

Fairness of Academic Performance Prediction for the Distribution of Support Measures for Students: Differences in Perceived Fairness of Distributive Justice Norms

Marco Lünich, Birte Keller & Frank Marcinkowski

Article - Version of Record



Suggested Citation:

Lünich, M., Keller, B., & Marcinkowski, F. (2023). Fairness of Academic Performance Prediction for the Distribution of Support Measures for Students: Differences in Perceived Fairness of Distributive Justice Norms. *Technology, Knowledge and Learning*, 29(2), 1079–1107.
<https://doi.org/10.1007/s10758-023-09698-y>

Wissen, wo das Wissen ist.

This version is available at:

URN: <https://nbn-resolving.org/urn:nbn:de:hbz:061-20250304-103009-2>

Terms of Use:

This work is licensed under the Creative Commons Attribution 4.0 International License.

For more information see: <https://creativecommons.org/licenses/by/4.0>



Fairness of Academic Performance Prediction for the Distribution of Support Measures for Students: Differences in Perceived Fairness of Distributive Justice Norms

Marco Lünich¹ · Birte Keller¹ · Frank Marcinkowski¹

Accepted: 15 October 2023 / Published online: 11 November 2023
© The Author(s) 2023

Abstract

Artificial intelligence in higher education is becoming more prevalent as it promises improvements and acceleration of administrative processes concerning student support, aiming for increasing student success and graduation rates. For instance, Academic Performance Prediction (APP) provides individual feedback and serves as the foundation for distributing student support measures. However, the use of APP with all its challenges (e.g., inherent biases) significantly impacts the future prospects of young adults. Therefore, it is important to weigh the opportunities and risks of such systems carefully and involve affected students in the development phase. This study addresses students' fairness perceptions of the distribution of support measures based on an APP system. First, we examine how students evaluate three different distributive justice norms, namely, *equality*, *equity*, and *need*. Second, we investigate whether fairness perceptions differ between APP based on human or algorithmic decision-making, and third, we address whether evaluations differ between students studying science, technology, engineering, and math (STEM) or social sciences, humanities, and the arts for people and the economy (SHAPE), respectively. To this end, we conducted a cross-sectional survey with a 2×3 factorial design among $n = 1378$ German students, in which we utilized the distinct distribution norms and decision-making agents as design factors. Our findings suggest that students prefer an equality-based distribution of support measures, and this preference is not influenced by whether APP is based on human or algorithmic decision-making. Moreover, the field of study does not influence the fairness perception, except that students of STEM subjects evaluate a distribution based on the need norm as more fair than students of SHAPE subjects. Based on these findings, higher education institutions should prioritize student-centric decisions when considering APP, weigh the actual need against potential risks, and establish continuous feedback through ongoing consultation with all stakeholders.

Keywords Algorithmic decision-making · Higher education · Algorithmic fairness · Distributive fairness · Student support measures

1 Introduction

Artificial Intelligence (AI) applications in the higher education sector have been acknowledged to have great and diverse potential (Attaran et al., 2018; Daniel, 2015). For example, there are already AI applications that assign grades (Kotsiantis, 2012), predict dropouts or academic performance (Alyahyan & Düşteğör, 2020; Alturki et al., 2022; Berens et al., 2019; Kemper et al., 2020; Armatas et al., 2022), or chatbots that can answer various questions from students (Nguyen et al., 2021; Pelletier et al., 2021; Vieira et al., 2022). Developers and users hope that Artificial Intelligence in Education (AIED) will lead to a higher number of graduates and improved student performance, for example, by offering students better individual feedback and support to reduce dropout rates (Attaran et al., 2018; Daniel, 2015; Zawacki-Richter et al., 2019). However, debates about the social justice of corresponding systems repeatedly come into focus (Baker & Hawn, 2022; Fazelpour & Danks, 2021; Hsu et al., 2021; Keller et al., 2022; Marcinkowski et al., 2020; Muñoz et al., 2016; Slade & Prinsloo, 2013). An illustrative example is provided by an automated study place allocation system from France called *ParcourSup* (Wenzelburger & Hartmann, 2022). The system was quickly criticized because, in addition to a lack of transparency, it was feared that individual student groups would be discriminated against in the allocation of study places, thus threatening to exacerbate existing social inequalities (Orwat, 2020). Further concerns have been voiced in instances where algorithms predict the likelihood of graduation even before students start studying. It is feared that such a system could lead to discrimination against students from lower-income backgrounds (Muñoz et al., 2016). Similar applications that predict academic performance based on past performance, among other factors, also point to the danger of different forms of bias, for example, regarding the input data (Fazelpour & Danks, 2021).

Accordingly, questions of algorithmic fairness are becoming virulent not only in the higher education sector but also in numerous other areas, such as criminal justice or hiring processes (Angwin et al., 2016; Kaibel et al., 2019). Consequently, the state of research in fair machine learning (FAIR ML) offers numerous mathematical fairness notions that enable equitable distribution (Verma & Rubin, 2018; Makhlof et al., 2021; Friedler et al., 2021). However, there has yet to be a consensus on which of these notions seems most appropriate in preventing systematic discrimination. Instead, the choice of an adequate notion of fairness is highly context-dependent, and as a result, the perceptions of fairness may vary (Lee et al., 2019; Starke et al., 2022; Wong, 2020). While the various fairness notions tend to be developed at the theoretical level, only scattered evidence on stakeholders' perceptions of the same is available (Saxena et al., 2020; Srivastava et al., 2019). Thereby, it is shown that the fairness perception of an algorithm has a significant impact on the trust in and acceptance of the AI system in question as well as the legitimacy of the ensuing decision-making (Lünich & Kieslich, 2022; Shin, 2020, 2021; Shin et al., 2020; Sun & Tang, 2021). Against this background, it is necessary to include the fairness perceptions of important stakeholders in the development of an AI system (Cheng et al., 2021; Keller et al., 2022; Webb et al., 2018) in order to achieve trustworthy AI (European Commission, 2019; Mäntymäki et al., 2022; Shneiderman, 2020).

To answer this call for research on fairness perceptions of those most affected by AIED, in this paper, we investigate students' preferences for different understandings of justice in the context of the distribution of support measures based on an academic performance prediction (APP) system. However, we do not focus on specific mathematical fairness notions but on students' underlying justice principles concerning the AI-based distribution

of support measures. We draw on the four-dimensional concept of *organizational justice* (Cropanzano & Ambrose, 2015; Greenberg, 1993) and focus primarily on the dimension of distributive fairness. Based on a factorial survey of German students, we provide an explanatory analysis for the three following questions: First, we investigate how students perceive different norms for just distribution as fair when evaluating the APP system. Furthermore, we investigate whether the fairness perceptions of AI-based APP systems (i.e., so-called algorithmic decision-making, ADM) differ from those of human decision-making (HDM). Third, we look at the different fields of study and investigate whether the choice of field of study impacts the fairness perception of the APP system. In answering these questions, this paper adds to the literature on AIEd and points to relevant concerns of students, as the arguably most important stakeholders in this field, concerning distributive justice, connecting to the literature on FAIR ML.

