Cognitive personalization for online microtask labor platforms: A systematic literature review

Dennis Paulino[1] · **António Correia**[1] · **João Barroso**[1] · **Hugo Paredes**[1]

Abstract

Online microtask labor has increased its role in the last few years and has provided the possibility of people who were usually excluded from the labor market to work anytime and without geographical barriers. While this brings new opportunities for people to work remotely, it can also pose challenges regarding the difficulty of assigning tasks to workers according to their abilities. To this end, cognitive personalization can be used to assess the cognitive profile of each worker and subsequently match those workers to the most appropriate type of work that is available on the digital labor market. In this regard, we believe that the time is ripe for a review of the current state of research on cognitive personalization for digital labor. The present study was conducted by following the recommended guidelines for the software engineering domain through a systematic literature review that led to the analysis of 20 primary studies published from 2010 to 2020. The results report the application of several cognition theories derived from the field of psychology, which in turn revealed an apparent presence of studies indicating accurate levels of cognitive personalization in digital labor in addition to a potential increase in the worker's performance, most frequently investigated in crowdsourcing settings. In view of this, the present essay seeks to contribute to the identification of several gaps and opportunities for future research in order to enhance the personalization of online labor, which has the potential of increasing both worker motivation and the quality of digital work.

Keywords Online microtasks labor · Crowdsourcing · Cognition · Personalization

1 Introduction

A significant advantage of digital labor platforms is the reduce of the regulatory impact between employer and worker, as pointed out in Graham and co-authors (2017). In

✉ Dennis Paulino
  dennis.l.paulino@inesctec.pt

[1] INESC TEC and University of Trás-os-Montes e Alto Douro, Vila Real, Portugal

🖄 Springer

this model, workers are self-employed and national labor laws are rarely applied. A potential benefit of this lack of regulation is the flexibility that can result in greater job autonomy. Thus, digital workers can push the boundaries of local labor markets and then, increase the conditions (e.g., price per task) for which they are willing to work. Responding to the adverse pandemic effects of COVID-19 outbreak that have left a significant part of the workforce unable to work in person (Vyas and Butakhieo 2021), the solution found was to put (almost) all workers working from home. For that reason, there is growing empirical evidence pointing to the need for flexible and inclusive teleworking practices. Crowdsourcing can be seen as a practice for engaging large and anonymous groups (crowds) of workers, especially from online communities, to fulfill a shared goal. In Estellés-Arolas and González-Ladrón-de-Guevara (2012), crowdsourcing is defined as an online activity where a person or organization proposes the assignment of a task to a group of individuals that must receive some compensation. The tasks performed by crowd workers typically comprise those that allow object detection and image classification, text translation, and questionnaire-based items (Kucherbaev et al. 2016). In a global context experiencing different work requirements, the need to crowdsource tasks is associated with human computation, which have the purpose of organizing the tasks executed by humans to perform computation processes (Law and Ahn, 2011). Through this lens, crowdsourcing can be seen as the optimal usage of human computation, which is particularly useful and helpful for companies (Ipeirotis 2010; Nguyen Hoang, Pedro, and David, 2017) and scientific institutions (Cooper et al. 2010; Raddick et al. 2019), while even contributing to advance artificial intelligence research (Chang et al. 2017; Correia et al. 2018; Muller et al. 2015; F. A. Schmidt 2019).

Actually, the work in crowdsourcing is characterized by having a wide range of necessary skills to accomplish a task (Webster 2016). In one hand, there are creative forms of digital labor that may require professional skills such as video editing (Roth and Kimani 2014), product design (Bayus 2013), or software development (Sarı et al. 2019; Stol and Fitzgerald 2014). On other hand, there are also small units of work associated with low payments under the umbrella of "microwork" (Zyskowski et al. 2015a). Under such circumstances, microwork has the potential for being the pinnacle of accessible digital work, for example, giving remote job opportunities to people with special needs from developing countries (Galpaya et al. 2018; Mtsweni and Burge 2014). However, the high amount of digital labor supply, particularly in the case of crowdsourcing, also leads to a wide range of skills' assumptions that raise challenges to employers and workers. While the former must handle with poor quality of the work done (Daniel et al. 2018), the latter deals with the frustration of disposal of time due to the fact that an employer can reject the work performed and in extreme cases do not even give any credit or compensation to workers (Deng and Joshi 2016).

Personalization in computer systems is defined as the process of adjusting the functionality or the interface to increase personal relevance from the user's perspective. Customization refers to providing the customization options for the user to customize as they prefer. A significant difference between personalization and customization is in the adaptation role: While personalization is done implicitly by the system, customization is done explicitly by the user. Both adaptations methods have seven common strengths of increasing interactivity, ease of use, usefulness, trust, credibility,

users' perception of a system's relevance and users' self-efficacy (Orji et al., 2017). Personalization can be used even on a large number of resources, such as YouTube large panoply of videos, where recommendations suggested to users are effective and engaging despite the wide range of videos available (Covington et al., 2016). While personalization allows to reduce the user's burden in the personal adaptation of the system, it is considered that it can raise problems related to privacy. One solution to tackle the privacy issues and reduce the user burden is to elicit user adaptation's preferences and needs through small, interactive tasks (Paulino et al., 2020). This can be applied for adaptation in microtask interface design, through cognitive micro-tasks. Thus, a correct assignment of tasks taking into account the particularities of digital work in terms of skills and cognitive abilities involved can solve both problems and alleviate the burden resulting from disproportionate assignments.

With the vast development of microwork, several studies have explored task assignment in crowdsourcing settings (Difallah et al. 2013; Fan et al. 2015; Gadiraju et al. 2019; Hu et al. 2016; Kazai et al. 2011, 2012; Lykourentzou et al. 2016; Rahman et al. 2015; Shaw et al. 2011; Zheng et al. 2015). Collaborative tasks such as translating documents can have better results if the working groups are assigned based on age (Rahman et al. 2015) and personality traits (Kazai et al. 2011; Lykourentzou et al. 2016). In addition, behavioral traces of the worker can be used for task assignment. In this regard, data collected from previous tasks (e.g., mouse movements, clicks, and keystrokes) can be used for measuring workers' performance and assign them to suitable tasks (Gadiraju et al. 2019). Behavior traces can also be identified on recommenders systems based on individual's interactions to elicit personalized user preferences (Quadrana et al., 2017; Yang et al., 2018). Additionally, the unintentional activity of users, based only on the mouse and keypresses data, can provide hints regarding that user, such as the prediction of age and gender, so the personalization made for the user can be improved (Pentel, 2017). Further experimentation in Fan et al., (2015) aimed to construct an adaptive framework for matching the worker abilities to the task properties as measured by his/her past performance by using qualification tests. Moreover, a related line of research examined the use of social networks of workers to extract their profile attributes and match them with the task properties (Difallah et al. 2013). Personalization systems may also affect users' privacy, such as a recommendation system based on the data captured from the social relationships (Guy et al., 2009). Although obtains good accuracy, personalization systems using social networks data may not comply with current workplace privacy policies (e.g., Gorm and Shklovski 2016).

Whereas most of past studies in line with worker-to-task assignment in crowdsourcing settings have focused in skill expertise and prior records of task execution (e.g., task accomplishment ratio), a solution for task assignment that has potential to be better comparing to the previous approaches is based on the assessment of the worker's cognitive profile. Cognition can be explained as a group of capacities and mental processes which are used for the fulfillment of a goal (Miller and Wallis 2009; Ramsey 2017). Furthermore, cognition comprises the mental structures involving perception, attention, thinking and reasoning, learning, memory, and communication (Montello 2009). The measurement of cognitive ability is known as a good predictor of the work performance (F. L. Schmidt and Hunter 2004; Schmitt 2014). As will be elaborated

upon in the following sections, the measurement of cognitive abilities for task assignment purposes can have additional benefits such as higher levels of accuracy (Germine et al. 2012), shorter measurement times (Danula et al., 2019), and even the ability to be applicable to different types of tasks (Hettiachchi et al. 2020). For instance, Eickhoff (2018) argued that task assignment could be used to mitigate cognitive biases and/or errors that influence thinking and judgment. With this in mind, Hettiachchi et al., (2020) proposed an online task assignment framework based on cognitive abilities. The results indicated that the proposed cognitive-based framework obtained better results than state-of-the-art task assignment methods relying on workers' previous task performance (Zheng et al. 2015). Although there is a growing tendency for studying cognitive assignment in crowdsourcing settings (Eickhoff 2018; Goncalves et al. 2017; Hettiachchi et al. 2019a, 2019b, 2020), the personalization of digital labor based on cognitive abilities is still an understudied area of research.

Regarding other similar systematic reviews on this field, one article described the state of the art on web personalization between 2005 and 2015 (Salonen and Karjaluoto, 2016). Although this article offered very interesting insights on web personalization, it had the downside of focus primarily on the marketing field and does not focus on the cognitive dimension. Nevertheless, it is discussed some user-centric issues, implementation details and theoretical foundations for the web personalization. From the perspective of personalization in the field of crowdsourcing, it can be used to help people with autism, through the correct annotation of PoIs (Points of Interest), which in addition to meeting the preferences of individuals, also takes into account their idiosyncratic aversions (Cena et al., 2018; Mauro et al., 2020). Personalization takes place through a recommendation model that integrates different measurement criteria into the balanced recommendation of PoIs for the users taking into account their interests and compatibility. Either analyzing solutions which may help crowd workers with cognitive disorders performing microtasks or even improving the microtask assignment on neurotypical crowd workers, there is a necessity of examine the state-of-the-art on this theme. The state-of-the-art can present personalization solutions that translate into an improvement in the quality of the work performed and additionally in an improvement in the satisfaction of the crowd workers. The purpose of this study is to identify opportunities, gaps, and future research paths by conducting a systematic literature review on examining the types of cognitive personalization proposed in the context of online microtasks labor.

2 Methodology

The main goal of this systematic literature review (SLR) is to conduct a methodical evaluation and interpretation of the research done on cognitive personalization for online labor microtasks. At this end, we believe that the present study may offer a scientific lens for looking at ways in which we can better support the tailoring of technological solutions to the needs of each individual based on the cognitive profile of the digital worker.

2.1 Research questions

Our work follows the guidelines proposed by Kitchenham (2004) for conducting and reporting literature reviews in the software engineering domain. In general terms, this method allows to evaluate, aggregate, and synthesize evidence from the available literature taking into account its relevance to understanding a phenomenon, topic area, or research problem using a systematic and rigorous approach. Before formulating the research questions (RQs) for a systematic literature review, Kitchenham (2004) recommends first defining a set of structural elements embedded into the PICO (Population, Intervention, Comparison, and Outcomes) strategy. The approach taken to planning the present SLR is detailed below, along with specificities of the PICO format:

- Population—workers on digital work platforms;
- Intervention—cognitive personalization in the online microtasks labor platforms;
- Comparison—other task assignment frameworks for online microtasks or absence of personalization;
- Outcomes—usability, accuracy, and efficiency regarding the online tasks performed.

With the move toward seeking solutions for enhancing worker-to-task matching models, this SLR covers an 11 year period (2010–2020) of research on cognitive personalization for online labor platforms. This range was selected based on the fast pace of web technology development such as the specification of HTML 5 (Anthes 2012), which may cause the older solutions to be less effective and sometimes outdated. Online microtasks labor was defined as a criterion for identifying studies conducted in online platforms that dispatch crowdsourcing and related digital work arrangements (e.g., microwork). We choose to focus on this object of study in order to exclude other types of crowdsourcing occurring offline like transportation or food delivery, where it would be more difficult to evaluate the presence of personalization strategies with accuracy. Although there is a wide diversity of online labor scenarios such as crowd sensing (Muller et al. 2015), where data collection is achieved using smartphone sensors, we will focus on online labor conducted in virtual workspaces. To this end, the research studies considered for scrutiny comprise an evaluation of the cognitive personalization in terms of usability and effect on the work performance. Based on the PICO strategy, the RQs that guide this article to accomplish the goal of this SLR are presented in Table 1.

2.2 Search process

The search process is focused on primary studies from search engines (i.e., Google Scholar, Scopus) and digital libraries (i.e., ACM Digital Library, IEEE Xplore, PubMed). Most of these electronic sources have been suggested as relevant in the field of software engineering (Brereton et al. 2007). In the pursuit of records in the health domain, PubMed was used with the purpose of gathering insights on cognition studies. The search string was then constructed taking into consideration the defined PICO structure and the following steps:

1. Analyze the RQs to identify search terms;

Table 1 Formulation of the research questions

ID	Research question (RQ)	Motivation
RQ1	How cognitive personalization on online labor microtasks platforms affects the task performance of workers?	Examine the effects caused on the task performance
RQ2	What are the personalization methods used in each primary study?	Explore the methods applied to create the personalization. Identify the programming libraries or frameworks applied for personalization, which can also include artificial intelligence techniques for mapping the cognitive abilities of workers
RQ3	What types of evaluation were used to assess the validity of these methods?	Provide information of the techniques used to evaluate the cognitive personalization methods
RQ3.1	What evaluation techniques were adopted?	Verify if the results obtained from the methods used indicate that personalization has better, similar, or worse results in comparison with other frameworks, task assignment or personalization methods
RQ3.2	Did real users evaluate these methods?	Explore if either a real user (i.e., online worker) or expert evaluated these methods

2. Identify search terms in relevant papers;
3. Identification of synonyms and alternative spellings for the search terms;
4. Construction of search strings using Boolean Operators (i.e., AND and OR).

