



HAL
open science

Fair Design of Learners Descriptive Cards

Inaya El Alaoui, Céline Treuillier, Anne Boyer

► **To cite this version:**

Inaya El Alaoui, Céline Treuillier, Anne Boyer. Fair Design of Learners Descriptive Cards. Fourth International Workshop on Human-Centred Learning Analytics (HCLA) held in conjunction with Learning Analytics and Knowledge Conference (LAK'23), Mar 2023, Online, France. hal-04412811

HAL Id: hal-04412811

<https://hal.science/hal-04412811>

Submitted on 23 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Fair Design of Learners Descriptive Cards

Inaya El Alaoui¹, Céline Treuillier¹ and Anne Boyer¹

¹Université de Lorraine, CNRS, LORIA, Vandoeuvre-lès-Nancy, France

Abstract

The analysis and processing of learning traces through Learning Analytics, although essential to help improve learning, requires addressing strong ethical issues. Tools provided to various educational stakeholders must ensure fair learning, and take into account the diversity of students and their behavior while avoiding bias and discrimination. In that context, descriptive cards are an emerging tool allowing for a representative description of the diversity of behaviors adopted within a class. However, their design raises a lot of questions about equity and fairness. To contribute to the need for fair learning, we present a study conducted in collaboration with teachers and propose a new design of descriptive cards, limiting as much as possible the introduction of bias.

Keywords

Descriptive Cards, Trustworthy Learning Analytics, Ethical Learning, Learning Behavior

1. Introduction

Nowadays, it is essential to consider the ethical concerns raised by the processing of the large quantity of data available about learners. In that context, learning analytics tools must be directed towards Trustworthy Learning Analytics (TLA) that respect the criteria of Trustworthy Artificial Intelligence (TAI) defined by the European Commission¹. In that way, LA systems must respect, among others, the diversity criterion by avoiding algorithmic bias and discrimination, as well as integrating stakeholders during their conception [1].

Among the large variety of existing LA systems, we find descriptive cards [2], which allow visualizing learning behaviors recorded in the learning traces. They help to improve explainability by taking the form of fictitious students to whom pedagogical experts can refer during their pedagogical tasks. However, the learners' behavior representation raises questions about bias and equity, leading to our research question **(RQ): How to design learner descriptive cards meeting the need for diverse, non-discriminative, and fair learning analytics?**

To address this question, we conducted a user study with teachers to propose a new way to design descriptive cards, participating in the dissemination of fairer learning analytics.

2. User Study: How to avoid representation bias?

The design of descriptive cards can lead to the introduction of several types of biases: algorithmic bias, with the under or over-representation of a class of population according to the studied

 inaya.elalaoui@gmail.com (I. E. Alaoui); celina.treuillier@loria.fr (C. Treuillier); anne.boyer@loria.fr (A. Boyer)



© 2022 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).



CEUR Workshop Proceedings (CEUR-WS.org)

¹<https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html>

data, but also human bias, with the introduction of stereotypes, whether cultural, gender or age-related. As personas represent users' archetypes, individuals who embody personas are often characterized by a photo, an age, and a gender, associated with a narrative description of their behavior. This can indicate ethnicity, socio-economic status, preferences, and sometimes education level which can clearly influence professionals in the design and understanding of the tool, because of the assumptions and stereotypes they may have [3, 4].

2.1. Methodology

We evaluated three different types of cards (See Fig. 1). The first one is the **avatar description card** (See Figure 1a) inspired by the one in Treuillier and Boyer [2]. To address our research question about stereotype bias, we represented learners from different ethnic backgrounds: North African, South American/Hispanic, Asian, African/Black Americans, and Caucasian, for both women and men. The second type of card is the **icons descriptive cards** (See Figure 1b), designed to counter the potential bias due to avatars. An icon reflecting the learner's result replaces the avatar, and a non-gendered first name is used. The third type of card is the **group descriptive cards** (See Figure 1c) which displays a group of avatars representing learners with the same learning behavior. All cards were colored with a neutral color, except for the icon cards with gender-neutral names, where colors could vary between pink and blue, to test whether it could influence the perception of gender.

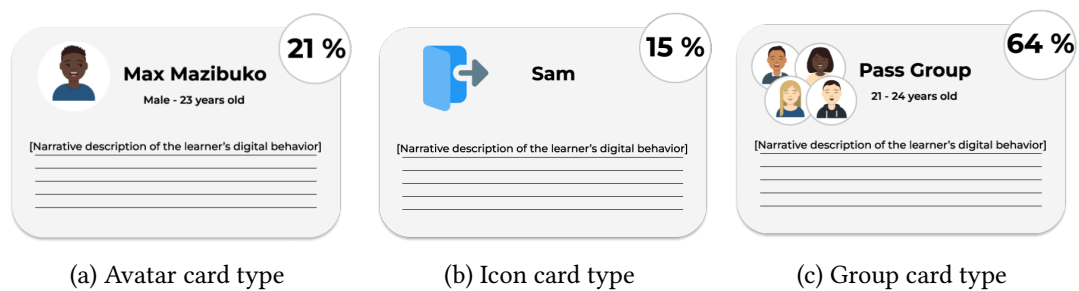


Figure 1: Evaluated Descriptive cards

Following our information campaign about our studies, 30 volunteer teachers were recruited. The first objective is to investigate whether teachers are actually influenced by stereotypes, and which descriptive cards representation are least likely to induce them. To do that, participants were asked to autonomously carry out a computer-assisted categorization task inspired by the Implicit Association Test (IAT) [5], which is intended for measuring the stereotyping biases linked to the representation of learners on the descriptive cards, in an implicit way. Next, the second objective is to assess the participants' perception about descriptive cards. Once the part on stereotyping bias was completed, participants were thus invited to fill out a questionnaire in order to evaluate how teachers receive and perceive the descriptive cards in a pedagogical context. This questionnaire was adapted from the Persona Perception Scale [6], and we evaluated five relevant factors: completeness, coherence, clarity, empathy and usefulness, and intention or willingness to use. Participants had to indicate their level of agreement for each type of descriptive card using a 7-point Likert scale ("strongly disagree" (1) to "strongly agree" (7)).