2 A Question of Justice: Perils and Pitfalls of Academic Performance Prediction

Concerning AIEd, a systematic review by Zawacki-Richter et al. (2019) shows four application areas: “profiling and prediction, intelligent tutoring systems, assessment and evaluation, and adaptive systems and personalisation” (p. 20). A so-called APP system is an example that falls within the first and third category. As the name suggests, the system’s function is to use predictive analytics to forecast students’ performance or, in some cases, to prognosticate whether they will complete their studies (so-called dropout detection). APP, like other systems of AIEd, is based on historical student performance data, sometimes supplemented in part by non-academic data (e.g., sociodemographics) (Abu Saa et al., 2019; Olaya et al., 2020). Universities, as the systems’ users, hope to increase retention and graduation rates based on the predictions (Attaran et al., 2018; Mah, 2016). APP is already being used at various higher education institutions (HEI) worldwide, for example, in the USA (Arnold & Pistilli, 2012; Ornelas & Ordonez, 2017), Australia (Adams et al., 2017; Niall & Mullan, 2017), Germany (Berens et al., 2019; Kemper et al., 2020), or Bangladesh (Ahmed & Khan, 2019). Based on the algorithmic predictions, HEIs may offer individual feedback and support measures that can be tailored explicitly to the needs of the students and automatically assigned to them with the help of APP (Ekowo et al., 2016; Muñoz et al., 2016; Pistilli & Arnold, 2010).

Universities use APP to allocate limited student support resources faster and more equitably and, as a result, increase student success and receive higher graduation rates. Such support resources could be academic and non-academic interventions (Keller et al., 2022), as Tinto (1975) distinguishes two types of integration that are important goals to achieve student retention: namely, *academic integration* and *social integration*. Interventions for academic integration are used to improve the student’s academic performance. Such improvement can be achieved, for example, through measures such as tutorials, advising sessions for study planning, or student coaching (Bettinger & Baker, 2014; Chiteng Kot, 2014) but also through individually tailored feedback (Brade et al., 2018; Tinto, 1975), which could be greatly facilitated and improved by an APP system. While these support interventions come *from above*, others come *from below* (Morosan et al., 2010), such as those for social integration. These latter interventions build a social support network and can be fostered especially through contact with experienced teachers and other students, for instance, through mentoring programs. In addition, however, the sense of belonging

to the institution—that is, the university—also plays a decisive role (Hausmann et al., 2009; Neugebauer et al., 2019; Tinto, 1975). Assuming limited organizational resources (financial and human) that a university can muster to support students, these interventions often cannot be made available to all students. Instead, they can only be deployed efficiently, for instance, if they are utilized (only) by distressed students in need or—should a university pursue a strategy of promoting only the best students—by promising and already successful students.

According to the 2021 EDUCAUSE Horizon Report, the COVID-19 pandemic has further accelerated the development and use of hybrid learning models and AIED (Pelletier et al., 2021). However, the authors further state that “the arrival of AI in higher education has opened a Pandora’s box. Going forward, higher education will need to be a careful and ethical user of AI” (Pelletier et al., 2021, 15). In this sense, APP’s implementation must be thoroughly reflected and systematically evaluated. Almost all commonly discussed issues that come along with AI, in general, can also be applied to AIED and the APP context, respectively. These concerns range from technological over methodological to societal challenges (Hagendorff & Wezel, 2020; Wirtz et al., 2019) and include, for example, questions of transparency (de Fine Licht, 2020; Ananny & Crawford, 2018), accountability (Busuioc, 2021; Diakopoulos, 2016), and fairness (Veale & Binns, 2017; Shin, 2019). Because the use of AIED can significantly impact the prospects of young adults, it is crucial to weigh the opportunities and risks of any proposed system carefully.

For instance, an automated assessment system in the UK offers a negative example of problematic AI systems in the educational sector. Since no A-level examinations could be conducted during the first lockdown of the COVID-19 pandemic in 2020, an algorithm was implemented to predict the A-level grades of students in the UK. However, this led to massive protests because students claimed the prediction fell far from their expectations and actual performance as they were downgraded (Edwards, 2021). The procedure was particularly problematic from a fairness perspective, as the algorithm disadvantaged students from more deprived families compared to those from private schools, and the grades awarded impact further education and working life (Adams & McIntyre, 2020; Smith, 2020). While the British algorithmic grading system was abandoned after the protests, a similar algorithm in France, which also predicted baccalaureate grades, was accepted. However, it resulted in higher pass rates than in prior years, leading to higher demand for university places (Smith, 2020).

These examples show that without reflection and adjustment, APP systems may perpetuate social injustices and biases inherent to human social structures (Attaran et al., 2018; Fazelpour & Danks, 2021). Thus, when it comes to the use of APP, it is essential to ensure not only that the system provides an appropriate level of transparency and explainability (of both process and outcome) but that it is clear who can be held accountable in the event of poor decisions, or that privacy concerns are addressed, but also that APP is free of bias and that no one is systematically disadvantaged (Keller et al., 2022). Faced with the threat of discrimination by AIED, as evidenced by the UK grading algorithm, much of the academic literature addresses the issue of fair distribution, the translation of fairness into mathematical notions, and the reduction of algorithmic bias in AIED (Jiang & Pardos, 2021; Martinez Neda et al., 2021). Beyond this, however, special attention must be paid to the perceptions of those affected. Therefore, many authors stress the importance of stakeholder involvement in developing and introducing AI systems (Cheng et al., 2021; Howell et al., 2018; Keller et al., 2022; Knox et al., 2022; Lee et al., 2019; Webb et al., 2018). Otherwise, an APP system introduced with initially good intentions and presumably bias-mitigating execution may still fail in the face of reality.

Consequently, in addition to the properties of the system itself, such as the use of sensitive attributes (e.g., sociodemographics) (Nyarko et al., 2021), it can be assumed that the consequences that follow from APP also influence people's perceptions of fairness. Such individual perceptions, however, are often neglected in the literature (Starke et al., 2022). So, what does the cognitive process entail in concrete terms for the students affected? In addition to their self-reflective fairness assessment of the APP assigned to them, students' perceptions may also relate to the resulting distribution of support measures based on the forecast. For example, students may perceive it as unfair if those measures are not open to everyone in principle. The prediction could then become problematic if it leads to unintended adverse perception effects, which affect students' attitudes and behavior; for instance, should non-prioritized students also demand access to support measures. Consequently, a positive and a negative performance prediction could hurt study motivation (Fazelpour & Danks, 2021; Legault et al., 2006; Mah, 2016) and potentially put students under additional pressure (Yilmaz et al., 2020). While high-performing students might rest on the positive feedback, lower-performing students might become frustrated given their previous study efforts. Furthermore, students might see the prediction as an incentive to adapt to the algorithm's logic and change their behavior accordingly (Dai et al., 2021; Fazelpour & Danks, 2021), which may also lead to positive and negative effects. Eventually, negative perceptions of prediction outcomes of APP and decisions based on such predictions may have detrimental effects on students' attitudes towards HEIs deploying AIED and the perceived legitimacy of AI systems (Lünich & Kieslich, 2022; Marcinkowski et al., 2020). Therefore, the question of a fair distribution of student support measures based on the APP system arises.