At a higher level, the search string presented in Table 2 was adapted to each electronic source and formulated in accordance with its main scope. After collecting the primary studies eligible for this SLR, we also used snowballing techniques to obtain more potentially relevant studies (Wohlin 2014). To this end, we screened the reference

Table 2 Search string used in the primary search phase

Scope	String
Work done in online microtasks labor platforms	("crowd work" OR "crowdwork" OR "microwork" OR "online labor" OR "online labor" OR "online outsourcing" OR "micro sourcing" OR "microsourcing" OR "elancing" OR "e-lancing" OR "gig work" OR "gig-work" OR "on-demand work" OR "digital labor" OR "digital labor" OR "platform labor" OR "platform labor" OR "digital work" OR "task assignment") AND
Cognition method	("cognitive styles" OR "cognitive abilities" OR "cognitive functions" OR "cognitive levels") AND
Personalization or customization	("personalization" OR "adaptive" OR "customization" OR "adaptable")

lists of studies identified through database searches in order to increase the number of studies available in the literature that investigate the phenomenon under study.

2.3 Inclusion criteria

The inclusion criteria (IC) comprehend the RQs defined previously and involve studies published in or after 2010 and before 2021. This date range has the purpose of characterizing the current state-of-the-art approaches for cognitive personalization in digital labor settings. The inclusion criteria followed in this study are presented as follows:

1. Papers that include cognitive personalization applied on online microtasks labor platforms;
2. Papers about personalization, customization or work assignment;
3. Studies published between 2010 and 2020.

2.4 Exclusion criteria

The exclusion criteria (EC) aim to screen the studies collected and then, remove the duplicates along with studies not published in English and that were not peer-reviewed. In order to achieve this, the following exclusion criteria were defined:

1. Duplicate reports of the same study (when several reports appeared, it was only accepted the most completed version);
2. Not written in English;
3. Full paper not available;

4. Editorials, keynotes, abstracts, tutorials, dissertations or theses;

5. Not related to the topics addressed in the research questions;

6. Not peer-reviewed;

7. Studies without an empirical validation or experimental results; and
8. Studies published before January 1, 2010 and after December 31, 2020.

2.5 Search strategy

Regarding the search strategy used to identify primary studies, the articles were included only if the following assumptions were met:

1. Proposal of a system that allows the cognitive personalization in association with online microtasks labor; and
2. Explanation of the evaluation method and interaction mechanism for the first assumption.

2.6 Data collection

The data collection form used in this study for extracting insights from each included paper consists in the following elements:

- Source of publication (conference, journal);
- Year;
- Authors;
- Research question/purpose of the study;
- Technologies used (i.e., tools used to construct the solution);
- Evaluation methods (e.g., usability testing);
- Target population (e.g., cognitive impaired people);
- Main findings;
- Summary of the paper; and
- Additional notes (complementary observations on the study);

2.7 Conducting the review

As depicted in Fig. 1, the proposed flow diagram of the search process also comprised a snowballing strategy that was used to obtain additional relevant studies from the final list of articles. There were a total of twelve articles identified after applying the inclusion and exclusion criteria.
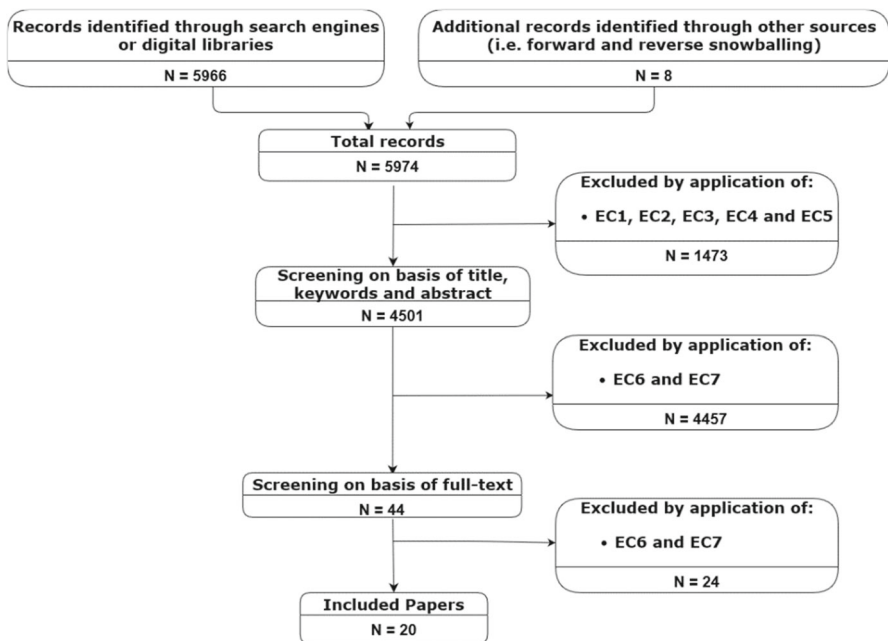

Fig. 1 Flowchart of the paper selection process

Therefore, eight more articles were selected after screening the citations (forward snowballing) and the reference lists (reverse snowballing) of each selected study, which in the first iteration were twelve. In addition, the flowchart of the paper selection process also presents a view of the records excluded taking into account the exclusion criteria defined in Sect. 2.4. Synthetizing and Analyzing. As a next step of this SLR, the results were analyzed and synthetized in order to answer the RQs formulated in Sect. 2.1. For performing correlations and thus find common aspects among the included studies, we explored several trends and common patterns in online microtasks labor based on the papers found, authors' expertise in the field, and validity of these aspects as presented in related work. From this process, we were able to compile the following list of key themes: nature of collaboration, evaluation of cognitive features, platforms, adaptation, and user testing.

2.7.1 Nature of collaboration

Looking at the found aspects in general, most studies addressed cognitive abilities considering the existence of collaboration between workers. Digital collaboration is defined as an experience that integrates people, processes and technology (Morabito 2014). Collaboration in digital work platforms can be analyzed from the perspective of collective intelligence systems with fundamentally two purposes in mind: i) to create something new, or ii) to decide on existing information (Malone et al. 2009). Additionally, online workers can collaborate independently or interdependently. Starting from the crowdsourcing systems scenario, Doan et al., (2011) have a similar view to that of Malone and co-authors (Malone et al. 2009) in the sense that collaboration can be grouped in two characteristics:

- Explicit—Online workers have the perception that they are collaborating with other people when carrying out tasks and their individual outcomes can be influenced by other responses (Huang and Sundar 2020; Koutrika et al. 2009).
- Implicit—The result of the task is the joint effort of multiple workers without having a clear perception of collaboration with other users. An example of implicit collaboration is the creation of a crowdsourcing campaign for annotating images that aggregates the results made individually by crowd workers (Müller et al. 2019).

Consequently, the studies found were grouped according to the explicit and implicit nature of collaboration proposed in Doan et al., (2011).

2.7.2 Cognitive features

At first sight, the International Classification of Functioning, Disability and Health (ICF) is a framework commonly used for the classification of health and disability (World Health Organization 2001). The SLR conducted by Gillespie et al., (2011) identified several concepts of ICF applied to cognitive functions, including attention, memory, perceptual, thought, higher level cognitive, calculation, mental or complex movement, experience of self, and time functions. Cognitive functions, also designated as cognitive abilities, are described as the mental capabilities used for learning and solving problems (Stanek and Ones 2018). Several cognitive dimensions have been

used for examining the impact that they can have on an individual's thinking. Cognitive style is an individual's characteristic that influence the perception and management of information in a systematic way (Littlemore 2001). There exist a main distinction between cognitive abilities and cognitive styles, especially when considering task performance (Riding 1997). In cognitive abilities, the task performance increases in a linear relation to the individual's abilities, while in cognitive styles, the task performance increases or decreases taking into account the presentation or type of task given to the individual (Riding 1997). Cognitive bias is defined as the cognition of individuals which regularly produces representations that are systematically inaccurate when compared to reality (Haselton et al. 2015). The cognitive engagement refers to the individual persistence and time spent when accomplishing a task (Rotgans and Schmidt 2011). As a result of this process, the selected studies were grouped based on the methods used for evaluating the previous cognitive features.

2.7.3 Adaptation

Adaptation of technology refers to matching the system attributes to the user abilities (Ahmi and Mohamad 2016), which can be done through personalization or customization. A definition for online personalization relies in delivering "the right content to the right person in the right format at the right time" (Ho and Tam 2005). Customization is categorized as providing the choices to the user for tailoring the content according to his/her preferences, giving a greater importance to the user control (Sundar 2008). The main difference between personalization and customization is that the adaptation of the former is system-driven, while the latter is user-driven. Both approaches can result on adaptation of task design and assignment. Task design refers primarily to how the interface and content are presented to the user. Furthermore, task design in digital work settings can affect the quality of results (Finnerty et al. 2013). On the other hand, task assignment consists in the optimized allocation of tasks to the online worker's abilities. The proper assignment of tasks in online labor platforms can upkeep the workers motivated and also save time and money for task requesters (Bhatti et al. 2020).

Furthermore, interactive adaptive systems have formative methods from which can be analyzed the digital work settings (Paramythis et al. 2010). The collection of input data considers different metrics such as accuracy, latency, or the sampling rate. Collected data must be interpreted under the light of the validity of interpretations (refers to whether the inferences/interpretations reflect the actual state of the entity being modeled), predictability (refers to whether users are capable of predicting the system's modeling behavior, given the system's interpretation of their actions) and scrutiny (refers to the users' capacity to inspect and modify the user model itself). Then, interactive adaptive systems must be examined on the determination upon adaptation decisions (necessity of adaptation, appropriateness of adaptation, predictability) and the application of these adaptation decisions (usability criteria, timeliness, and acceptance by user).

2.7.4 Platforms

The type of platform presented in the sample chosen for analysis is directly related to the nature of collaboration. Explicit collaboration emphasizes the support for communication between workers (e.g., videocall or messaging platforms). For aiding implicit collaboration, the platforms only need to provide the interface for performing the digital labor tasks (e.g., crowdsourcing platforms), where the worker is unaware of collaborating with other people.

2.7.5 User testing

On digital labor platforms, usability has a fundamental role and constitutes an inherent design aspect which can affect substantially the work performance (Zyskowski et al. 2015a). Usability evaluation has been incorporated in the development of websites, and user testing is considered one of the most popular techniques (Maguire and Isherwood 2018). In this respect, user testing consists primarily on the experience of users and is frequently conducted in a scenario-based setting (Tan et al. 2009) which can be then categorized as laboratory or remote testing. That is, user testing can be conducted in a laboratory, usually in a designated test room where users perform specific tasks in individual test sessions (Bastien 2010). On the other hand, a remote user testing implies a condition where the evaluators are disconnected in time and/or space from users (Andreasen et al. 2007). Additionally, to usability, there are other variables that can be applied in the context of user testing such as the appropriateness of adaptation, the user behavior, or user performance (Van Velsen et al. 2008). Appropriateness of adaptation refers to a correct adaptation having in consideration the requirements of the current interaction (Paramythis et al. 2001). The identification of user behavior can increase the comprehension of why an action was performed and supports a better personalization of the system. User performance reports with comparison of personalized and non-personalized systems. The methods used in user testing can be used to provide triangulated data, and they could also collect subjective feedback when evaluating users. The methods frequently used are questionnaires, data log analysis and think-aloud protocols.

3 Results

In the following subsections, we present the results of our systematic review (see Table 3 for an integrated view on the results grouped by the aspects of online digital work defined in the previous section). On a general level, the results allow to characterize the research conducted on cognitive personalization through a critical lens that comprises five different dimensions.

Table 3 Overview of the results obtained in this systematic literature review

Reference	Collaboration		Cognitive features				Adaptation of task		Platforms	User Testing
	Implicit	Explicit	Abilities	Bias	Engage	Style	Assignment	Design		
(Goncalves et al. 2017)	×		×				×		–	Laboratory
(Sakurai et al. 2010)		×	×					×	Wonderland	
(Stewart et al. 2019)		×	×				×		Zoom	
(Engin and Vetschera 2017)	×		×			×		×	–	
(Alagarai Sampath et al. 2014)	×		×					×	MTurk	Remote
(Eickhoff 2018)	×	×		×			×		MTurk	
(Hettiachchi et al. 2019)	×		×				×		MTurk	
(Hettiachchi et al. 2020)	×		×				×		MTurk	
(Evangelos Mourelatos and Tzagarakis 2016)	×		×					×	microWorkers	
(Mourelatos and Giannakopoulos 2020)	×		×					×	microWorkers	
(Ravana et al. 2018)	×		×				×		CrowdFlower	
(Kosinski et al. 2012)	×		×		×			×	MTurk	
(Gadiraju et al. 2017)	×			×			×		CrowdFlower	
(Feldman and Bernstein 2014)	×		×				×		eLance	
(Hassan and Curry 2013)	×		×				×		ShortTask	
(Erskine et al. 2019)	×		×					×	GISCloud	
(Davier and v., Hao, J., Liu, L., Kyllonen, P. , 2017)		×	×				×		MTurk	

Table 3 (continued)

Reference	Collaboration		Cognitive features				Adaptation of task		Platforms	User Testing
	Implicit	Explicit	Abilities	Bias	Engage	Style	Assignment	Design		
(Saab et al. 2019)	×			×				×	–	
(Ponciano and Brasileiro 2015)	×				×			×	–	
(Chujfi and Meinel 2020)		×				×	×		–	

3.1 Nature of collaboration

Starting from the typologies presented in the literature (e.g., Doan et al., 2011; Malone et al. 2009), the nature of collaboration refers to the collaboration that occurs in digital work settings explicitly (when workers have a clear perception of collaborating with others) or implicitly (the result of the digital work performed is the aggregation of all individual contributions). Among the 20 studies selected for this review, 15 articles involved implicit collaboration, while only 5 articles addressed explicit collaboration. A total of 86.67% of the studies involving implicit collaboration were based on crowd-sourcing with an emphasis on microtasks. On the other hand, only 20 percent of the studies addressing explicit collaboration were related to crowd work. These numbers are aligned with previous findings from the literature (e.g., Ghezzi et al., (2018)) where there has been pointed a steadily increase in crowdsourcing research since the inception of the term in 2006 (Wazny 2017), and since then it has created a new form of digital work. Furthermore, the high number of studies focused on microtask crowd-sourcing by means of implicit collaboration is explained by the fact that microtasks are frequently designed to be executed by workers in a short amount of time, which does not foster the establishment of explicit collaboration, as noted in Zyskowski et al. (2015a).