2.2. Results

2.2.1. Representation bias

Results related to the categorization task show that, when facing avatar descriptive cards (See Figure 1a), participants associate positive terms with all learners, regardless of gender or ethnicity. However, this association is more important for Caucasian and Asian avatars, while students from the North African, Black (especially the Black male student), and Hispanic minorities were more often linked to negative terms than their counterparts mentioned above (only 75% to 90% of participants associate positive terms to those ethnicities). Concerning the icon descriptive cards (See Figure 1b), results are coherent with most of the icons displayed on the descriptive cards. However, few participants expressed their reluctance to use several icons, which in their words "*does not seem clear*", because they are not associated with the right concept. Finally, for the group descriptive cards (See Figure 1c), the names of the groups seem to be fairly explicit. Besides, about gender bias, which can be introduced by the names on the cards or by their colors, results indicate that some of the non-gendered names are very strongly associated with the male gender, regardless of the color of the card (for about 70% of participants). Regarding what guided their association choice, only 39% felt influenced by the cards' colors. Among this sample, half associated the color with a gender whereas the other half thought it was related to the learner's success when it was blue or failure when it was pink. Also, 10 participants said that they automatically associated failure with the male gender and success with the female gender. Finally, a large majority of the participants stated that they based their choice on their interpretation of the icons and the group names. They all added that this information was the most explicit, making the descriptive cards more understandable.

2.2.2. Persona Perception Scale

Concerning the persona perception scale results, scores obtained for each factor (completeness, coherence, clarity, empathy, usefulness) were high for all types of descriptive cards. However, the group card type (See Fig.1c) is the one with better scores of coherence, empathy, and intention to use, while the icon card type (See Fig. 1b) has better results for clarity and completeness scores. Finally, concerning the users' preferences for a type of card and the reasons for their choice, it appears that 10% of the 30 participants stated that they preferred the avatar cards (See Fig. 1a), against 66% for the group cards (See Fig.1c) and 24% for the icon cards (See Fig. 1b). We can therefore see that a majority of participants expressed a preference for group cards. Among this majority, 25% of them expressed the fact that the visual aspect of the icon cards was interesting to combine with the formulation and presentation of the group cards. In their opinion, this would help to identify learners needing special attention more quickly and easily.

3. Conclusion

With this user study, we wanted to determine how descriptive cards could enhance pedagogical monitoring while respecting Trustworthy guidelines, especially by minimizing representation bias. Accordingly, we have highlighted stereotype bias that could emerge while using avatars, names, or colors and determine the teachers' preferred cards' features. In order to endorse

the diversity and non-discrimination criterion of TAI as well as the transparency one, we characterized how to design learner descriptive cards avoiding data representation bias, and thus allowing a fairer learners' behaviors characterization.

We then propose an updated version of descriptive cards, minimizing bias as much as possible. All presented results have been taken into account to propose this new version of descriptive cards. This user study carried out with the end-users of the studied LA tool enabled us to highlight some representation bias and stereotypes, and thus to propose a solution resulting from a joint reflection with educational stakeholders. This collaborative approach is essential to foster the development of fairer learning analytics tools.

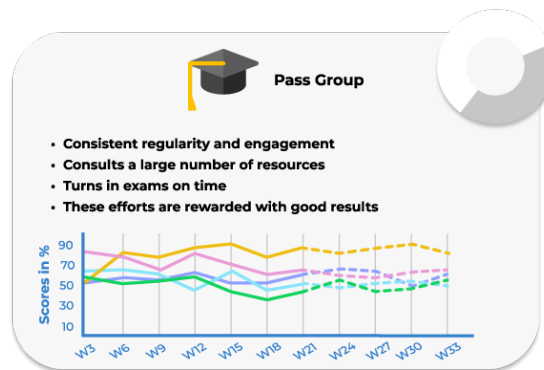


Figure 2: Enhanced version of descriptive card

References

- [1] H. Veljanova, C. Barreiros, N. Gosch, E. Staudegger, M. Ebner, S. Lindstaedt, Towards trustworthy learning analytics applications: An interdisciplinary approach using the example of learning diaries, in: *International Conference on Human-Computer Interaction*, Springer, 2022, pp. 138–145.
- [2] C. Treuillier, A. Boyer, A new way to characterize learning datasets., in: *CSEDU (2)*, 2022, pp. 35–44.
- [3] D. Spiliotopoulos, D. Margaris, C. Vassilakis, Data-assisted persona construction using social media data, *Big Data and Cognitive Computing* 4 (2020) 21.
- [4] P. Turner, S. Turner, Is stereotyping inevitable when designing with personas?, *Design studies* 32 (2011) 30–44.
- [5] A. G. Greenwald, D. E. McGhee, J. L. Schwartz, Measuring individual differences in implicit cognition: the implicit association test., *Journal of personality and social psychology* 74 (1998) 1464.
- [6] J. Salminen, H. Kwak, J. M. Santos, S.-G. Jung, J. An, B. J. Jansen, Persona perception scale: developing and validating an instrument for human-like representations of data, in: *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, 2018, pp. 1–6.