3 Justice, Fairness and the Distribution of Support Measures

Rawls (1999) lays the foundation of the question of fair distribution by describing fairness as the basis of justice. In his understanding, negotiating principles of justice requires a fair societal system that allows cooperation between equal and free individuals (Rawls, 1999; Sen, 2009). Murphy (2011) also highlights the link between both constructs by describing: "If proposed principles are deemed unfair, they are unjust" (S. 337). Accordingly, fairness judgments arise from a perceived injustice, whereby a person must be held accountable for that injustice (Folger & Cropanzano, 2001). A perceived (in-)justice usually refers to the distribution of goods but can also include the distribution of rights and freedoms (Rawls, 1999). While the distribution of support measures, in our case, refers to goods, denied access could also be perceived as a restriction on student freedom. As HEIs may see APP as a basis for decision-making on how to support and enable students to improve their academic performance, the ensuing distribution of said support measures poses a distribution issue, if only because it must be partly about the distribution of limited goods and resources. For example, as mentioned before, due to limited resources, offering tutoring and additional personal counseling may not be available to all students.

The perception of an equitable distribution of goods is also addressed by the *organizational justice* literature, which examines how intra-organizational decision processes should be designed to achieve a maximum degree of satisfaction and commitment of organizational members (Greenberg, 1987), which in turn is necessary to prevent students from protesting or, at worst, leaving the university as a result of using or being dissatisfied with APP (Marcinkowski et al., 2020). Concerning the concept of *organizational justice*

four dimensions of fairness have emerged: In addition to distributive fairness, there is procedural, interactional, and informational fairness (Greenberg, 1993). In this paper, however, we focus exclusively on distributive fairness to capture the essence of the distributive justice norms of a fair distribution of support measures.

3.1 Distributive Justice Norms

In light of the question of how different norms of just distribution are perceived by students when evaluating the APP system, we focus on the dimension of distributive fairness that asks for the validity of a decision, respectively an outcome (Cropanzano et al., 2001). This dimension is based on equity theory (Adams, 1965). According to equity theory, employees compare the ratio of their individually received outcome and the individually perceived input with the corresponding ratio of other employees. If this comparison of the two ratios does not match, the employee with the greater share is considered to be unjustly overpaid. In comparison, the employee with a lower share is considered to be unjustly underpaid (Greenberg, 1990). According to this, *equity* is oriented toward merit and is based on comparing the input produced by oneself with the output received (Cropanzano & Ambrose, 2015). However, such a meritocratic distribution is only one of three different distributive justice norms besides *equality* and *need* Deutsch (1975). In the case of APP, a distribution according to the *equity* principle could mean that only students who have already performed well so far are entitled to the support measures presented. The *need* principle offers an alternative option. This norm stipulates that only those students in particular need are provided with support interventions. Accordingly, only students with poor APP outcomes could benefit from support measures. Furthermore, a distribution according to the *equality* principle is also conceivable. In this case, no distinction is made, and all students receive access to support measures, which could reach its limits in everyday life because of the institution's limited resources. Of course, when it comes to applied distribution rules, these ideal-typical distribution norms may come in any combination and with different gradations and overlaps.

3.2 Fairness Perceptions of Distributive Justice Norms

Empirical research gives evidence that the choice of the *right* (i.e., publicly approved) distributive justice norm is highly context-dependent (Lee et al., 2017; Starke et al., 2022; Wong, 2020) and can also vary among different stakeholders (Cheng et al., 2021; Lee et al., 2019). To date, however, the literature on FAIR ML focuses mainly on the formulation of mathematical fairness notions (Dwork et al., 2012; Kusner et al., 2017; Verma & Rubin, 2018), their evaluation by those affected (Cheng et al., 2021; Debjani et al., 2020; Srivastava et al., 2019; Zhou et al., 2021) or takes distributive fairness in general into consideration (Hannan et al., 2021; Marcinkowski et al., 2020; Wang et al., 2020). However, the underlying norms of just distribution and respective preferences for distinct norms should be addressed.

Nevertheless, some authors conclude that an egalitarian distribution, as suggested by the *equality* principle, is preferred in many contexts (Cropanzano & Ambrose, 2015; van Hootegem et al., 2020). Such a preference seems to be especially the case when stakeholders see themselves as part of a (small) group, and resources must be shared collectively (Allison et al., 1992), as empathy is also a predictor of a preference for *equality* as a distributive justice norm (Huppert et al., 2019). This preference is explained with the help of decision

heuristics, which are based on the assumption that all parties concerned unspokenly accept the most simple solution of an equal distribution as the correct decision since they assume that this represents a satisfactory outcome for all (Allison & Messick, 1990; Allison et al., 1992). Therefore, this type of heuristic is also called equality heuristic or inequality aversion and is shown to come into play in different contexts (Asaria et al., 2023; Morand et al., 2020).

Concerning ADM systems, Lee and colleagues (Lee & Baykal, 2017), for example, show within qualitative interviews that some stakeholders prefer altruistic distributions and are even willing to compromise and forgo unequally distributed allocations. Some of them reasoned for this preference by prioritizing the group's happiness. While the authors substantiated this result in a second study, an international comparison between India and America also showed cultural differences in evaluating distributive justice norms. In India, *need* was preferred, while for Americans, the *equality* or *equity* norm was more in favor (Lee et al., 2017). In the case of a healthcare scheduling system, employees primarily preferred an equal distribution norm, although consideration of individual needs received equal attention (Uhde et al., 2020). In a similar use case of a denied vacation request, however, Schlicker et al. (2021) do not find any evidence for the assumption that *equality* explanations lead to a higher distributive fairness perception than *equity* explanations. Saxena et al. (2020), on the other hand, come to a different conclusion when asking for a preferred loan decision. Their results indicate that respondents favor a proportional 'ratio' distribution compared to meritocratic or equality norms. Kasinidou et al. (2021) confirm this insight by surveying computer science students who prefer a proportional distribution to an *equality* norm. However, due to the contextual nature of fairness perceptions and the inconsistent empirical insights, the potential perceptions of students toward a distribution of support interventions based on an APP system can only be inferred with qualification based on these findings. Therefore, evaluating different fairness norms in the specific context of APP and the distribution of student support measures is crucial. As a result, in RQ1, we ask: *How fair are different norms for fair distribution perceived by students when evaluating an APP system?*