3.2 Cognitive *features*

In this section, we describe the methods used for the evaluation of cognitive features found in the selected studies with a focus on the optimization of online labor to the cognitive aptitudes of workers.

3.2.1 Cognitive abilities

Cognitive abilities can be described as the mental capabilities used for learning and solving problems (Stanek and Ones 2018). Next, we present the methods for evaluating cognitive abilities as found in this systematic literature review.

Kit of factor-referenced cognitive tests   The Kit of Factor-Referenced Cognitive Tests (Ekstrom et al. 1976) is a manual of pencil-and-paper tests for the identification of 23 aptitude factors through 72 cognitive tests. This kit of cognitive tests was published as a landmark instrument in 1976 and has been subject of validation and reliability over the years (Herreen and Zajac 2018; Schaie et al. 1991). Furthermore, these cognitive tests have been applied in several domains such as mental disease research (Crucian et al. 2010; M. Müller et al. 2013), decision-making support (Fallon et al. 2014; Finucane and Gullion 2010), or even technology adoption (Mitzner et al. 2019; L. I. Schmidt and Wahl 2018; S. Zhang et al. 2017). For instance, Goncalves and co-authors (Goncalves et al. 2017) used the Kit of Factor-Referenced Cognitive Tests in a laboratory setting for measuring fluency and visual-oriented cognitive abilities across different types of tasks (e.g., item classification, and text distortion). Accordingly, the authors of the study noted that although the analysis of the capacities can achieve good
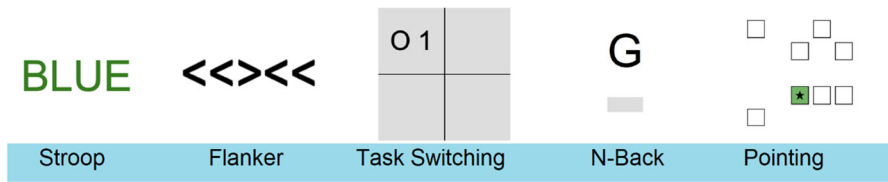
Fig. 2 Examples of cognitive tests proposed in Hettiachchi et al., (2020)

results, the tests are very time consuming to be used in a crowdsourcing environment. Furthermore, another similar study examined a set of visual cognitive abilities in an online setting based on 6 cognitive tests (Feldman and Bernstein 2014).

In summary, the results indicated that the estimated time required by the participants to perform the tasks would be 80 min. A further study carried out by Ravana et al., (2018) analyzed the logical reasoning of crowd workers through an online task. In this experiment, crowd workers had to infer the logical reasoning of a series of sentences. This led to describe the positive relationship between logical reasoning and the quality of crowd work. However, the average time required for performing the tests was not reported.

Microtasks Consistent with previous experiments, two related studies (Danula Hettiachchi et al. 2019a, b; Hettiachchi et al. 2020) used several microtasks for assessing the capabilities of crowd workers in online digital work settings. The main purpose of these studies was to improve task assignment in crowdsourcing based on the cognitive abilities of workers. At a glance, these studies assessed three different cognitive abilities (i.e., inhibition control, cognitive flexibility, and working memory) by using five cognitive tests (Stroop (MacLeod 1991), Flanker (Eriksen and Eriksen 1974), Task Switching (Monsell 2003), N-Back (Owen et al. 2005), and Self-ordered Pointing (Petrides et al. 1993). Figure 2 depicts an example of the cognitive tests performed. Before each test, it was included a set of instructions and an example to ensure that the workers understood the experiment. Besides the Self-ordered Pointing Test, each test had an expiration time of 3.5 s. This was applied to confirm that crowd workers did not pause their activity during a test. Results indicate a significant increase in terms of performance when applying these short online tests to support task assignment in crowdsourcing settings.

Operation Span (OSPAN) Task has been widely used for assessing the working memory capacity. With this framing, the tasks comprise solving arithmetic challenges while memorizing unrelated words (Turner and Engle 1989). In this regard, Graf et al., (2006) created an online version of the OSPAN Task (Web-OSPAN) which register several metrics such as response latency, efficiency on calculations, and number of correct words. OSPAN Task consists in 60 simple arithmetic problems to solve, each one including a word to memorize. A research study on how to enhance collaboration in online settings used the Web-OSPAN to measure the working memory capacity of collaborators (Sakurai et al. 2010). The findings suggested that the information presented could be dynamically adjusted by taking into account the working memory capacity.

In a similar vein, intelligence tests are frequently used to predict an individual performance in a job context (Murtza et al. 2020; Nguyen et al. 2019). Furthermore, these tests can be performed to evaluate intelligence aspects in a crowdsourcing platform. Kosinski et al., (2012) proposed a different approach for measuring the intelligence of crowd workers through a holistic perspective instead of individual assessments. The method used consisted in splitting an IQ questionnaire based on the Raven's Standard Progressive Matrices (Raven 2000), converting each question into a crowdsourcing task. Therefore, the intelligence of the crowd was used to analyze several elements such as the effect of worker reputation, payment, and aggregation of results. At this level, an important issue that arises is the fact that crowdsourcing performance can have different outcomes based on how tasks are designed. In (Alagarai Sampath et al. 2014), a cognitive-inspired task design was proposed to increase the performance of crowd workers. Several experiments were then conducted to examine different cognitive parameters on crowd workers performing a form-digitization task. Through a set of experiments on text-extraction tasks, the authors demonstrated that highlighting the text fields is essential for the tasks associated with visual attention. On the other hand, moving the answer text box near to the target fields can effectively increase the crowdsourcing performance when executing tasks involving working memory.

Analysis of transcripts   In practice, we can identify collaboration processes by examining the spoken language of online workers. Stewart et al., (2019) proposed a model for the automatic identification of collaborative problem-solving (CPS) skills. The study consisted in a manual analysis of transcripts to detect collaboration skills using video conferencing and an online collaborative task. The maximum time to accomplish the task was 20 min. As a result, an automated support mechanism for this process was then proposed with suggestive evidence for constructing intelligent collaborative interfaces and supporting the guidance of collaboration task activities.

Historical records of crowd worker's performance   The performance of crowd workers taking into consideration their previous tasks can be used as a means to assess their cognitive abilities. Hassan and Curry (2013) examined task assignment in crowdsourcing settings based on the historical crowd worker's performance. The study involved several steps which started from modeling tasks based on human abilities through a validated taxonomy (Edwin A Fleishman 1975; Edwin A. Fleishman et al. 1999). Afterward, the authors proceeded to capture the ability traces of crowd workers in order to better predict their performance. Some of the human abilities identified are considered as cognitive abilities (e.g., comprehension and reasoning). The results obtained show that a capability prediction strategy based on such traces are similar to the baseline metrics used based on task accuracy.

Questionnaires   Questionnaires are often used to evaluate the cognitive abilities of digital workers. The Geospatial Reasoning Ability (GRA) supports individuals on evaluating geospatial information (e.g., interactive web maps) to make accurate decisions (Jarupathirun and Zahedi 2007). With this in mind, a questionnaire was then created and validated for the measurement of GRA (M. A. Erskine et al. 2015). This

GRA questionnaire was used by M. Erskine et al., (2019) to study user and task characteristics on spatial decision support systems. The results indicate that GRA has a significant effect on decision-making performance. Similarly, the Scientific Literacy Measurement (SliM) is a questionnaire proposed for evaluating civic scientific literacy based on scientific keywords and similar data extracted from educational textbooks and newspapers (Rundgren et al. 2012). In (Davier et al. 2017), and the SliM questionnaire was used as a measurement of cognitive abilities for the evaluation of CPS skills through a web collaborative science evaluation prototype. As we look into the possibilities of measuring the cognitive abilities of digital workers, the Cognitive Reflection Test (CRT) also appears as a valid instrument that uses a three-item task for assessing the reasoning of the individual when seeking for an unintuitive correct answer (Frederick 2005). Specifically, CRT is considered a quick and easy test to conduct with a moderated positive association with cognitive abilities (Oechssler et al. 2009; Toplak et al. 2011). A study on the effects of information representation on decision-making used the CRT for the measurement of cognitive abilities (Engin and Vetschera 2017). In its general form, short demographic items can also be used to assess cognitive abilities. Two related studies on crowd work (E Mourelatos et al., 2020; Evangelos Mourelatos and Tzagarakis 2016) measured cognitive abilities through the information of participants in terms of education level and computer skills, which obtained a positive relation with the performance of workers. Furthermore, the education level is positively correlated with the personality trait of extraversion (E Mourelatos et al., 2020).

3.2.2 Cognitive styles

As mentioned before, a cognitive style is an individual characteristic that influences systematically the perception and management of information (Littlemore 2001). The Cognitive Style Index (CSI) questionnaire measures the individual differences when managing information, specifically in work settings (Allinson and Hayes 1996). Furthermore, the questionnaire identifies characteristics that belong between the abilities and personality domains. Further experimentation in (Engin and Vetschera 2017) examined the effect of different graphical settings in solving ranking problems. Accordingly, the impact of cognitive styles on the decision-making process was evaluated using the CSI questionnaire. The authors pointed out to the importance of information representation to match not only the task attributes but also the cognitive style of the user. A somewhat similar body of work has claimed for the use of online self-assessment questionnaires to identify cognitive styles. In view of this, Chujfi and Meinel (2020) created a questionnaire to identify the cognitive preferences of teleworkers. The questionnaire was constructed using the Sternberg's thinking style methodology (Sternberg 1997) which classifies the functions, forms, levels, scopes, and leanings of governance in individual cognitive preferences. Moreover, the research explored how organizations that provide support for digital work can adjust the cognitive abilities of workers to optimize the task assignment from a collective intelligence perspective. The study concluded that the patterns of self-organization are linked positively with matching each stage of knowledge management to the individual's cognitive style (Chujfi and Meinel 2020).

3.2.3 Cognitive bias

In a broad sense, cognitive bias is understood as the cognition of individuals which regularly produces representations that are systematically inaccurate when compared to reality (Haselton et al. 2015). Concerning the self-assessment confidence of the worker, the Dunning–Kruger effect describes a cognitive bias whereby individuals with less abilities have an optimistic outlook that does not correspond to the reality about their own abilities (Kruger and Dunning 1999). Furthermore, this cognitive bias causes individuals to make mistakes without being aware of it. Additionally, high-ability individuals can also suffer from a cognitive bias when they undervalue their abilities. Prior studies have examined the Dunning–Kruger effect on crowdsourcing participants to self-assess their confidence levels regarding task performance. From this point, Saab et al., (2019) constructed a model for evaluating the Dunning–Kruger effect on the crowd taking into perspective different aggregation methods. The results were based on an existing dataset of volunteer crowdsourcing on quiz answering (Aydin et al. 2017). This study's findings reported that the plurality voting aggregation method obtained better results when compared to confidence assessment approaches. Finally, the researchers proposed a competence-weighted approach based on the confidence of crowd workers that outperforms most of the baseline aggregation methods. In connection with this aspect, Gadiraju et al., (2017) investigated the Dunning–Kruger effect on crowd workers by conducting two experiments which evaluated if the crowd had accurate self-assessment answers and if showing the results of other workers influence their performance. In addition, two more experiments evaluated if self-assessment could be a good predictor of worker competence. The main findings were that the self-assessment of cognitive bias affected primarily less competent workers in easier tasks. However, according to the Dunning–Kruger effect (Kruger and Dunning 1999), competent workers were also affected when the difficulty was higher than their abilities. In summary, self-assessment was considered an integral component of worker competence and including self-assessment as a pre-screening strategy can significantly improve the overall results.

When we look at task design in microwork scenarios, a body of work has studied the value of changes in the layout of microtasks through the lens of cognitive bias (Eickhoff 2018). In this line of research, the purpose is to study the prevalence of cognitive bias in crowd workers, specifically in the case of document relevance assessment tasks. Some of the studied cognitive biases are:

- Ambiguity Effect—The lack of information makes the decision-making process seem more difficult (Ellsberg 1961);
- Anchoring occurs when individuals focus excessively on a specific piece of information (often the first one they observe) disregarding additional contradictory evidence (Tversky and Kahneman 1974);
- Bandwagon Effect—The presentation of an existing group of results can influence one individual to follow the group behavior (Bikhchandani et al., 1992). Figure 3 presents the interface proposed by Eickhoff (2018) to evaluate the Bandwagon effect; and

Description:

Document mentions a leveraged buyout valued at or above 200 million dollars.

Title:

Stop Reject Latest Dart Offer

Content:

Dart Group Corp. says it will press billions-dollar buyot bid for Stop andamp; Shop Cos. despite a unanimousrejection of the offer by the Stop andamp; Shop board of directors (...)

Relevance:

○ Relevant (7/10 people think so)
○ Not Relevant (3/10 people think so)

Fig. 3 Bandwagon effect task design which shows results of other crowd workers. Adapted from Eickhoff (2018)

- Decoy Effect—Overall, this effect is related to the preferences of each individual and happens when between choosing options A and B individuals will choose B when a third option C is presented and is obviously inferior to option B (Huber et al., 1982).

Going deeper into the issues regarding the Bandwagon effect, Eickhoff (2018) created several tasks with smooth changes in the layout in order to examine the effects of cognitive bias. The study compared each cognitive bias with a baseline (simple design with no cognitive bias effect) and the results were then compared with those from expert annotators. From these experiments, it was possible to observe a significant decrease in the quality of work when the task is designed without taking into account the cognitive bias of a crowd worker.

3.2.4 Cognitive*engagement*

When considered from a motivational viewpoint, cognitive engagement can indicate the state of an individual when he or she is motivated to perform a task. In a crowdsourcing study, Ponciano and Brasileiro (2015) examined cognitive engagement patterns based on two datasets from validated citizen science projects (i.e., Lintott et al. 2011; Simpson et al. 2012). The authors defined engagement as a metric for measuring the duration and number of times that a volunteer contributed to a citizen science project. Consequently, volunteers classified as persistent (low levels of engagement but working for a longer time) had the higher percentage of contribution. Moreover, Kosinski and co-authors (Kosinski et al. 2012) observed that the rewards given during a crowdsourcing campaign may have a significant impact on the engagement behavior of crowd workers. Among the aspects identified, a highlight of this study was the fact that a crowd worker can feel a psychological pressure with adverse repercussions on his/her cognitive skills if the reward received is too high.

