J. Park, H. H. Cho, Developing a Basic Scale for Workers' Psychological Burden from the Perspective of Occupational Safety and Health, Safety and Health at Work 9 (2018) 224–231. doi:10.1016/j.shaw.2018.02.004.

[7] G. Blandino, F. Montagna, M. Cantamessa, Workload and stress evaluation in advanced manufacturing systems, Materials Research Proceedings 35 (2023).

[8] J. Pejtersen, T. Kristensen, V. Borg, J. Bjorner, The second version of the copenhagen psychosocial questionnaire (copsoq ii), Scandinavian Journal of Public Health 38 (2010) 8–24. doi:10.1177/1403494809349858.

[9] A. Shahid, K. Wilkinson, S. Marcu, C. M. Shapiro, Fatigue Assessment Scale (FAS), in: STOP, THAT and One Hundred Other Sleep Scales, Springer, New York, NY, 2012, pp. 161–162. doi:10.1007/978-1-4419-9893-4_33.

[10] H. J. Michielsen, J. De Vries, G. L. Van Heck, Psychometric qualities of a brief self-rated fatigue measure: The Fatigue Assessment Scale, Journal of Psychosomatic Research 54 (2003) 345–352. doi:10.1016/S0022-3999(02)00392-6.

[11] A. J. H. M. Beurskens, U. Bültmann, I. Kant, J. H. M. M. Vercoulen, G. Bleijenberg, G. M. H. Swaen, Fatigue among working people: validity of a questionnaire measure, Occupational and Environmental Medicine 57 (2000) 353–357. doi:10.1136/oem.57.5.353.

[12] T. Gnambs, T. Staufenbiel, The structure of the general health questionnaire (ghq-12): two meta-analytic factor analyses, Health Psychology Review 12 (2018) 179–194. doi:10.1080/17437199.2018.1426484.

[13] U. Bültmann, I. Kant, S. V. Kasl, A. J. H. M. Beurskens, P. A. van den Brandt, Fatigue and psychological distress in the working population: Psychometrics, prevalence, and correlates, Journal of Psychosomatic Research 52 (2002) 445–452. doi:10.1016/S0022-3999(01)00228-8.
[14] G. Blandino, How to measure stress in smart and intelligent manufacturing systems: A systematic review, Systems 11 (2023) 167. doi:10.3390/systems11040167.
[15] S. Rubio, E. Díaz, J. Martín, J. M. Puente, Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods, Applied Psychology 53 (2004) 61–86. doi:10.1111/j.1464-0597.2004.00161.x.
[16] F. D. Davis, Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, MIS Quarterly 13 (1989) 319–340. doi:10.2307/249008.
[17] V. Venkatesh, M. G. Morris, G. B. Davis, F. D. Davis, User acceptance of information technology: Toward a unified view, MIS quarterly (2003) 425–478.
[18] J. Brooke, SUS - a quick and dirty usability scale, CRC Press, 1996, pp. 189–194.
[19] B.-J. Kim, J. Lee, The mental health implications of artificial intelligence adoption: the crucial role of self-efficacy, Humanities and Social Sciences Communications 11 (2024) 1–15. doi:10.1057/s41599-024-04018-w.
[20] B.-J. Kim, M.-J. Kim, J. Lee, Code green: ethical leadership's role in reconciling AI-induced job insecurity with pro-environmental behavior in the digital workplace, Humanities and Social Sciences Communications 11 (2024) 1–16. doi:10.1057/s41599-024-04139-2.
[21] J. Xu, X. Tang, E.-C. Chang, H. Peng, Working with AI: the impact of organizational intelligent service strategy on employees' perception of career achievement, Humanities and Social Sciences Communications 11 (2024) 1–20. doi:10.1057/s41599-024-03265-1.
[22] J.-P. Stein, T. Messingschlager, T. Gnambs, F. Hutmacher, M. Appel, Attitudes towards AI: measurement and associations with personality, Scientific Reports 14 (2024). doi:10.1038/s41598-024-53335-2.
[23] A. Schepman, P. Rodway, Initial validation of the general attitudes towards artificial intelligence scale, Computers in Human Behavior Reports 1 (2020) 100014. doi:https://doi.org/10.1016/j.chbr.2020.100014.
[24] F. Kaya, F. Aydin, A. Schepman, P. Rodway, O. Yetişensoy, M. Demir Kaya, The Roles of Personality Traits, AI Anxiety, and Demographic Factors in Attitudes toward Artificial Intelligence, International Journal of Human–Computer Interaction 40 (2024) 497–514. doi:10.1080/10447318.2022.2151730.
[25] Y.-Y. Wang, Y.-S. Wang, Development and validation of an artificial intelligence anxiety scale: an initial application in predicting motivated learning behavior, Interactive Learning Environments 30 (2022) 619–634. doi:10.1080/10494820.2019.1674887.
[26] AI-HLEG, Assessment list for trustworthy artificial intelligence (altai) for self-assessment, http://digital-strategy.ec.europa.eu/en/library/assessment-list-trustworthy-artificial-intelligence-altai-self-assessment, 2019.
[27] W. Saeed, C. Omlin, Explainable AI (XAI): A systematic meta-survey of current challenges and future opportunities, Knowledge-Based Systems 263 (2023). doi:10.1016/j.knosys.2023.110273.
[28] N. Al-Ansari, D. Al-Thani, R. S. Al-Mansoori, User-Centered Evaluation of Explainable Artificial Intelligence (XAI): A Systematic Literature Review, Human Behavior and Emerging Technologies 2024 (2024). doi:10.1155/2024/4628855.
[29] R. R. Hoffman, S. T. Mueller, G. Klein, J. Litman, Metrics for Explainable AI: Challenges and Prospects, 2019. doi:10.48550/arXiv.1812.04608.
[30] I. Baroni, G. Re Calegari, D. Scandolari, I. Celino, AI-TAM: a model to investigate user acceptance and collaborative intention in human-in-the-loop AI applications, Human Computation 9 (2022) 1–21. doi:10.15346/hc.v9i1.134, 1.
[31] J. D. Patón-Romero, R. Vinuesa, L. Jaccheri, M. T. Baldassarre, State of gender equality in and by artificial intellicenge, 2022. doi:10.33965/ijcsis_2022170203.
[32] D. Delaram Golpaygani, H. J. Pandit, D. Lewis, Vocabulary of AI Risks - VAIR, 2024. URL: https:

//w3id.org/vair.
[33] D. Delaram Golpaygani, H. J. Pandit, D. Lewis, The ai risk ontology - airo, 2000. URL: https://w3id.org/airo.
[34] I. Celino, V. A. Carriero, A. Azzini, I. Baroni, M. Scrocca, Procedural knowledge management in industry 5.0: Challenges and opportunities for knowledge graphs, Web Semantics: Science, Services and Agents on the World Wide Web 84 (2025). doi:10.1016/j.websem.2024.100850.