3.3 Algorithmic vs. Human—Based Distribution of Support Measures

The main driver of the implementation of AI technologies, especially when it comes to AI in public administration, are expectations for AI to lead to cheaper, faster, more objective, and reliable results in decision-making compared to humans (König & Wenzelburger, 2022; Wirtz & Müller, 2019; Gomes de Sousa et al., 2019). While the superiority of temporary AI applications may often be dubious, even if AI proves to make better decisions, this must not necessarily lead to positive evaluations by the persons affected. In analogy to the potential difference between notions of factual and perceived fairness, preferences regarding the nature of the decision-makers may not solely be based on the quality or goodness of the outcome. Two contrasting strands of literature address whether people prefer AI systems over humans regarding decision-making. On the one hand, there is algorithm aversion (Dietvorst et al. 2015; Dietvorst & Bharti 2020; for an overview, see Burton et al. 2020), and on the other hand, there is algorithmic appreciation (Araujo et al., 2020; Logg et al., 2019; Thurman et al., 2019). Here, the former suggests a preference for HDM over ADM. Such preference is especially the case when errors are observed, as people lose confidence in algorithms more quickly than in other humans (Dietvorst et al., 2015). This rejection is particularly evident in the case of unpredictable and risky decisions (Dietvorst

& Bharti, 2020). The latter, algorithmic appreciation, however, sees ADM as superior and therefore assumes that people are more likely to follow the recommendation of algorithms than that of humans (Logg et al., 2019), based on the assumption that algorithms are less biased and decide more objectively (Araujo et al., 2020).

Closely connected to aversion or appreciation of ADM is whether AIED is wanted by students, administrators, and other educational sector stakeholders at all. Again, several studies indicate that the answer to this question can vary in different contexts and regarding the stakeholders involved (Starke et al., 2022; Wong, 2020; Marcinkowski et al., 2020). Empirical research, on the one hand, gives evidence that ADM systems are perceived as fairer than HDM, for instance, in logistics, even if those results do not explicitly focus on the dimension of distributive fairness (Bai et al., 2022). Nevertheless, higher distributive fairness is also attested to ADM in university admission processes (Marcinkowski et al., 2020) and without reference to a specific use case (Helberger et al., 2020). On the other hand, there are some use cases where HDM systems are attested to have higher distributive fairness than ADM, such as resource allocation in social division tasks (Lee & Baykal, 2017) or criminal justice (Harrison et al., 2020). Another example is the decision on skin cancer detection, where the fairness perception of ADM is also lower than those in HDM, which is also reflected in the level of trust towards the decision. Nevertheless, this is only true for people with low mistrust in human systems, whereas a high mistrust leads to equally (un-)fair evaluations of ADM and HDM (Lee & Rich, 2021). In this context, the fairness perception of ADM and HDM also depends on the nature of the decision. For example, it plays a role in whether mechanical or human-attributable skills are required (Lee, 2018). Furthermore, whether the decision's impact is high or low is equally important (Araujo et al., 2020).

However, there are also reports of ambivalent results, where no preference is given to the decision taken by an algorithmic system or a human decision-maker, for example, in the already mentioned denial of a vacation request (Schlicker et al., 2021) or the allocation process of COVID-19 vaccines (Lünich & Kieslich, 2022). All of these inconsistent findings underscore the importance of reflecting on and examining ADM fairness perceptions on a case-by-case basis, not just using ADM systems because the data are accessible and it is technically possible (boyd, danah, and Kate Crawford., 2012). In line with the ideas of *human-in-the-loop* (Holzinger, 2016; Starke & Lünich, 2020), the decision to use or not to use ADM systems offers a range of possible hybrid courses of action, for example, introducing cooperation between humans and algorithms. Thus, appropriate cooperation may also lead to higher perceived fairness, although it is still preferred that humans retain control over the decision (de Cremer & McGuire, 2022). Therefore, our RQ2 asks: *To what extent do perceptions of fairness APP systems using ADM differ from those using HDM?*

3.4 Differences in Perceived Fairness Based on Students' Study Subject

Focusing on the specific context of AIED, we also need to consider the varying perceptions of computer applications within different study fields and thus ask whether the study subject of students impacts their fairness perception of the APP system. Even though it can be assumed that today's students all have grown up with computers and are familiar with their use, there are different digital literacy levels between study subjects, which may shape perceptions of computer applications (Gibson & Mourad, 2018). As Lee and Baykal (2017) show, fairness perceptions of ADM systems vary depending on the respondents' computer programming knowledge. Contrary to their hypothesis,

their findings indicate that individuals with lower programming skills perceive ADM systems as fairer than respondents with higher knowledge. The authors assume that this evaluation corresponds with a feeling of loss of control over the ADM as students with more pronounced programming skills presumably know more about the algorithm's limitations. Logg et al. (2019) come to a similar conclusion, as their results indicate that experts relied less on algorithmic advice than laypeople. Jang et al. (2022) also point out that "students who had prior experience with AI education had a more sensitive attitude toward the fairness of AI compared to students without experience" (p. 20). In addition, and as mentioned before, Kasinidou et al. (2021) point to differences in the evaluation of different distributional norms, with computer science students preferring a ratio distribution to an equal distribution. However, they do not examine differences between different fields of study. Instead, they find marginal differences between undergraduate and postgraduate students, with the former viewing a given decision on loan allocation as more unfair than the latter when sociodemographic factors are considered.

To our knowledge, however, there is no evidence yet on the fairness perceptions of different fields of study towards AIED and APP, respectively. Against this background, when it comes to APP, it seems to be fruitful to take a closer look at the comparison of students from so-called *STEM* (science, technology, engineering, and mathematics) and *SHAPE* (social sciences, humanities, the arts for people and the economy) fields of study (Jones et al., 2020). On the one hand, we reason that *STEM* students are more likely to be familiar with the mathematical and engineering foundations of computer systems, algorithms, and AI techniques such as machine learning, which may shape their perspective on APP. On the other hand, we assume that *SHAPE* students have a greater familiarity and sensitivity with social questions of human nature and society that may reflect in their perceptions of a sociotechnical system such as APP.