3.3 Adaptation

Viewed through a lens of cognitive aspects, subsequent adaptation can be made outside or inside the scope of an online task in a digital work environment. The former is based on task assignment toward matching tasks to the most appropriate workers, while the latter is related to task design as presented to the worker.

3.3.1 Task assignment

Pre-screening in crowdsourcing is a popular strategy to filter out unsuitable crowd workers participating in online tasks (Oleson et al. 2011). At a higher level, the pre-screening methods consist in the performance of crowd workers when executing prototypical microtasks. If the pre-screening results are satisfactory, then, a worker can perform the actual task. Gadiraju et al., (2017) studied the cognitive bias regarding the self-assessment of crowd workers in order to complement the pre-screening methods. Self-assessment was considered an integral component of the crowd worker's competence. Among other findings, the study found that a pre-screening strategy complemented with self-assessment could significantly improve the performance of workers. In light of these findings, it is suggested that the cognitive characteristics of each worker can be used for enhancing pre-screening methods. A study based on crowdsourcing relevance judgment indicated that logical reasoning is related with the quality of the crowd worker's inputs (Ravana et al. 2018). Additionally, other factors such as English proficiency and education level can also have a positive relation with the performance of crowd workers (E Mourelatos et al., 2020; Evangelos Mourelatos and Tzagarakis 2016).

Going beyond pre-screening, task assignment can also involve the posterior matching of each worker's abilities to the task characteristics. Using the identification of workers' cognitive abilities as a basis for task assignment in crowdsourcing, two preliminary studies (Feldman and Bernstein 2014; Goncalves et al. 2017) revealed that the application of the Kit of Factor-Referenced Cognitive Tests had a strong correlation with the performance of crowd workers. Following these preliminary insights, a study (Hettiachchi et al. 2020) was conducted in a real crowdsourcing setting where workers had to perform cognitive evaluation tests in the form of microtasks. In a subsequent paper (Difallah et al. 2013), crowd workers performed microtasks selected to be representative of the typical tasks available in crowd work and the results were compared with several state-of-the-art task assignment methods (e.g., (Zheng et al. 2015)). The main findings indicated that short-length cognitive tasks help to achieve better results regarding the crowd worker assignment when comparing to other task assignment methods. Furthermore, improving the assignment of tasks to workers is validated in the literature as promoting a higher worker satisfaction (Edwards 1991).

In the studies of task assignment, we examined several interactive adaptive features. In the case of the input data collected, it was analyzed the latency in the adaptation process. For example, in the study by Goncalves et al., (2017), it was estimated that pencil-and-paper tests to measure cognitive abilities would take at least 60 min, so even the authors themselves considered that it was not feasible to perform online adaptation with this delay. However, this study served as the basis for further studies,

adapting pencil-and-paper tests to short microtasks online, where in less than 10 min, the executive functions of crowd workers could be assessed and then, the subsequent task assignment adapted to the crowd workers (Hettiachchi et al. 2020). In another direction, (Ravana et al. 2018) perform an analysis on the work made by crowd workers, by excluding tasks performed below than two minutes, which should indicate that the crowd worker did not had enough time to perform the microtask appropriately. Although this screening can be easy to implement, it could however exclude efficient crowd workers and not properly assess the adaptation perform in the task assignment.

Regarding the validity of interpretations on the collected data, Goncalves et al., (2017) and Hettiachchi et al., (2020) based their studies on standard cognitive tests that have been thoroughly tested their validity and reliability in the scientific literature, such as the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al. 1976; Herreen and Zajac 2018). This provides a solid basis to then provide accurate task assignment. Moreover, Hettiachchi et al., (2020) validated the task assignments by comparing the accuracy of the work performed with a baseline (neither task assignment nor task recommendation was performed). Their conduction of several studies indicated the usage of self-assessment measures by crowd workers was a good predictor of the worker competence and thus being applied to better task assignment. The validity of the self-assessment measures was performing by correlating the microtasks effectiveness and efficiency. Another feature of interactive adaptive systems is predictability, which is based on user predicting the system's modeling behavior; Gadiraju et al., (2017) discuss about the predictability of crowd workers on what would be their microtask performance, stating that self-assessment cognitive bias affected primarily less competent workers in easier tasks, but also competent workers when the difficulty was higher than their abilities. The interactive adaptive systems must be examined on the determination upon adaptation decisions the application of these adaptation decisions, Hettiachchi et al., (2020) provided in one case study an option to the crowd workers to choose the preferred microtasks and even discard entirely the suggestions provided by the system.

3.3.2 Task design

Some work has drawn attention to task design in crowdsourcing for facilitating a better user experience (Alagarai Sampath et al. 2014; Eickhoff 2018; M. Erskine et al. 2019; Sakurai et al. 2010). Basing their approach on the use of a collaborative virtual world, Sakurai et al., (2010) introduced a method for content personalization that takes into account the user context and the cognitive profile of each individual as assessed through the Web-OSPAN test (Graf et al. 2006). The results indicated that the enriched collaborative environment could reduce the misunderstandings between users in remote collaboration. However, in this study, the cognitive profile of users was assessed separately in a laboratory setting. Therefore, M. Erskine et al., (2019) conducted a study focused on spatial decision support with the main goal of determining if the user and task characteristics can be enhanced for obtaining a better performance in decision-making. Among the characteristics examined, the authors explored the Geospatial Reasoning Ability (M. A. Erskine et al. 2015) and the effects of problem

complexity and perceived task-technology fit on decision performance. It was highlighted that a suitable visualization and reduced problem complexity can enhance the performance of workers in geospatial tasks. Similarly, Engin and Vetschera (2017) performed an experiment to evaluate the effect of information representation concerning the relation of cognitive styles and decision-making performance. The results underlined the negative effects that a mismatch between information representation, task aspects and cognitive styles will cause in solving problems. Another set of significant contributions regards the prevalence of cognitive bias in crowd workers, specifically in the case of document relevance assessment tasks (Eickhoff 2018). As aforementioned, Alagarai Sampath et al., (2014) studied the improvement of crowdsourcing task design from a cognitive standpoint with an emphasis on text-transcription tasks. During this investigation, a set of experiments were undertaken to analyze cognitive parameters from crowd workers in terms of working memory requisites and visual saliency of the intended fields. Others have examined the characteristics that can lead to a successful task design for online workers. Specifically, Stewart et al., (2019) analyzed a set of transcripts from video meetings to model the processes of collaborative problem-solving taking into account the skills required.

Interactive adaptive features could also be analyzed in the selected task design studies. Collected data were interpreted under the light of the validity of interpretations, using several techniques. In Alagarai Sampath et al., (2014), it was used eye-tracking data to compare the cognitive load on different inputs fields presented in the microtasks. This helped the validation of some input fields to optimize the microtasks. In Eickhoff (2018), the validity of the inferences made on crowd workers was based on several psychology cognitive bias features (Ambiguity, Anchoring, Bandwagon and the Decoy effects). Predictability was also examined in one study of task design. The key aspect in the proposed method of Sakurai et al., (2010), is when it is considered a misunderstanding in the user situation, so it would adapt an avatar in the proposed cyberspace world. As this system had in their adaptability mechanism the usage of web camera feed, the user could predict and adjust the adaptations performed by the system.

3.4 Platforms

Looking at the platforms described in the articles selected for this SLR, our findings demonstrate that they differ depending on the nature of collaboration. We observe that most of the platforms involving implicit collaboration were related to crowdsourcing, while the platforms supporting explicit collaboration were mostly used as a communication channel between workers, as it can be perceived from Fig. 4. In two of the included studies (Chujfi and Meinel 2020; Engin and Vetschera 2017), no specific platform was used.
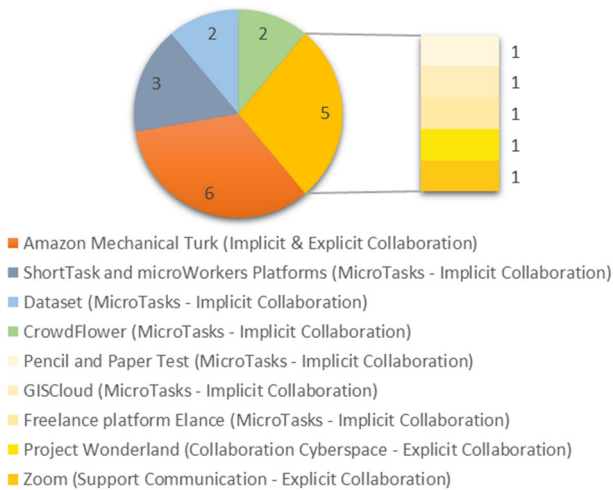
Platforms used in the selected articles

- ■ Amazon Mechanical Turk (Implicit & Explicit Collaboration)
- ■ ShortTask and microWorkers Platforms (MicroTasks - Implicit Collaboration)
- ■ Dataset (MicroTasks - Implicit Collaboration)
- ■ CrowdFlower (MicroTasks - Implicit Collaboration)
- ■ Pencil and Paper Test (MicroTasks - Implicit Collaboration)
- ■ GISCloud (MicroTasks - Implicit Collaboration)
- ■ Freelance platform Elance (MicroTasks - Implicit Collaboration)
- ■ Project Wonderland (Collaboration Cyberspace - Explicit Collaboration)
- ■ Zoom (Support Communication - Explicit Collaboration)

Fig. 4 Pie chart of the platforms used in the selected articles

3.4.1 Platforms used in implicit collaboration

From a technological point of view, Amazon Mechanical Turk (usually abbreviated as Mturk)[1] is a crowdsourcing marketplace that allows requesters (e.g., researchers) to publish tasks to be solved virtually by online workers. Furthermore, Mturk is frequently used in multiple domains such as psychology research (Cheung et al. 2017; Paolacci et al. 2010), business data collection (Keith et al., 2017), or even for the improvement of artificial intelligence techniques (Zhang et al. 2019). From the records found in this SLR with a clear focus on the use of Mturk for implicit collaboration, a total of 4 articles had as their main purpose the recruitment of crowd workers. Additionally, some studies opted to conduct experiments in their own server (Danula Hettiachchi et al. 2019a, b; Hettiachchi et al. 2020) since Mturk has a feature that allows tasks to be accomplished externally. By contrast, other researchers hosted their tasks in Mturk. From this point, Kosinski et al., (2012) focused on evaluating the effect of different task designs on cognitive abilities. At its foundational level, one disadvantage of Mturk is the difficulty of registration of crowd workers from countries outside the USA (de Winter et al. 2015).

Other crowdsourcing platforms cited in the articles found use similar features supported by Mturk but are more open to accept requesters and crowd workers worldwide. For instance, CrowdFlower[2] (now named as Appen) was used in 2 articles as the preferred crowdsourcing platform for task assignment (Gadiraju et al. 2017; Ravana et al. 2018). Concomitantly, Gadiraju et al., (2017) used a pre-screening test for their case study in order to select trustworthy crowd workers. Then, the authors adopted

---

[1] https://www.mturk.com/.

[2] https://appen.com/figure-eight-is-now-appen/.

the CrowdFlower's internal channel[3] to directly contact the selected crowd workers. Adding on to this line of research, Microworkers[4] was used in (E Mourelatos et al., 2020; Evangelos Mourelatos and Tzagarakis 2016) as a crowdsourcing platform able to support a huge number of templates and due to its easiness to adapt for different types of case study requirements. At this end, the process of decision-making can be aided with the help of this platform in order to compile a large amount of data. Thereafter, one of the included studies used GISCloud[5] as a spatial decision support system for studying the task-technology fit on web maps (M. Erskine et al. 2019). That is, on a general level, GISCloud provides a viewer with geographic features and supports crowdsourced data annotation. With respect to the study conducted by M. Erskine et al., (2019), GISCloud allowed to customize different representation types of geographic data on a web map.

3.4.2 Platforms used in explicit collaboration

From the above literature review, two articles used 3D collaborative virtual worlds in their case studies to enhance the collaboration between workers. In particular, Sakurai et al., (2010) proposed a 3D context-aware collaborative environment based on a virtual space, an extension of the Project Wonderland[6] (now known as Open Wonderland) where the collaborators were represented as avatars. This platform also supports sharing and interaction with artifacts (e.g., documents or screen sharing) and facilitates the integration with sensors. Subsequently, Davier et al., (2017) focused on a virtual scenario to assess collaboration skills based on a web-based simulation task. In contrast, a study used the video call platform Zoom[7] for the assessment of CPS skills (Stewart et al. 2019). The latter performed posterior manual analysis on the video call recording, which would be interesting to combine with a natural language interface in a recommender system that allows an individual to make recommendation requests, with a taxonomy that encompassed the personalized user's requests (Kang et al., 2017). Surprisingly, an exception to most of the studies addressing explicit collaboration occurred in an experiment using microtasks (Eickhoff 2018). In this specific case, a microtask crowdsourcing campaign was launched in Mturk (which normally happens in implicit collaboration) to present the results of other workers that already answered the task.

3.5 User testing

In view of the types of evaluation for the proposed methods involving user testing, there is a clear distinction between laboratory and remote studies. In fact, there is only one experimentally verified example of the usage of both testing methods (Alagarai Sampath et al. 2014). Table 4 summarizes the user testing characteristics identified in the studies.

---

[3] https://success.appen.com/hc/en-us/articles/202703355-Contributors-CrowdFlower-s-Internal-Channel

[4] https://www.microworkers.com

[5] https://www.giscloud.com/

[6] http://openwonderland.org/

[7] https://zoom.us/

Table 4 Types of user testing addressed in the selected studies

Settings of User Testing		Collaboration		Article	Testing characteristics							
		Implicit	Explicit		Population		Cognitive Features				Time Supervision	
					Group	Size	Abilities	Bias	Engage	Style	Tracking	Limit
Laboratory	Paper and Pencil	×		(Goncalves et al. 2017)	Tertiary Students	24	×				×	
	Laptop/Desktop	×	×	(Sakurai et al. 2010)		30	×					
			×	(Stewart et al. 2019)		96	×				×	×
		×		(Engin and Vetschera 2017)		227	×			×		
		×		(Alagarai Sampath et al. 2014)		38	×					
Remote	Crowdsourcing Platform	×	×	(Eickhoff 2018)	Crowd Workers	–		×				
		×		(D. Hettiachchi et al. 2019a, b)		102	×				×	×
		×		(Hettiachchi et al. 2020)		574	×				×	×
		×		(Evangelos Mourelatos and Tzagarakis 2016)		50	×				×	

Table 4 (continued)

Settings of User Testing	Collaboration		Article	Testing characteristics								
	Implicit	Explicit		Population		Cognitive Features				Time Supervision	
				Group	Size	Abilities	Bias	Engage	Style	Tracking	Limit
	×		(E Mourelatos et al., 2020)		250	×				×	
	×		(Ravana et al. 2018)		519	×				×	×
	×		(Kosinski et al. 2012)		175	×		×			
	×		(Gadiraju et al. 2017)		2000		×			×	
	×		(Feldman and Bernstein 2014)		37	×					
	×		(Hassan and Curry 2013)	Tertiary Stu-dents	34	×					
External Platform	×		(M. Erskine et al. 2019)		200	×					
		×	(Davier et al. 2017)	Crowd Work-ers	1500	×					
Dataset	×		(Saab et al. 2019)	–	–		×				

Table 4 (continued)

Settings of User Testing	Collaboration		Article	Testing characteristics							
	Implicit	Explicit		Population		Cognitive Features				Time Supervision	
				Group	Size	Abilities	Bias	Engage	Style	Tracking	Limit
	×		(Ponciano and Brasileiro 2015)	–	–			×			
Interview		×	(Chujfi and Meinel 2020)	Remote Workers	83				×		

3.5.1 Laboratory user testing

In five of the included studies, the user testing was conducted in laboratory by providing a laptop or desktop PC for allowing participants to perform the desired activities. In [51], the authors developed the experiment by using a paper and pencil cognitive test as a validated measure tool for identifying cognitive abilities. In two studies, the equipment used was not feasible in a remote experiment, being more seamless to setup the material in a laboratory facility. For instance, Sakurai et al., (2010) collected sensor data like accelerometer and webcam images to detect the face of participants, while Alagarai Sampath et al., (2014) performed several remote experiments. Then, a laboratory experiment was conducted by applying the eye-tracking technique (Alhadreti et al. 2017) to measure the visual attention of participants taking into account their on-screen activity (Alagarai Sampath et al. 2014). Considering the population attributes, all laboratory experiments recruited tertiary students (Mean (SD): 83 (85.469)) instead of targeting a real population (e.g., online crowd workers). Furthermore, the cognitive evaluation consisted essentially in measuring the cognitive abilities.

Drawing from the findings of our SLR, only two studies described the time duration and/or imposed time limit of experiments. Regarding the appropriateness of adaptation, only Sakurai et al., (2010) evaluated the appropriateness by conducting questionnaire and interviews but unfortunately they didn't specify what questionnaire was used and which method of interview was chosen. For analyzing the user behavior, Alagarai Sampath et al., (2014) created different versions of the testing interface to reflect different parameters of visual saliency and working memory. These authors evaluated the user performance using eye tracking and could identify the necessary changes to make in the proposed interface to improve performance.

3.5.2 Remote user testing

At a glance, most of the studies conducted in crowdsourcing platforms (e.g., Mturk) occurred predominantly in remote settings. Moreover, it is also noticeable that most of those studies indicated a time limit or expiration time for registration and task execution when using the crowdsourcing platform. Although this is possible in crowdsourcing platforms, some studies complement it with the usage of an external library. In particular, two related crowdsourcing studies conducted on Mturk (D. Hettiachchi et al. 2019a, b; Hettiachchi et al. 2020) used jsPsych, a JavaScript library that enables the design of web-based behavioral experiments by supporting the creation of tasks with the presentation of stimulus or the registration of response time (de Leeuw 2015). In the literature, there is evidence that this tool obtains precise response time (Chandler and Shapiro 2016) and achieves similar results when compared to laboratory studies (Hilbig 2016). Considering the population attributes, most of the remote experiments hired crowd workers with polarized sample values (Mean (SD): 540.7 (672.695)). These numbers can be explained by the high number of crowdsourcing studies, which can easily escalate the number of participants involved in their studies. Furthermore, the remote experiments had several cognitive evaluation methods applied with a clear exception in the case of cognitive styles. Looking inside their foundational elements, cognitive styles deal with the correct representation of content to each participant,

being mostly studied in experiments involving direct supervision with participants. Regarding the appropriateness of adaptation, none of the studies reported the evaluation of this variable which can be explained due to the difficulties emerged on evaluating them in a remote testing context. In user behavior, Eickhoff (2018) used different versions of the proposed interface for document relevance assessment microtasks, with versions differing in small changes on the input and output fields, to study the cognitive bias of crowdworkers. The user performance was evaluated in almost every study, based primarily on data log analysis. Hettiachchi et al., (2020) used the cognitive tasks and microtasks accuracy to evaluate the comparison of the adaptation performed (based on the correct task assignment) with other state-of-the-art methods. Regarding user performance, Hassan and Curry (2013) also made adaptations on microtasks accuracy, with the addition that this study focused on modeling several tasks based on human abilities from a validated taxonomy.

4 Discussion

The discussion will be presented separately in the next subsections, following the order of the research questions defined for this systematic review. In this regard, we provide some insights with an explanation of the significance and implications of the research in light of the results obtained.

4.1 RQ1—How cognitive personalization on digital platforms affects the task performance of workers?

In a broad sense, the effect produced by cognitive personalization varies depending on the type of cognitive features investigated. Cognitive abilities were the feature most found in the included articles with great results in the assignment of tasks to online workers, especially in the crowdsourcing scenario (Hettiachchi et al. 2020). Comparing with state-of-the-art methods, the personalization of task assignment based on cognitive abilities has managed to increase the performance of crowd workers. Apparently, Hettiachchi et al., (2020) investigation was possibly the one that came closest to truly personalizing tasks, as it mapped tasks according to the capabilities of crowd workers. Using this rationale, any worker can have access to tasks without being constrained by selection mechanisms determining the best workers based on performance information. This is the case of other studies in the scientific literature (e.g., Fan et al. 2015; Zheng et al. 2015)) that addressed task assignment by filtering out workers who had apparently more intelligence or capacity. This should be avoided as it can encourage the exclusion of people who could have the opportunity to perform digital work in online virtual spaces. Another study was based on the historical performance of crowd workers and tried to map the cognitive abilities without satisfactory results (Hassan and Curry 2013). However, the reason for these poor or inconclusive results comes from several limitations identified by the authors such as the small group size or the task design that could negatively affect the results. In addition to the literature on task assignment, task design was also studied through cognitive skills. In line with this, it

is possible to observe that performance can be improved when the tasks are designed taking into consideration the visual attention or working memory of the crowd workers (Alagarai Sampath et al. 2014). Nonetheless, although this study obtained good results, the design was static, and subsequently, no adaptation to each worker was noticed.

At a higher level, other cognitive features achieved positive results for improving performance in online work settings. Having conversational interfaces in a crowdsourcing setting, that supports the personalization of microtasks, substantially improve the performance of crowdworker (Mavridis et al., 2019). Concerning the study of cognitive bias, there were two studies that obtained remarkable outcomes in the selection of more competent crowd workers through the use of self-assessment quizzes based on the Dunning–Kruger effect (Kruger and Dunning 1999). Extrapolating to task design, it was reported that there is a direct influence between the cognitive bias of workers and their performance. When we look at the cognitive styles, two articles gave recommendations on how to improve the task assignment and task design in digital labor environments. However, such studies did not perform any kind of personalization, and we believe that this should be addressed in future work, as argued by Sternberg (1997). Nevertheless, a study conducted by Raptis et al., (2017) developed a method that allows the assessment of cognitive styles, based on an eye-tracking model. Based on the collected data, an implicit evaluation process was developed in order to identify the cognitive styles, and this process was highly effective. A similar study was done posterior but this time to infer cognitive abilities, regarding the research of the personalization of textual documents with embedded visualizations using eye-tracking data (Toker et al. 2019). One point to consider from these methods is that unfortunately it may not work well in a crowdsourcing context, in particular microtasks that require little duration, apart from that not all crowd workers would have the equipment needed for this method to work. In our vision, cognitive styles are equally or even more important to cognitive abilities as they refer to the correct adaptation of the information format to the cognition style of each individual. This importance is corroborated by Alagarai Sampath et al., (2014) through their work on cognitive abilities as a way to resemble the cognitive styles. For instance, the authors designed a set of to minimize the cognitive load of workers. To tackle this problem and achieve an appropriate customization based on the cognitive style of workers, the performance could be evaluated when they have to perform the tasks using different representations of the interfaces in order to map the tasks to the style of cognition of each person in run-time. Similar work has already been done in other research areas, such as web or mobile accessibility (Gajos et al. 2008; Goel, Findlater and Wobbrock, 2012).

4.2 RQ2—What are the personalization methods used in each primary study?

With regard to the cognitive personalization of digital labor, current methods vary depending on whether a study is conducted within or outside the laboratory. As mentioned previously, one of the cognitive tests that has been considered valid and reliable is the Kit of Factor-Referenced Cognitive Tests (Ekstrom et al. 1976; Herreen and Zajac 2018; Schaie et al. 1991). Despite the validity of these tests, it was considered

that its application in digital work settings could not be viable due to the due time that participants need to spend using a pencil-and-paper approach. The latency in the input data collection of interactive adaptive systems is crucial to obtain optimal results. Considering this, Hettiachchi et al., (2020) turned some cognitive tests into micro-tasks to get measurements on the cognitive abilities of crowd workers in a quick and seamless fashion. Moreover, a limitation of 3.5 s was imposed to complete each test. This allowed to guarantee greater effectiveness in the measurement of capacities. The transformation of time-consuming (yet valid and reliable) cognitive tests into quick microtasks provides a path to create a personalization method for task assignment and task design in crowdsourcing settings. In addition, other studies have also used online tasks based on psychological tests to measure cognitive abilities (Kosinski et al. 2012; Sakurai et al. 2010). In Stewart et al., (2019), the authors chose to implicitly measure the historical performance of each worker throughout the use of personalized inter-faces taking into account the measurement of CPS. Similar to the analysis of historical performance, a study predicted the future performance of crowd workers (Hassan and Curry 2013). The predictability is important in the context of personalization meth-ods, which as shown by Sakurai et al., (2010) had in their adaptability mechanism the usage of web camera feed, the crowd worker could predict and adjust the adaptations performed by the system.

Although both approaches seem to provide promising results to the automatic mea-surement of the cognitive abilities of online workers, it is necessary to consider the application of regular tests. Along this view, Zyskowski et al., (2015c) indicated that capacities of crowd workers are susceptible to fluctuations, and it is important to regu-larly measure their capacities with periodic tests, which in the perspective of interactive adaptive system is important to guarantee the timeliness and appropriateness of the adaptations.

Cognitive personalization can be aided if we distribute test-set questions, especially in the moments before task execution. There have been several studies that applied questionnaires to assess the cognitive abilities or cognitive styles of workers (Chujfi and Meinel 2020; Engin and Vetschera 2017; M. Erskine et al. 2019). Although the questionnaires are validated and effective for measurement purposes, they also require workers to allocate some time to answer, which may discourage the completion of tasks in digital labor platforms. An alternative to work around this problem was presented in some studies that asked the workers few questions. An example of this is the CRT instrument that relies in a three-item task to measure the cognitive abilities of each individual (Engin and Vetschera 2017; Toplak et al. 2011). This tool should be better explored to assess whether it is possible to achieve personalization based on the cognitive abilities of workers. Two studies also mentioned the use of a self-assessment question to choose the most suitable crowd workers (Gadiraju et al. 2017; Saab et al. 2019). The method used was to ask crowd workers to perform a microtask and then, answer what their self-assessment represents in relation to the answer given. Then, the self-assessment was compared to the effectiveness of the answer given, which can be an indicator of the worker's performance. The self-assessment techniques can be applied on the creation of personalization methods that in an online microtask, labor platform can enhance the predictability of the crowd workers on how the adaptation will be performed.

An alternative to the dynamic personalization of digital labor can be the a priori design of tasks to maximize the efficiency of workers. In this context, two crowdsourcing studies referred to the application of psychology theories to improve microtask design. First, a study based on text-transcription tasks (Alagarai Sampath et al. 2014) pointed out an improvement in crowdsourcing performance when considering small characteristics like placing the text insertion fields close to the target fields or underlining the essential fields. Other study highlighted that knowing a priori the cognitive bias of crowd workers can avoid the design of tasks that could result in lower work quality. The study identified critical cognitive biases such as presenting contradictory information (i.e., Anchoring bias (Tversky and Kahneman 1974)) or presenting an existing group of results toward making an individual to follow the group behavior (i.e., Bandwagon effect (Bikhchandani et al. 1992)).