That this distinction between study subjects may have an important influence on fairness perceptions is supported by evidence concerning additional differences in traits between the respective student groups. Such factors and their consequences for individual approaches to social questions in the application of data and prediction systems may influence distinct evaluations of data-driven prediction systems. For instance, according to the literature, personality traits and cognitive skills predict study choice. While the former mainly refers to the *Big Five personality traits* (eg., Humburg 2017; Vedel 2016), the latter is, for example, focused on mathematical skills but also on expected future earnings (eg., Arcidiacono et al., 2012). In this sense, it shows that higher expected future earnings, compared to current prospects, would also increase the probability of choosing *SHAPE* (Arcidiacono et al., 2012). Regarding personality traits, it can be stated that compared to *STEM* students, students of *SHAPE* subjects are more often willing to experiment and show higher levels of emotionality but are less conscientious (Verbree et al., 2021). Next to this, however, neuroticism also is attested to these students (Vedel, 2016), and they also seem to show more flexibility than students of *STEM* (Sherrick et al., 1971), who are instead less extrovert and more emotionally stable (Coenen et al., 2021; Humburg, 2017; Pozzebon et al., 2014). Nevertheless, Coenen et al., (2021) find that openness to experience also positively influences the study choice for *STEM*. Besides, the most interesting finding for our case is that *SHAPE* students show a higher preference for altruistic views compared to other study fields (Pozzebon et al., 2014). Following this, a study of German students shows that students who consider social engagement particularly important are underrepresented in *STEM* subjects (Stegmann, 1980). Accordingly, we ask RQ3: *To what extent does the field of study of students have an impact on the fairness perception of the APP system?*

4 Method

We conducted a cross-sectional factorial survey using a questionnaire with standardized response options to address the research questions. Assessing the results, we performed the data analysis in R (version 4.0.3) using the packages *lavaan* (Rosseel, 2012) and *semTools* (Jorgensen et al., 2019). The data and code of our analyses are shared online via OSF (<https://osf.io/djt39/>). Due to the exploratory nature of the study, there was no pre-registration. Accordingly, for our analysis, in which we sought to identify even small effect sizes that may hold consequential implications, we did not perform an a priori power analysis. First, we still needed sufficient information regarding the fit of our measurement models and the distribution of the underlying variables. On the other hand, against this background, the power analysis of covariance-based structural equation modeling (SEM) is also rather intricate. Given this, we turned to G*Power 3.1.9.7 for a post hoc sensitivity analysis targeting the effect size (Faul et al., 2009). With an alpha error probability set at 0.05 and a high power (1-beta error probability) of 0.999, our sample size in the smallest group was 453, and for the whole sample was 1378. First, our latent means model compared means for three groups. With a numerator degrees of freedom (df) of 2 and a denominator df of 1384, the smallest effect we could have reliably detected with our sample size was determined to be a $f \sim 0.15$ (i.e., a small effect). Second, our regression model encompassed six predictors. This analysis rendered the smallest effect we could have reliably detected with our sample size to be a f^2 of ~ 0.08 (i.e., a small effect). These parameters guide our inquiry, emphasizing our interest in detecting even subtle effects that may be overlooked in less rigorously designed analyses. Remember that due to the different statistical approaches of SEM, our results below may deviate from this sensitivity analysis.

4.1 Procedure

For screening purposes, respondents had to indicate their occupational status to determine if they could participate in the student survey. Afterward, they were informed about the nature and aim of the study, the duration of approximately 15 min to answer the questionnaire, and the use and protection of their data. After giving informed consent, respondents started to answer the questionnaire that included the vignettes of the factorial survey design, the ensuing fairness perception of a university's APP procedure, questions regarding their self-assessed propensity for dropout, their expected probability that an ADM would assign to their dropout, and their sociodemographics—i.e., age, gender, study subject, numbers of semesters of study, their intended degree, and the type of their recent HEI.

Lastly, participants were thanked, debriefed, and redirected to the provider of the online access panel (OAP), where they received monetary compensation for participation. The average time to complete the questionnaire was 13.72 min ($SD = 5.6$).

4.1.1 Vignettes of the Factorial Survey Design

Concerning the nature of the factorial survey, following a 2×3 design, each participant was randomly presented with one of six possible scenarios (see the wording of all vignettes in the appendix). In all vignettes, participants first received information that

an HEI offers support measures for students but, due to limited resources, cannot offer those to every student. The institution either deploys 1) an ADM system or 2) university staff (HDM) to analyze student performance data, aiming to assess and predict students' performance.

In the experimental condition with the vignette regarding the use of AI for the distribution of support measures, participants were given a brief explanation of the term AI, in which ADM was explained as a specific form of AI. Next, the use case was introduced in which an ADM system was used to assess and predict student performance based on student performance data. Regarding the use of university staff, the vignette read that human members of the university were in charge of this assessment and the prediction of performance.

Eventually, in each scenario, respondents learned that the university decides to use or not use this information to either 1) allow all students to sign up for support measures irrespective of the available APP outcome (i.e., the *equality* norm), 2) allow only students with a good APP outcome to sign up for support measures (i.e., the *equity* norm), or 3) allow only students with a poor APP outcome to sign up for support measures (i.e., the *need* norm).

4.2 Sample

Participants were recruited with the online access panel (OAP) of the market research institute INNOFACT AG. Survey field time was between January 27 and February 1, 2022.

All in all, 43776 respondents from the OAP were invited to participate in the survey.¹ The questionnaire was accessed 3906 times, and 3680 persons started answering the questionnaire. Of those respondents, 1806 persons were screened out as they were not eligible for our survey as they did not belong to the investigated population of German students. The survey's non-completion rate was 6.1%, and dropouts were equally distributed over all questionnaire pages. Additionally, in the middle of the questionnaire, there was an attention check where respondents were asked to click on a specific scale point ("Please click the 'tend to disagree' item here to confirm your attention."). Altogether, 333 respondents did not answer correctly and were screened out. The final sample consists of 1378 participants.

The average respondent age was 23.22 ($SD = 3.32$). All in all, 1012 (73.4%) respondents identified as women, 358 (26%) as men, and 8 (0.6%) identified as non-binary.

4.3 Measurements

Perceived Distributive Fairness

Assessing the fairness involved respondents rating four statements on a five-point Likert scale ranging from 1 = do not agree at all to 5 = totally agree.

¹ We note that the high number of invitations results from the highly dynamic and often short durations of academic careers of panel members. According to the panel provider, respondent information on educational attainment is regularly updated. However, it cannot be guaranteed that a person who signed up as a student in higher education is necessarily enrolled at the field time of the survey. Hence, invitations were distributed extensively, and the first question was then used for screening purposes.

Table 1 Reliability values of perceived fairness

	Perceived Fairness
alpha	0.89
omega	0.90
omega2	0.90
omega3	0.89
avevar	0.68

The four statements translated from German to English are as follows:

- “A fair balance of interests between different concerns has been achieved.”
- “The distribution of support measures is fair.”
- “No one is unduly disadvantaged by the distribution of support measures.”
- “The distribution of funding measures is just.”

The four indicators suggest good factorial validity (see Table 1).