4.3 RQ3—What types of evaluation were used to assess the validity of these methods?

4.3.1 RQ3.1—What evaluation techniques were adopted?

The evaluation techniques used in the selected articles can be grouped into laboratory and remote tests. When we focus on laboratory tests, Ekstrom et al., (Ekstrom et al. 1976) used a pencil-and-paper test to measure cognitive abilities, as previously validated as feasible in other studies (Herreen and Zajac 2018; Schaie et al. 1991). Another study mentioned the use of eye-tracking techniques that have been used to measure the visual attention of participants on the screen with effective results (e.g., (Alhadreti et al. 2017)). On the other hand, when we pass to remote studies, the functionalities already incorporated in crowdsourcing platforms for the tracking of time or actions taken by users were frequently used. Additionally, two studies mentioned the use of jsPsych JavaScript library, which allows the implementation of web-based behavioral experiments (de Leeuw 2015). Also in remote studies, the results obtained using state-of-the-art techniques were frequently compared, as for example in the cases of task assignment (Hettiachchi et al. 2020) or pre-screening of crowd workers (Gadiraju et al. 2017; Saab et al. 2019).

Although the studies presented have carried out evaluations to analyze the results, none of the included articles mentioned the application of usability testing (Tan et al. 2009). Usability makes it possible to assess whether a system is usable by the target population. Subsequently, it is necessary to take usability into account to allow a greater adequacy of the tools developed for workers. In the case of cognitive personalization, it also allows widening the range of tasks proposed for people with disabilities, who may find in digital work an opportunity to be integrated into the labor market (Zyskowski et al. 2015b).

4.3.2 RQ3.2—Did real users evaluate these methods?

All selected studies mentioned the analysis of results from real users to evaluate the proposed methods. Although two studies have used pre-existing datasets, the data

came from studies with real users (Ponciano and Brasileiro 2015; Saab et al. 2019). A problem found in these studies relies in the fact that a significant part of them used students instead of target people (i.e., online workers). This has been a recurring problem in studies with experimental results, pointing to a possible skew in the findings (Al-Ubaydli et al. 2017). From a digital labor personalization perspective, a possible solution to this problem is to perform tests with crowd workers as a low cost, effective, and scalable approach to obtain experimental results (Chandler and Shapiro 2016; Deng and Joshi 2016).

5 Concluding remarks, limitations, and future work

The aim of this systematic literature review was to evaluate the research done on cognitive personalization for digital labor in virtual workspaces. The purpose of the review was to identify possible research paths for improving the efficiency of digital work, both from the perspective of the requester and the worker. In particular, the research analyzed was focused on the cognitive profile of workers to allow a more effective analysis of the their real capacities. The methodology outlined allowed to obtain 20 studies on cognitive personalization for online microtasks labor. From the selected studies, some managed to make the customization at the level of task design and task assignment. Other studies have identified characteristics that provide interesting directions to follow as a guidance to achieve cognitive personalization in digital labor platforms. It is worth mentioning that in the cognitive domain, there exists a panoply of cognitive features derived from psychology theories that could be then implemented for achieving personalization. In this systematic review, most of the studies mentioned these theories, which were then grouped in four cognitive features. Other research aspects identified in the studies were also analyzed for giving a more comprehensive overview of the research conducted (e.g., user testing methods and techniques). Most of the studies focused on crowdsourcing scenarios, which corroborates a research trend in the latest years (Ghezzi et al. 2018). Subsequently, in consonance with the endeavors observed in microwork settings, cognitive personalization can be achieved quickly through cognitive evaluation in short and seamless microtasks (Hettiachchi et al. 2020), which reveal a potential line of further research following this paradigm.

The considerations that resulted from this systematic review allow us to elaborate some guidelines for carrying out cognitive personalization in online digital labor platforms. First, the analysis of cognitive abilities shows a promising path with solid results for the optimal task assignment. This allows each worker to be selected for their abilities with adequate evaluation efficiency, through the application of short-duration online cognitive tests. For example, the transformation of time-consuming (yet valid and reliable) cognitive tests into quick microtasks provides a path to create a personalization method for task assignment and task design in crowdsourcing settings. Another cognitive feature that should be taken into account are cognitive biases that can negatively affect worker performance (e.g., through the use of self-assessment quizzes to predict the Dunning–Kruger effect). Cognitive styles may help to personalize how information is properly presented to workers (e.g., preference for more textual or visual information). This can be accomplished through the implementation

of online cognitive tests that can be inferred to personalize the online digital labor tasks. Regarding the methods used for personalization, it should be taken into account that the workers' cognitive profile is susceptible to fluctuations and that it is important to regularly measure their capacities with periodic tests. Moreover, it is necessary to evaluate frequently the methods used by carrying out usability tests, which was not verified in any of the studies found. Usability makes it possible to assess whether a system is usable by the target population. In the case of cognitive personalization, it also allows the validation of online digital labor tasks. To perform usability tests, crowd workers can be recruited as a low cost, effective, and scalable approach to obtain valid results.

Our systematic review has some limitations in relation to the bias in the selection of articles. One limitation comes from the difficulty in choosing the right keywords to cover cognitive features, which comes from observing a wide range of psychology theories. Additionally, another limitation comes from the data collection of this review. As it was only one author to perform this step, some errors could have occurred in categorizing the findings of the selected articles. Furthermore, this limitation also arises from a few number of articles did not report the technologies used in their experiments, which difficult the accurate categorization. Regarding the ethical implication of the studies found in the cognitive personalization, one limitation arises: what does it mean ethically to not assign tasks to crowd workers based on their cognitive profile? In the task assignment, it is true that there exist pre-screening techniques that filter unsuitable crowd workers, fomenting a meritocracy model where inevitable would lead to excluded crowd workers. However, two studies addressed the task assignment having in consideration that every crowd worker could have the access their most suitable task based on their cognitive abilities (Hettiachchi et al. 2019a, 2020, 2019b). Another limitation was the scope of this study only comprised on microtasks and not in a set of more creative or collaborative crowdsourcing tasks also known as macrotasks. Currently it is difficult to find a significant number of macrotasks studies that reported cognitive personalization. Macrotasks are still an embryonic line of research, and most of the current focus is directed to the decomposition of macrotasks into microtasks. However, in the future, there may be potential to exploit macrotasks, as areas such as crowdsourcing software will continue to grow (Sarı et al. 2019). Another reason may be that the distribution of macrotasks does not require cognitive personalization, or it is not yet feasible, being deprecated to other techniques such as analysis of the history of results or self-evaluation (Samimi et al. 2016; Zheng et al. 2015). In a context of microtasks, personalization can work well, however, for macrotasks as complex as software development, this type of personalization can lose validity due to the wide range of skills necessary to assess for producing a subsequent adaptation of features.

Future research work could aim to implement cognitive theories identified in this review into microtasks for determining the cognitive profile of workers. Extending the work from Hettiachchi et al., (2020), both cognitive styles and cognitive bias could represent important features to be explored in terms of dynamic personalization in crowdsourcing settings. These cognitive features provide advantages not only on task assignment but also in terms of task design that is also an essential aspect in digital work. Additionally, it is important to conduct usability tests in order to get more

insights as the studies mentioned in this systematic literature review lack an objective view regarding it.

Declarations

Competing interests The authors declare no competing interests.

References

Ahmi, A., Mohamad, R.: Evaluating accessibility of Malaysian ministries websites using WCAG 2.0 and section 508 guideline. J. Telecommun. Electr. Comput. Eng. **8**(8), 177–183 (2016)

Alagarai Sampath, H., Rajeshuni, R., Indurkhya, B.: Cognitively inspired task design to improve user performance on crowdsourcing platforms. Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems (2014)

Alhadreti, O., Elbabour, F., Mayhew, P.: Eye tracking in retrospective think-aloud usability testing: is there added value? J. Usability Stud. **12**(3), 95–110 (2017)

Allinson, C.W., Hayes, J.: The cognitive style index: a measure of intuition-analysis for organizational research. J. Manag. Stud. **33**(1), 119–135 (1996)

Al-Ubaydli, O., List, J.A., Suskind, D.L.: What can we learn from experiments? Understanding the threats to the scalability of experimental results. Am. Econ. Rev. **107**(5), 282–286 (2017)

Andreasen, M.S., Nielsen, H.V., Schrøder, S.O., Stage, J.: What happened to remote usability testing? An empirical study of three methods. Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems (2007)

Anthes, G.: HTML5 leads a web revolution. Commun. ACM **55**(7), 16–17 (2012)

Aydin, B.I., Yilmaz, Y.S., Demirbas, M.: A crowdsourced who wants to be a millionaire player. Concurr. Comput. Pract. Exp. **33**, e4168 (2017)

Bastien, J.M.C.: Usability testing: a review of some methodological and technical aspects of the method. Int. J. Med. Inform. **79**(4), e18–e23 (2010). https://doi.org/10.1016/j.ijmedinf.2008.12.004

Bayus, B.L.: Crowdsourcing new product ideas over time: an analysis of the dell ideastorm community. Manag. Sci. **59**(1), 226–244 (2013). https://doi.org/10.1287/mnsc.1120.1599

Bhatti, S.S., Gao, X., Chen, G.: General framework, opportunities and challenges for crowdsourcing techniques: a comprehensive survey. J. Syst. Softw. **167**, 110611 (2020)

Bikhchandani, S., Hirshleifer, D., Welch, I.: A theory of fads, fashion, custom, and cultural change as informational cascades. J. Polit. Econ. **100**(5), 992–1026 (1992)

Brereton, P., Kitchenham, B.A., Budgen, D., Turner, M., Khalil, M.: Lessons from applying the systematic literature review process within the software engineering domain. J. Syst. Softw. **80**(4), 571–583 (2007). https://doi.org/10.1016/j.jss.2006.07.009

Chandler, J., Shapiro, D.: Conducting clinical research using crowdsourced convenience samples. Ann. Rev. Clin. Psychol. **12**, 53 (2016)

Chang, J.C., Amershi, S., Kamar, E.: Revolt: collaborative crowdsourcing for labeling machine learning datasets. Paper presented at the proceedings of the 2017 CHI conference on human factors in computing systems, Denver, Colorado, USA. https://doi.org/10.1145/3025453.3026044(2017)

Cheung, J.H., Burns, D.K., Sinclair, R.R., Sliter, M.: Amazon mechanical turk in organizational psychology: an evaluation and practical recommendations. J. Bus. Psychol. **32**(4), 347–361 (2017). https://doi.org/10.1007/s10869-016-9458-5

Chujfi, S., Meinel, C.: Matching cognitively sympathetic individual styles to develop collective intelligence in digital communities. AI Soc. **35**(1), 5–15 (2020). https://doi.org/10.1007/s00146-017-0780-x

Cooper, S., Khatib, F., Treuille, A., Barbero, J., Lee, J., Beenen, M., players, F.: Predicting protein structures with a multiplayer online game. Nature **466**(7307), 756–760 (2010). https://doi.org/10.1038/nature09304

Correia, A., Paredes, H., Fonseca, B.: Scientometric analysis of scientific publications in CSCW. Scientometrics **114**(1), 31–89 (2018). https://doi.org/10.1007/s11192-017-2562-0

Crucian, G.P., Armaghani, S., Armaghani, A., Foster, P.S., Burks, D.W., Skoblar, B., Heilman, K.M.: Visual–spatial disembedding in Parkinson's disease. J. Clin. Exp. Neuropsychol. **32**(2), 190–200 (2010)

Daniel, F., Kucherbaev, P., Cappiello, C., Benatallah, B., Allahbakhsh, M.: Quality control in crowdsourcing: a survey of quality attributes, assessment techniques, and assurance actions. ACM Comput. Surv. **51**(1), 7 (2018). https://doi.org/10.1145/3148148

Davier, A.V., Hao, J., Liu, L., Kyllonen, P.: Interdisciplinary research agenda in support of assessment of collaborative problem solving: Lessons learned from developing a collaborative science Elsevier (2017)

de Leeuw, J.R.: jsPsych: a javascript library for creating behavioral experiments in a web browser. Behav. Res. Methods **47**(1), 1–12 (2015). https://doi.org/10.3758/s13428-014-0458-y

de Winter, J.C.F., Kyriakidis, M., Dodou, D., Happee, R.: Using crowdflower to study the relationship between self-reported violations and traffic accidents. Proc. Manuf. **3**, 2518–2525 (2015). https://doi.org/10.1016/j.promfg.2015.07.514

Deng, X.N., Joshi, K.: Why individuals participate in micro-task crowdsourcing work environment: revealing crowdworkers' perceptions. J. Assoc. Inform. Syst. **17**(10), 3 (2016)

Difallah, D.E., Demartini, G., Cudré-Mauroux, P.: Pick-a-crowd: tell me what you like, and i'll tell you what to do. Paper presented at the proceedings of the 22nd international conference on World Wide Web, Rio de Janeiro, Brazil. https://doi.org/10.1145/2488388.2488421(2013)

Doan, A., Ramakrishnan, R., Halevy, A.Y.: Crowdsourcing systems on the world-wide web. Commun. ACM **54**(4), 86–96 (2011)

Edwards, J.R.: Person-Job Fit: A Conceptual Integration, Literature Review, and Methodological Critique: Wiley (1991)

Eickhoff, C.: Cognitive biases in crowdsourcing. Paper presented at the proceedings of the eleventh ACM international conference on web search and data mining (2018)

Ekstrom, R.B., Dermen, D., Harman, H.H.: Manual for kit of factor-referenced cognitive tests (Vol. 102): Educational testing service Princeton, NJ. (1976)

Ellsberg, D.: Risk, ambiguity, and the savage axioms. Q. J. Econ. **75**, 643–669 (1961)

Engin, A., Vetschera, R.: Information Representation in Decision Making: The Impact of Cognitive Style and Depletion Effects: Elsevier. (2017)

Eriksen, B.A., Eriksen, C.W.: Effects of noise letters upon the identification of a target letter in a nonsearch task. Percept. Psychophys. **16**(1), 143–149 (1974)

Erskine, M.A., Gregg, D.G., Karimi, J., Scott, J.E.: Geospatial reasoning ability: definition, measurement and validation. Int. J. Hum. Comput. Interact. **31**(6), 402–412 (2015)

Erskine, M., Gregg, D., Karimi, J., Scott, J.:Individual decision-performance using spatial decision support systems: a geospatial reasoning ability and perceived task-technology fit perspective: Springer. (2019)

Estellés-Arolas, E., González-Ladrón-de-Guevara, F.: Towards an integrated crowdsourcing definition. J. Inform. Sci. **38**(2), 189–200 (2012). https://doi.org/10.1177/0165551512437638

Fallon, C.K., Panganiban, A.R., Wohleber, R., Matthews, G., Kustubayeva, A.M., Roberts, R.: Emotional intelligence, cognitive ability and information search in tactical decision-making. Person. Individ. Differ. **65**, 24–29 (2014) https://doi.org/10.1016/j.paid.2014.01.029

Fan, J., Li, G., Ooi, B.C., Tan, K.-l., Feng, J.: icrowd: An adaptive crowdsourcing framework. Paper presented at the Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data. (2015)

Feldman, M., Bernstein, A.: Cognition-based task routing: towards highly-effective task-assignments in crowdsourcing settings. 35th international conference on information systems "building a better world through information systems", ICIS 2014. (2014)

Finnerty, A., Kucherbaev, P., Tranquillini, S., Convertino, G.: Keep it simple: reward and task design in crowdsourcing. Paper presented at the proceedings of the biannual conference of the italian chapter of SIGCHI, Trento, Italy. https://doi.org/10.1145/2499149.2499168 (2013)

Finucane, M.L., Gullion, C.M.: Developing a tool for measuring the decision-making competence of older adults. Psychol. Aging **25**(2), 271–288 (2010). https://doi.org/10.1037/a0019106

Fleishman, E.A.: Toward a taxonomy of human performance. Am. Psychol. **30**(12), 1127 (1975)

Fleishman, E.A., Costanza, D.P., Marshall-Mies, J.: Abilities *an Occupational Information System for the 21st Century: The Development Of O*NET*, pp. 175–195. Am. Psychol. Assoc., Washington, DC, US (1999)

Frederick, S.: Cognitive reflection and decision making. J. Econ. Perspect. **19**(4), 25–42 (2005)

Gadiraju, U., Demartini, G., Kawase, R., Dietze, S.: Crowd anatomy beyond the good and bad: behavioral traces for crowd worker modeling and pre-selection. Comput. Support. Cooperat. Work (CSCW) **28**(5), 815–841 (2019). https://doi.org/10.1007/s10606-018-9336-y

Gadiraju, U., Fetahu, B., Kawase, R., Siehndel, P., Dietze, S.: Using worker self-assessments for competence-based pre-selection in crowdsourcing microtasks. ACM Trans. Comput. Hum. Interact. **24**(4), 30 (2017). https://doi.org/10.1145/3119930

Gajos, K.Z., Wobbrock, J.O., Weld, D.S.: Improving the performance of motor-impaired users with automatically-generated, ability-based interfaces. Paper presented at the proceedings of the SIGCHI conference on human factors in computing systems. (2008)

Galpaya, H., Perampalam, S., Senanayake, L.: Investigating the Potential for Micro-work and Online-Freelancing in Sri Lanka. In: Pupillo, L., Noam, E., Waverman, L. (eds.) Digitized Labor: The Impact of the Internet on Employment, pp. 229–250. Springer International Publishing, Cham (2018)

Germine, L., Nakayama, K., Duchaine, B.C., Chabris, C.F., Chatterjee, G., Wilmer, J.B.: Is the Web as good as the lab? Comparable performance from Web and lab in cognitive/perceptual experiments. Psychon. Bull. Rev. **19**(5), 847–857 (2012). https://doi.org/10.3758/s13423-012-0296-9

Ghezzi, A., Gabelloni, D., Martini, A., Natalicchio, A.: Crowdsourcing: a review and suggestions for future research. Int. J. Manag. Rev. **20**(2), 343–363 (2018)

Gillespie, A., Best, C., O'Neill, B.: Cognitive function and assistive technology for cognition: a systematic review. J. Int. Neuropsychol. Soc. **18**(1), 1–19 (2011). https://doi.org/10.1017/S1355617711001548

Goncalves, J., Feldman, M., Hu, S., Kostakos, V., Bernstein, A.: Task routing and assignment in crowd-sourcing based on cognitive abilities. Paper presented at the proceedings of the 26th international conference on world wide web companion. (2017)

Gorm, N., Shklovski, I.:Sharing steps in the workplace: Changing privacy concerns over time. Paper presented at the proceedings of the 2016 CHI conference on human factors in computing systems. (2016)

Graf, S., Lin, T., Jeffrey, L., Kinshuk.: An exploratory study of the relationship between learning styles and cognitive traits. Paper presented at the innovative approaches for learning and knowledge sharing, Berlin, Heidelberg. (2006)

Graham, M., Hjorth, I., Lehdonvirta, V.: Digital labour and development: impacts of global digital labour platforms and the gig economy on worker livelihoods. Transf. Eur. Rev. Labour Res. **23**(2), 135–162 (2017). https://doi.org/10.1177/1024258916687250

Haselton, M.G., Nettle, D., Murray, D.R.: The evolution of cognitive bias. The handbook of evolutionary psychology, pp. 1–20 (2015)

Hassan, U., Curry, E.: A capability requirements approach for predicting worker performance in crowdso*urcing.* Paper presented at the 9th IEEE international conference on collaborative computing: networking, applications and worksharing. (2013)

Herreen, D., Zajac, I.T.: The reliability and validity of a self-report measure of cognitive abilities in older adults: more personality than cognitive function. J. Intell. **6**(1), 1 (2018)

Hettiachchi, D., van Berkel, N., Hosio, S., Kostakos, V., Goncalves, J.: Effect of cognitive abilities on crowd-sourcing task performance. Paper presented at the IFIP conference on human-computer interaction. (2019a)

Hettiachchi, D., van Berkel, N., Hosio, S., Kostakos, V., Goncalves, J.: Effect of cognitive abilities on crowdsourcing task performance. *Vol. 11746 LNCS*. lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics) (pp. 442–464). (2019b)

Hettiachchi, D., van Berkel, N., Kostakos, V., Goncalves, J.: CrowdCog: a cognitive skill based system for heterogeneous task assignment and recommendation in crowdsourcing. Paper presented at the proceedings of the ACM on human-computer interaction. (2020)

Hilbig, B.E.: Reaction time effects in lab-versus Web-based research: experimental evidence. Behav. Res. Methods **48**(4), 1718–1724 (2016)

Ho, S.Y., Tam, K.Y.: An empirical examination of the effects of web personalization at different stages of decision making. Int. J. Hum. Comput. Interact. **19**(1), 95–112 (2005). https://doi.org/10.1207/s15327590ijhc1901_7

Hu, H., Zheng, Y., Bao, Z., Li, G., Feng, J., Cheng, R.: Crowdsourced POI labelling: location-aware result inference and task assignment. Paper presented at the 2016 IEEE 32nd international conference on data engineering (ICDE). (2016)

Huang, Y., Sundar, S.S.: Do we trust the crowd effects of crowdsourcing on perceived credibility of online health information. Health Commun. (2020). https://doi.org/10.1080/10410236.2020.1824662

Huber, J., Payne, J.W., Puto, C.: Adding asymmetrically dominated alternatives: violations of regularity and the similarity hypothesis. J. Consum. Res. **9**(1), 90–98 (1982)

Ipeirotis, P.G.: Analyzing the amazon mechanical turk marketplace. XRDS Crossroads ACM Mag. Stud. **17**(2), 16–21 (2010)

Jarupathirun, S., Zahedi, F.M.: Exploring the influence of perceptual factors in the success of web-based spatial DSS. Decis. Support Syst. **43**(3), 933–951 (2007). https://doi.org/10.1016/j.dss.2005.05.024

Kazai, G., Kamps, J., Milic-Frayling, N.: Worker types and personality traits in crowdsourcing relevance labels. Paper presented at the Proceedings of the 20th ACM international conference on Information and knowledge management. (2011)

Kazai, G., Kamps, J., Milic-Frayling, N.:The face of quality in crowdsourcing relevance labels: demographics, personality and labeling accuracy. Paper presented at the proceedings of the 21st ACM international conference on Information and knowledge management, Maui, Hawaii, USA. https://doi.org/10.1145/2396761.2398697(2012)

Keith, M.G., Tay, L., Harms, P.D.: Systems perspective of amazon mechanical turk for organizational research: review and recommendations. Front. Psychol. (2017). https://doi.org/10.3389/fpsyg.2017.01359

Kitchenham, B.: Procedures for performing systematic reviews. Keele, UK, Keele Univ. **33**(2004), 1–26 (2004)

Kosinski, M., Bachrach, Y., Kasneci, G., Van-Gael, J., Graepel, T.: Crowd IQ: measuring the intelligence of crowdsourcing platforms. Proceedings of the 4th annual ACM web science conference (pp. 151–160). (2012)

Koutrika, G., Bercovitz, B., Kaliszan, F., Liou, H., Garcia-Molina, H.: CourseRank: a closed-community social system through the magnifying glass. Proc. Int. AAAI Conf. Web Soc. Med. **3**(1), 98 (2009)

Kruger, J., Dunning, D.: Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments. J. Pers. Soc. Psychol. **77**(6), 1121 (1999)

Kucherbaev, P., Daniel, F., Tranquillini, S., Marchese, M.: Crowdsourcing processes: a survey of approaches and opportunities. IEEE Internet Comput. **20**(2), 50–56 (2016). https://doi.org/10.1109/MIC.2015.96

Law, E., Ahn, L.V.: Human computation. Synth. Lectures Artif. Intell. Machine Learn. **5**(3), 1–121 (2011)

Lintott, C., Schawinski, K., Bamford, S., Slosar, A., Land, K., Thomas, D., Raddick, M.J.: Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. Monthly Notices r. Astronom. Soc. **410**(1), 166–178 (2011)

Littlemore, J.: Metaphoric competence: a language learning strength of students with a holistic cognitive style? TESOL q. **35**(3), 459–491 (2001). https://doi.org/10.2307/3588031

Lykourentzou, I., Antoniou, A., Naudet, Y., Dow, S.P.: Personality matters: Balancing for personality types leads to better outcomes for crowd teams. Paper presented at the proceedings of the 19th ACM conference on computer-supported cooperative work & social computing. (2016)

MacLeod, C.M.: Half a century of research on the Stroop effect: an integrative review. Psychol. Bull. **109**(2), 163 (1991)

Maguire, M., Isherwood, P.: A comparison of user testing and heuristic evaluation methods for identifying website usability problems. Paper presented at the international conference of design, user experience, and usability. (2018)

Malone, T.W., Laubacher, R., Dellarocas, C.: Harnessing crowds: mapping the genome of collective intelligence. (2009)

Miller, E., Wallis, J.: Executive function and higher-order cognition: definition and neural substrates. Encycloped. Neurosci. **4**, 99–104 (2009)

Mitzner, T.L., Savla, J., Boot, W.R., Sharit, J., Charness, N., Czaja, S.J., Rogers, W.A.: Technology adoption by older adults: findings from the PRISM Trial. Gerontologist **59**(1), 34–44 (2019). https://doi.org/10.1093/geront/gny113

Monsell, S.: Task switching. Trends Cogn. Sci. **7**(3), 134–140 (2003). https://doi.org/10.1016/S1364-6613(03)00028-7

Montello, D.R.: Cognitive Research in GIScience: recent achievements and future prospects. Geogr. Compass **3**(5), 1824–1840 (2009). https://doi.org/10.1111/j.1749-8198.2009.00273.x

Morabito, V.: Digital Work and Collaboration Trends and Challenges in Digital Business Innovation, pp. 113–131. Springer International Publishing, Cham (2014)

Mourelatos, E., Giannakopoulos, N., Tzagarakis, M.: Personality traits and performance in online labour markets. Behav. Inform. Technol. **41**(3), 468–484 (2022)

Mourelatos, E., Tzagarakis, M.: Worker's cognitive abilities and personality traits as predictors of effective task performance in crowdsourcing tasks. Paper presented at the proceedings of 5th ISCA/DEGA workshop on perceptual quality of systems (PQS 2016). (2016)

Mtsweni, J., Burge, L.: The potential benefits of mobile microwork services in developing nations: Research opportunities and challenges. Paper presented at the 2014 IST-Africa Conference Proceedings. (2014)

Müller, M., Vetter, S., Weiser, M., Frey, F., Ajdacic-Gross, V., Stieglitz, R.-D., Rössler, W.: Precursors of cognitive impairments in psychotic disorders: a population-based study. Psychiatry Res. **210**(1), 329–337 (2013). https://doi.org/10.1016/j.psychres.2013.05.035

Müller, H., Kalpathy-Cramer, J., Seco, G., de Herrera, A.: Experiences from the ImageCLEF Medical Retrieval and Annotation Tasks. In: Ferro, N., Peters, C. (eds.) Information Retrieval Evaluation in a Changing World: Lessons Learned from 20 Years of CLEF, pp. 231–250. Springer International Publishing, Cham (2019)

Muller, C.L., Chapman, L., Johnston, S., Kidd, C., Illingworth, S., Foody, G., Leigh, R.R.: Crowdsourcing for climate and atmospheric sciences: current status and future potential. Int. J. Climatol. **35**(11), 3185–3203 (2015). https://doi.org/10.1002/joc.4210

Murtza, M.H., Gill, S.A., Aslam, H.D., Noor, A.: Intelligence quotient, job satisfaction, and job performance: the moderating role of personality type. J. Pub. Affairs (2020). https://doi.org/10.1002/pa.2318

Nguyen, N.N., Nham, P.T., Takahashi, Y.: Relationship between ability-based emotional intelligence, cognitive intelligence, and job performance. Sustainability **11**(8), 2299 (2019)

Oechssler, J., Roider, A., Schmitz, P.W.: Cognitive abilities and behavioral biases. J. Econ. Behav. Organ. **72**(1), 147–152 (2009)

Oleson, D., Sorokin, A., Laughlin, G., Hester, V., Le, J., & Biewald, L.: Programmatic gold: targeted and scalable quality assurance in crowdsourcing. Paper presented at the workshops at the twenty-fifth AAAI conference on artificial intelligence. (2011)