Study subject

To inquire about the study subject of the respondents, we followed the study subject classification of the German Statistical Office (Destatis, 2021). Respondents could choose up to the second level of the classification from a list of 79 study subjects for brevity. On the first level of the classification, there were eight fields of study, of which we counted four as *STEM*-related (i.e., mathematics and natural sciences; human medicine and health sciences; agricultural, forestry, and nutritional sciences, veterinary medicine; engineering) and five as *SHAPE*-related (humanities; sports; law, economics, and social sciences; arts; as well as the category “other”).

Eventually, based on the indicated study subject, respondents were assigned to either a group that studies a *STEM* subject or a *SHAPE* subject. This procedure results in a dummy variable coded as 0 ‘SHAPE’ and 1 ‘STEM’.

Self-assessment of the individual expected probability that an ADM would assign to one’s own dropout

The self-assessment individual expected probability that an ADM would assign to one’s own dropout was used as a control variable for the analysis of RQ3. Respondents were asked the following question: “Thinking about your own current studies, what do you think is the likelihood that an AI would predict that you would drop out?”. Respondents could indicate a percentage from 0% to 100% using a slider. The average expected probability that an ADM would assign to one’s dropout was 32.54% ($SD = 29.33$). For the analysis and interpretation, the variable was divided by ten, so an increase of one unit reflects an increase of ten percent in a respondent’s self-assessment.

5 Results

To address RQ1 and RQ2, we ran a latent factor analysis for the perceived fairness in each condition, and to address RQ3, we estimated a structural regression model with perceived fairness as the dependent variable. In all analyses, effect coding was used for factor scaling, a procedure that “constrains the set of indicator intercepts to sum to zero for each construct

Table 2 Measurement invariance of perceived fairness

	Chisq (df)	CFI	RMSEA	comp	Chisq diff (df)	CFI diff	RMSEA diff
M1: Configural Invariance	24.32* (12)	1.00	0.07 (0.03–0.11)				
M2: Metric Invariance	58.21* (27)	0.99	0.07 (0.05–0.10)	M1	33.88* (15)	−0.01	0.00
M3: Scalar Invariance	80.76* (42)	0.99	0.06 (0.04–0.08)	M2	22.56 (15)	0.00	−0.01
M4: Residual Invariance	115.65* (62)	0.98	0.06 (0.04–0.08)	M3	34.89* (20)	0.00	0.00

and the set of loadings for a given construct to average 1.0” (Little et al., 2006, p. 62). The estimated latent factor model is scaled like the indicators, which helps with interpretation.

Before assessing the results, it is essential to test for measurement invariance between our analysis groups. A test for measurement invariance suggests strong factorial invariance of the latent factor in each condition of the factorial survey (see Table 2).

5.1 Mean Differences of Perceived Fairness Between the Distribution Norms and Between ADM and HDM

Subsequently, the model was estimated for each of the six experimental groups of the factorial survey, estimating six means. The model shows good fit (χ^2 (62) = 115.65, p = < 0.001; $RMSEA$ = 0.06 CI [0.04, 0.08]; TLI = 0.99).

Perceived fairness of the distribution of support measures for students is highest in the *equality* condition, followed by the *need* condition (see Table 3). In the condition that *equity* is the distributive justice norm of choice, perceived fairness is lowest. Accordingly, the mean difference between *equality* and *need* and *equality* and *equity* suggest a medium-sized significant effect, irrespective of whether APP was based on ADM or HDM. Concerning the mean difference between *equity* and *need*, results suggest a significant difference in the ADM condition but not in the HDM condition.

Concerning the differences in perceived fairness of the distinct distributive justice norms between ADM and HDM, as seen on the right-hand side of Table 3, the results suggest no difference.

5.2 Effect of Study Subject on Perceived Fairness

To address RQ3, we estimated a structural regression model in which the perceived fairness as the dependent variable is regressed on the study subject of a respondent as the independent variable among control variables such as their age, gender, and the expected probability that an ADM would assign to their drop out (i.e., the self-assessed likelihood of AI drop out prediction). For a more accessible overview of the effects in light of the results of RQ1 and RQ2, the condition of whether ADM or HDM performed APP was also included as a dummy variable (0 = ‘ADM’ and 1 = ‘HDM’). Furthermore, a moderator variable was included to assess whether the study subject had a different effect depending on the condition of ADM or HDM (i.e., the scaled and mean-centered product of the two

Table 3 Latent means and mean differences for the perceived fairness

	ADM				HDM				Mean Differences Between ADM & HDM			
	B	SE	95% CI LL	95% CI UL	B	SE	95% CI LL	95% CI UL	Δ Mean	p	Cohen's d	
Equality	3.26	0.05	3.15	3.37	3.31	0.05	3.20	3.41	-0.05	0.541	0.01	
Equity	2.72	0.06	2.61	2.84	2.69	0.06	2.58	2.80	0.03	0.685	0.00	
Need	2.88	0.05	2.78	2.98	2.82	0.06	2.71	2.93	0.06	0.464	-0.04	
Mean Differences Between Norms												
	Δ Mean		p		Cohen's d		Δ Mean		p		Cohen's d	
Equality vs. Equity	0.54		0.000		0.64		0.62		0.000		0.79	
Equality vs. Need	0.38		0.000		0.48		0.48		0.000		0.57	
Equity vs. Need	-0.16		0.044		-0.19		-0.13		0.087		-0.30	

Table 4 Structural regression model for the relationship between study subject and the perceived fairness of support measure distribution

	Equality		Equity		Need	
	Estimate (Std.Err.)	p	Estimate (Std.Err.)	p	Estimate (Std.Err.)	p
Regression Slopes						
Perceived Fairness						
Condition of decision-making: ADM vs. HDM (1 = HDM)	0.04 (0.08)	0.560	-0.05 (0.08)	0.553	-0.05 (0.08)	0.468
Age	-0.00 (0.01)	0.826	-0.03 (0.01)	0.021	-0.01 (0.01)	0.220
Gender (1 = female)	-0.04 (0.08)	0.641	-0.32 (0.09)	0.000	0.04 (0.09)	0.631
Self-assessed likelihood of AI drop out prediction	-0.03 (0.01)	0.026	0.03 (0.01)	0.032	0.02 (0.01)	0.060
Study subject: SHAPE vs. STEM (1 = STEM)	0.02 (0.08)	0.801	0.07 (0.09)	0.384	0.16 (0.08)	0.039
Moderation Effect: 'ADM vs. HDM *'SHAPE vs. STEM'	-0.00 (0.04)	0.980	-0.01 (0.04)	0.900	0.02 (0.04)	0.649
Fit Indices						
χ^2 (df)	112.93 (84)	0.019				
CFI	0.99					
RMSEA	0.03					
RMSEA.CI LOWER	0.01					
RMSEA.CI UPPER	0.04					

dummy variables *study subject* and the respective *condition of decision-making*). The estimated model shows a good fit (see Table 4).