Owen, A.M., McMillan, K.M., Laird, A.R., Bullmore, E.: N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum. Brain Mapp. **25**(1), 46–59 (2005)

Paolacci, G., Chandler, J., Ipeirotis, P.G.: Running experiments on amazon mechanical turk. Judgm. Decis. Mak. **5**(5), 411–419 (2010)

Paramythis, A., Weibelzahl, S., Masthoff, J.: Layered evaluation of interactive adaptive systems: framework and formative methods. User Model. User-Adap. Inter. **20**(5), 383–453 (2010)

Paramythis, A., Totter, A., Stephanidis, C.: modular approach to the evaluation of adaptive user interfaces. Paper presented at the empirical evaluation of adaptive systems. Proceedings of workshop at the eighth international conference on user modeling, UM2001. (2001)

Paulino, D., Pinheiro, P., Rocha, J., Martins, P., Rocha, T., Barroso, J., Paredes, H.: Assessment of wizards for eliciting users' accessibility preferences. In *Proceedings of the 9th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-Exclusion* (pp. 135–140) (2020)

Petrides, M., Alivisatos, B., Evans, A.C., Meyer, E.: Dissociation of human mid-dorsolateral from posterior dorsolateral frontal cortex in memory processing. Proc. Natl. Acad. Sci. **90**(3), 873–877 (1993). https://doi.org/10.1073/pnas.90.3.873

Ponciano, L., Brasileiro, F.: Finding volunteers' engagement profiles in human computation for citizen science projects. *arXiv preprint* arXiv:1501.02134. (2015)

Raddick, M.J., Prather, E.E., Wallace, C.S.: Galaxy zoo: science content knowledge of citizen scientists. Pub. Underst. Sci. **28**(6), 636–651 (2019). https://doi.org/10.1177/0963662519840222

Rahman, H., Roy, S.B., Thirumuruganathan, S., Amer-Yahia, S., Das, G.: Task assignment optimization in collaborative crowdsourcing. Paper presented at the 2015 IEEE international conference on data mining. (2015)

Ramsey, W.: Must cognition be representational? Synthese **194**(11), 4197–4214 (2017). https://doi.org/10.1007/s11229-014-0644-6

Ravana, S.D., Samimi, P., Rajagopal, P.: Quality of crowdsourced relevance judgments in association with logical reasoning ability. Malays. J. Comput. Sci., pp. 73–86 (2018)

Raven, J.: The Raven's progressive matrices: change and stability over culture and time. Cogn. Psychol. **41**(1), 1–48 (2000)

Riding, R.J.: On the nature of cognitive style. Educ. Psychol. **17**(1–2), 29–49 (1997)

Rotgans, J.I., Schmidt, H.G.: Cognitive engagement in the problem-based learning classroom. Adv. Health Sci. Educ. **16**(4), 465–479 (2011)

Roth, Y., Kimani, R.: Crowdsourcing in the production of video advertising: the emerging roles of crowd-sourcing platforms International perspectives on business innovation and disruption in the creative industries: Edward Elgar Publishing. (2014)

Rundgren, C.-J., Rundgren, S.-N.C., Tseng, Y.-H., Lin, P.-L., Chang, C.-Y.: Are you SLiM? Developing an instrument for civic scientific literacy measurement (SLiM) based on media coverage. Pub. Underst. Sci. **21**(6), 759–773 (2012)

Saab, F., Elhajj, I.H., Kayssi, A., Chehab, A.: Modelling cognitive bias in crowdsourcing systems. Cogn. Syst. Res. **58**, 1–18 (2019). https://doi.org/10.1016/j.cogsys.2019.04.004

Sakurai, Y., Kinshuk, Takada, K., Kawabe, T., Knauf, R., Tsuruta, S.: Evaluation of enriched cyberspace for adaptive support of remote collaboration. Paper presented at the 2010 sixth international conference on signal-image technology and internet based systems. (2010)

Samimi, P., Ravana, S.D., Koh, Y.S.: Effect of verbal comprehension skill and self-reported features on reliability of crowdsourced relevance judgments. Comput. Hum. Behav. **64**, 793–804 (2016). https://doi.org/10.1016/j.chb.2016.07.058

Sarı, A., Tosun, A., Alptekin, G.I.: A systematic literature review on crowdsourcing in software engineering. J. Syst. Softw. **153**, 200–219 (2019). https://doi.org/10.1016/j.jss.2019.04.027

Schaie, K.W., Dutta, R., Willis, S.L.: Relationship between rigidity-flexibility and cognitive abilities in adulthood. Psychol. Aging **6**(3), 371 (1991)

Schmidt, F.L., Hunter, J.: General mental ability in the world of work: occupational attainment and job performance. J. Pers. Soc. Psychol. **86**(1), 162 (2004)

Schmidt, L.I., Wahl, H.-W.: Predictors of performance in everyday technology tasks in older adults with and without mild cognitive impairment. Gerontologist **59**(1), 90–100 (2018). https://doi.org/10.1093/geront/gny062

Schmidt, F.A.: Crowdsourced production of AI Training Data: How human workers teach self-driving cars how to see. (2019)

Schmitt, N.: Personality and cognitive ability as predictors of effective performance at work. Ann. Rev. Organ. Psychol. Organ. Behav. **1**(1), 45–65 (2014)

Shaw, A.D., Horton, J.J., Chen, D.L.: Designing incentives for inexpert human raters. Paper presented at the proceedings of the ACM 2011 conference on computer supported cooperative work, Hangzhou, China. https://doi.org/10.1145/1958824.1958865(2011)

Simpson, R., Povich, M., Kendrew, S., Lintott, C., Bressert, E., Arvidsson, K., Sherman, R.: The milky way project first data release: a bubblier galactic disc. Monthly Not. r. Astronom. Soc. **424**(4), 2442–2460 (2012)

Stanek, K.C., Ones, D.S.: Taxonomies and compendia of cognitive ability and personality constructs and measures relevant to industrial, work and organizational psychology. *D. S. Ones, N. Anderson, C. Viswesvaran, & H. K. Sinangil (Eds.), The SAGE handbook of industrial, work & organizational psychology: Personnel psychology and employee performance*, pp. 366–407.(2018)

Sternberg, R.J.: Thinking Styles. Cambridge University Press, Cambridge (1997)

Stewart, A.E.B., Vrzakova, H., Sun, C., Yonehiro, J., Stone, C.A., Duran, N.D., D'Mello, S.K.: I say, you say, we say: using spoken language to model socio-cognitive processes during computer-supported collaborative problem solving. Proc. ACM Hum. Comput. Interact. **3**, 194 (2019). https://doi.org/10.1145/3359296

Stol, K.-J., Fitzgerald, B.:*Two's company, three's a crowd: a case study of crowdsourcing software development*. Paper presented at the proceedings of the 36th international conference on software engineering, Hyderabad, India. https://doi.org/10.1145/2568225.2568249 (2014)

Sundar, S.S.: Self as source: Agency and customization in interactive media *Mediated interpersonal communication* (pp. 72–88): Routledge. (2008)

Tan, W.-S., Liu, D., Bishu, R.: Web evaluation: Heuristic evaluation vs. user testing. Int. J. Ind. Ergon. **39**(4), 621–627 (2009)

Toker, D., Moro, R., Simko, J., Bielikova, M., Conati, C.: Impact of english reading comprehension abilities on processing magazine style narrative visualizations and implications for personalization. Paper presented at the proceedings of the 27th ACM conference on user modeling, adaptation and personalization, Larnaca, Cyprus. https://doi.org/10.1145/3320435.3320447 (2019)

Toplak, M.E., West, R.F., Stanovich, K.E.: The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. Mem. Cognit. **39**(7), 1275 (2011)

Turner, M.L., Engle, R.W.: Is working memory capacity task dependent? J. Mem. Lang. **28**(2), 127–154 (1989). https://doi.org/10.1016/0749-596X(89)90040-5

Tversky A., Kahneman D.: Judgment under uncertainty Heuristics and biases, science, 185(4157): 1124–1131 (1974)

Van Velsen, L., Van Der Geest, T., Klaassen, R., Steehouder, M.: User-centered evaluation of adaptive and adaptable systems: a literature review. Knowl. Eng. Rev. **23**(3), 261–281 (2008)

Vyas, L., Butakhieo, N.: The impact of working from home during COVID-19 on work and life domains: an exploratory study on Hong Kong. Policy Des. Pract. **4**(1), 59–76 (2021). https://doi.org/10.1080/25741292.2020.1863560

Wazny, K.:Crowdsourcing ten years in: a review. J. Glob. Health, 7(2): 020602–020602. https://doi.org/10.7189/jogh.07.020602. (2017)

Webster, J.: Microworkers of the gig economy: separate and precarious. New Labor Forum **25**(3), 56–64 (2016). https://doi.org/10.1177/1095796016661511

Wohlin, C.: Guidelines for snowballing in systematic literature studies and a replication in software engineering. Paper presented at the Proceedings of the 18th international conference on evaluation and assessment in software engineering. (2014)

World Health Organization.: International classification of functioning, disability and health: ICF: World Health Organization. (2001)

Zhang, S., Grenhart, W.C.M., McLaughlin, A.C., Allaire, J.C.: Predicting computer proficiency in older adults. Comput. Hum. Behav. **67**, 106–112 (2017). https://doi.org/10.1016/j.chb.2016.11.006

Zhang, D., Zhang, Y., Li, Q., Plummer, T., Wang, D.: Crowdlearn: a crowd-AI hybrid system for deep learning-based damage assessment applications. Paper presented at the 2019 IEEE 39th international conference on distributed computing systems (ICDCS). (2019)

Zheng, Y., Wang, J., Li, G., Cheng, R., Feng, J.: QASCA: A quality-aware task assignment system for crowdsourcing applications. Paper presented at the Proceedings of the 2015 ACM SIGMOD international conference on management of data. (2015)

Zyskowski, K., Morris, M.R., Bigham, J.P., Gray, M.L., Kane, S.K.: Accessible crowdwork? Understanding the value in and challenge of microtask employment for people with disabilities. Paper presented at the proceedings of the 18th ACM conference on computer supported cooperative Work & Social Computing, Vancouver, BC, Canada. https://doi.org/10.1145/2675133.2675158 (2015a)

Zyskowski, K., Morris, M.R., Bigham, J.P., Gray, M.L., Kane, S.K.: Accessible crowdwork? Understanding the value in and challenge of microtask employment for people with disabilities. Paper presented at the Proceedings of the 18th ACM conference on computer supported cooperative work & social computing. (2015b)

Zyskowski, K., Morris, M. R., Bigham, J.P., Gray, M. L., & Kane, S. K.: Accessible crowdwork?: Understanding the value in and challenge of microtask employment for people with disabilities. Paper presented at the proceedings of the 18th ACM conference on computer supported cooperative work, Vancouver, BC, Canada. (2015c)

Dennis Paulino received the Master's Degree in Informatic Engineer, at University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal in December 2018. He participated since September 2016 until December 2018 in the project NanoSTIMA RL2 - Passus Mobile, responsible for developing a system that makes exercise supervision of people with peripheral arterial disease. From January 2019 until November 2019, he participated in the project eCSAAP, responsible for the construction of a informatic system to help in the visualization and detection of meteorological phenomena. In Dezember 2019 he ingress in PhD in Informatics at UTAD, with a scholarship financed by FCT.

António Correia is a former Microsoft Research fellow and works as a Research Assistant at INESC TEC, Porto, Portugal. He received M.Sc. degree in Information and Communication Technologies from the University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal, where he recently obtained a Ph.D. degree in Computer Science, funded by the Portuguese Foundation for Science and Technology (FCT). Furthermore, he formerly worked as a Visiting Postgraduate Researcher at University of Kent, Medway, UK. António holds more than ten years of experience in research and scientific writing, and his research interests are mainly in the field of Human-Computer Interaction (HCI), including the areas of Computer-Supported Cooperative Work (CSCW), crowdsourcing, artificial intelligence, and sciento-metrics. He has authored or co-authored more than 40 publications, including journal articles, conference papers, book chapters, and posters. Moreover, he has also participated in research projects conducted at national and international level and has been executing functions as external reviewer and scientific committee member for top-tier journals and conferences covering aspects of computer science.

João Barroso earned a doctorate in University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal in 2002 in Electrical Engineering and held in 2008 the Habilitation in Informatics/Accessibility and became an Associate Professor in December 2012. He was Pro-Rector for Innovation and Information Management at UTAD from July 2010 to July 2013, Pro-Rector for Innovation and Technology Transfer from May 2017 and June 2020, and Vice-Rector for Innovation, Technology Transfer and Digital University since May 2021. He produced over 150 scientific papers, including book chapters, journal articles and articles in proceedings of scientific events. He supervised 40 postgraduate students (masters and doctorates). He was member of the research team in 35 research and development projects. He was member of several organizing committees of the international scientific meetings. In 2006 he directed the team that created the conference "Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion (www.dsai.ws/2016) and in 2016 the conference Technology and Innovation is Sports, Health and Wellbeing (www.tishw.ws/2016). and Human Computer Interaction".

Hugo Paredes received the B.Eng. and Ph.D. degrees in computer science from the University of Minho, Braga, Portugal, in 2000 and 2008, respectively, and the Habilitation title from the University of Trás-os-Montes e Alto Douro (UTAD), Vila Real, Portugal, in 2016. Since 2003, he has been with UTAD, where he is currently an Associate Professor with Habilitation. In 2017, he was a Visiting Faculty with Human Computer Interaction Institute, Carnegie Mellon University, Pittsburgh, PA, USA. He is currently the Director of the Master Program with Informatics Engineering, UTAD, and a Member of the Managing Board of Ph.D. in Computer Science with UTAD. He is also an Assistant Coordinator with the Centre for Computer Graphics and Information Systems (CSIG), Institute for Systems and Computer Engineering, Technology and Science (INESC TEC), Porto, Portugal, where he is a Senior Researcher.