The results suggest that in the *Equality* condition ($B = 0.02$, $SE = 0.08$, $p = 0.8$, $\beta = 0.01$) and the *Equity* condition ($B = 0.07$, $SE = 0.09$, $p = 0.38$, $\beta = 0.04$) the study subject did not affect the perceived fairness of the distribution decision. Only in the *Need* condition there was small significant effect ($B = 0.16$, $SE = 0.08$, $p = 0.04$, $\beta = 0.10$). Whether ADM or HDM was used, the condition did not moderate this effect ($B = 0.02$, $SE = 0.04$, $p = 0.65$, $\beta = 0.02$). Accordingly, students studying a *STEM* subject judged the latter scenario, in which students in need were eligible for support measures as fairer than students studying a *SHAPE* subject. There was no difference in whether ADM or HDM served as a basis for decision-making.

6 Discussion

This study aimed to explore students' perceptions of the deployment of APP as a basis for the distribution of student support measures in an attempt of HEIs to improve student performance and retention using AIEd. The study adds to the body of empirical evidence of fairness perceptions of ADM, an important FAIR ML research field. More specifically, the factorial survey design addressed German students' distributive fairness perceptions of the application of different distributive justice norms (i.e., *equality*, *equity*, and *need*) based on student performance assessments by either ADM or HDM.

The results of the factorial survey show that *equality* appears to be the favored distributive justice norm when it comes to the distribution of student support measures. Compared to the distributive justice norms of *equity* and *need*, students perceive unqualified access to resources as substantially more fair. This is in line with the literature that suggests that in situations where collectives need to distribute limited resources, the low-threshold heuristic suggests to allocate an equal-sized share to each member (Allison & Messick, 1990; Allison et al., 1992; Morand et al., 2020). In addition, it can be assumed that students see themselves as part of a group and, in this sense, are concerned with social cohesion (Cropanzano & Ambrose, 2015). It may also be speculated that some loss aversion of students may contribute to their preference for *equality* as the distribution norm of choice (Smith et al., 2019; Tversky & Kahneman, 1991). After all, students prefer a situation in which everyone—at least in theory—has the same chance to receive support measures, irrespective of support measures being a limited resource that not everyone can draw from. Thus, despite having additional information on the individual's performance from APP, German students remained in this 'fallback option' of equal distribution. As a result, the results of our study are consistent with some previous research demonstrating the preference for altruistic and equal distributions (Lee & Baykal, 2017; Uhde et al., 2020). On the other hand, some disagreement with other studies was revealed that either indicates a preference for proportional distributions, for example, in the context of loan decisions (Saxena et al., 2020), or no preference at all (Schlicker et al., 2021).

As of now, even though decision-makers on the university side have high hopes for the use of AI systems and computer scientists go to great lengths to develop and improve them, the idea of APP using ADM (as well as HDM) does not fall on fertile grounds when proposed to students in higher education. After all, one does not need APP to distribute support measures when *equality* is demanded of the students affected by APP. The same distribution can be ensured by straightforward means, even if, in the end, not everyone

can access and benefit from support measures because of limited resources, especially those students who might well deserve additional assistance according to the *equity* norm or those who could genuinely use it under the *need* norm. Such evidence points to a conflict of interest between more or less well-intentioned aspirations for improving student performance by administrators and the affected students that may not favor the ultimately suggested solution. It remains to be seen whether more detailed information on the goals and effects of APP or other additional communicative measures may improve preferences for distributions that differentiate between students based on their performance. Further research should thus investigate under what specific conditions students may judge the distribution logic based on assessing performance as more or less fair.

Altogether, students were ambivalent regarding the distributive fairness perceptions of the distribution norms of *equity* and *need*. While the reported fairness perception in the *equality* condition was above the center of the scale, for these two conditions, the results suggest a more negative fairness perception below the center of the scale. Accordingly, not only were distribution norms of *equity* and *need* perceived as less fair than *equality* but their overall assessment suggests that they were deemed somewhat unfair. From this finding, one might conclude that those situations are evaluated more negatively, in which some students that—based on a distributive justice norm—are deemed worthy of support benefit more than all others. However, it remains to be seen whether such fairness perceptions will change if, for example, a combination of different norms focusing on the most effective distribution of support measures or fine-tuning APP to suggest only those measures with the highest individual payoff in terms of increasing student performance are implemented.

Moreover, results show that students are similar in their fairness perception when contrasting HDM and ADM. As it is usually suggested within the literature on AIED that there is a qualitative difference between humans and AI systems concerning the assessment of vast amounts of input data, our findings suggest that students do not draw a distinction when it comes to their perceived fairness. This finding raises the question, *why* students are the same in their assessment. In aggregate, do the individual strengths and weaknesses of the respective approach employing HDM and ADM even out so that they are both eventually perceived as equally fair or unfair? Moreover, what are the specific drivers of perceptions of (un)fairness concerning APP if this were the case? Are the underlying preferences of students concerning the application of systematic student performance assessment opposed to the introduction of further means of bureaucratic management and organizational control?

While the inclusion of sociodemographic variables attests that there are inconclusive effects of individual characteristics on perceived fairness, the presumed central predictor of the students' study subject suggests that the field of study has hardly any effect on fairness perceptions. Only in the *need* condition did *STEM* students perceive the proposed distribution as more fair than students from *SHAPE* studies. However, this effect was not moderated by whether humans or algorithms were the basis of decision-making. While we can only speculate about reasons for this effect limited to the norm of *need*, the absence of the effects of interest suggests that there is no systematic difference with respect to the field of study when it comes to APP in general.

The study adds to the literature in two important regards. First, it contributes empirically to the understanding of how students perceive different distributive justice norms, particularly focusing on the application of ADM and HDM in distributing student support measures. This addresses a gap in existing research, uncovering nuanced insights into the preference for equality over other norms like equity and need in the context of higher education. Second, the study challenges commonly held assumptions about the potential for

technological solutions like APP to enhance fairness in educational support allocation. It emphasizes the complexity of human perception and the importance of social cohesion, altruism, and the potential conflict of interest between administrative aspirations and student preferences. These findings not only underline the need for careful consideration and transparent communication in implementing AI systems in education but also pave the way for future research exploring how to align technological advancements with societal values and expectations.

Limitations of our study include the choice of the empirical design and the sampling of students. First, using vignettes to illustrate the possible scenario of the deployment of APP comes with limited room for the depiction of a usually detailed sociotechnical distribution process. It may also be detached from the current study situation of the surveyed students at their respective universities. After all, only three distinct distributive justice norms were used for illustration. The systems and the decision-making processes HEIs deploy are more varied and elaborate than can be easily depicted and understood in a brief written paragraph within the factorial design. Further research should thus use an actual APP system and give a vivid display of its inner procedural workings, especially in inviting students' evaluation of ADM. Moreover, non-standardized empirical approaches may prove helpful in assessing the nuances and points of reference of students' fairness perceptions more closely than our standardized design and measure of fairness perceptions. Second, the sample was limited to German students and was skewed towards students who identified as female. While this sample bias is less of an issue concerning the experimental logic within a factorial survey design, we suggest that additional studies broaden the scope and include a more diverse set of students with a special focus on cross-country comparisons, considering the diverse academic systems and institutions across nations. Such an approach is critical given that previous research on FAIR ML has focused almost exclusively on study samples from Western countries (Starke et al., 2022). As delineated by Lee et al. (2017), these cultural subtleties can be discerned when contrasting American and Indian respondents regarding their distinct fairness perceptions. Furthermore, the heterogeneous expectations of students, shaped by the differing academic systems and sociocultural contexts they hail from, warrant further exploration to ensure the comprehensive applicability of any model or application.

7 Conclusion and Implications

The deployment of AIED is an ongoing process that aims to introduce sociotechnical AI systems such as ADM for APP within HEIs. Consequently, it may change the social fabric of educational organizations and the life trajectories of generations of students. While the intentions to use APP to improve student performance and retention arguably appear understandable from the perspective of university administrators, it is suggested to also focus on the people affected by introducing sociotechnical AI systems. Research on the social consequences and evaluation of the affected students thus deems necessary, and our study answers the call for a deeper understanding of the distributive fairness perceptions of distinct norms of distributive justice. Our findings show that German students prefer a distribution of support measures based on an APP system according to the *equality* norm. In this regard, there are no differences in fairness perceptions of an APP comparing HDM and ADM and concerning the students' study subject.

Eventually, the results of our study link to the literature on AIEd and thus contribute essential findings to the research on perceptions of distributive justice in this field. Consequently, these findings have implications for HEIs and administrators and future research in this area. In light of evidence that students prefer a distribution according to the *equality* norm when distributing support measures based on an algorithmic APP system, the question arises as to whether it makes sense to use an APP system at all since support measures that are—in theory—intended to be open to all students do not require particular classification by an algorithm in the first place. Should individual universities nevertheless decide to use an APP system to benefit from its presumed potential—for example, to enable individual feedback or to increase the graduation rates (Attaran et al., 2018; Mah, 2016)—this use of AIEd could evoke students' disapproval. In light of potential negative consequences for the university's reputation with students and student retention (Marcinkowski et al., 2020), the pros and cons of using APP must be carefully weighed. Should universities make plans to introduce APP, we suggest that they conduct research that investigates the evaluation of APP by its current and future student body. This evaluation may take the form of an ongoing consultation process with all stakeholders within the university to navigate potential fairness issues and their respective contextuality.

Further research in the field of AIEd, particularly regarding the deployment of APP within HEIs, beckons multidimensional exploration. Firstly, studies should delve deeper into understanding the various norms of distributive justice across diverse student populations. While our study focused on German students' preferences towards the equality norm, it remains to be seen if such preferences resonate similarly across different cultural and educational contexts. Moreover, as suggested above, the practical ramifications of these findings for universities worldwide need empirical substantiation. A targeted examination of whether supposedly innovative APP systems align with the core ethos and strategic visions of HEIs could provide more clarity on the matter. Similarly, the potential for negative repercussions for universities, in terms of reputation and student retention, stemming from student perceptions of AIEd, necessitates a comprehensive risk assessment. Lastly, a co-creative approach, wherein students and other key stakeholders are involved in APP deployment's developmental and decision-making processes, might offer a blueprint for future AIEd endeavors. Investigations into such participatory methodologies, gauging their feasibility, efficiency, and efficacy, could shape the future discourse on AI in education.

Appendix

Vignette of the Factorial Survey

Text Vignette ADM - German Original

“Derzeit wird in der Öffentlichkeit viel über, Künstliche Intelligenz‘ (KI) gesprochen. Gemeint sind hier Computeranwendungen, die automatisiert digitale Daten auswerten. Die Auswertung großer Datenmengen stellt für KI einen Lernprozess dar, aus dem sie laufend neue Informationen verarbeitet und so im Zeitverlauf immer präzisere Muster erkennt. Auf Grundlage dieser Analyse können Sachverhalte festgestellt sowie zukünftige Entwicklungen prognostiziert werden. Systeme mit Künstlicher Intelligenz können dem Menschen Handlungsempfehlungen vorschlagen oder autonom Entscheidungen treffen und diese auch

direkt ausführen. KI gewinnt mit der Zeit auch an Hochschulen immer mehr an Bedeutung und Relevanz.

Versetzen Sie sich bitte einmal in die folgende Situation, zu der wir Ihnen auf der nächsten Seite ein paar Fragen stellen möchten: Eine Hochschule bietet Ihren Studierenden ein Angebot an Fördermaßnahmen, kann jedoch, aufgrund begrenzter Ressourcen, nicht allen Studierenden eine Teilnahme ermöglichen.

Eine KI-Anwendung sichtet und analysiert systematisch alle vorliegenden Leistungsdaten (z. B. Prüfungsergebnisse) der Studierenden. Auf dieser Basis erstellt die KI eine Leistungsvorhersage für alle Studierenden.

Vor diesem Hintergrund trifft die Hochschule folgende Entscheidung für die Verteilung der Fördermaßnahmen:

Text Vignette HDM - German Original

“Versetzen Sie sich bitte einmal in die folgende Situation, zu der wir Ihnen auf der nächsten Seite ein paar Fragen stellen möchten: Eine Hochschule bietet Ihren Studierenden ein Angebot an Fördermaßnahmen, kann jedoch, aufgrund begrenzter Ressourcen, nicht allen Studierenden eine Teilnahme ermöglichen.

Beschäftigte der Hochschule sichten und analysieren systematisch alle vorliegenden Leistungsdaten (z. B. Prüfungsergebnisse) der Studierenden. Auf dieser Basis erstellen die Hochschulbeschäftigten Leistungsvorhersagen für die Studierenden.

Vor diesem Hintergrund trifft die Hochschule folgende Entscheidung für die Verteilung der Fördermaßnahmen:

Vignette Addendum - Group Equality - German Original

“Alle Studierenden können sich für Fördermaßnahmen anmelden, ungeachtet des Ergebnisses der Vorhersage der KI.” (emphasis in original)

Vignette Addendum - Group Equity - German Original

“Auf Grundlage der Vorhersage der KI, können nur Studierende, die eine gute Leistungsvorhersage erhalten haben, sich für das Angebot an Fördermaßnahmen anmelden.” (emphasis in original)

Vignette Addendum - Group Need - German Original

“Auf Grundlage der Vorhersage der KI, können nur Studierende, die eine schlechte Leistungsvorhersage erhalten haben, sich für das Angebot an Fördermaßnahmen anmelden.” (emphasis in original)

Text Vignette ADM - English Translation

“Currently, there is a lot of public talk about ‘artificial intelligence’ (AI). This refers to computer applications that automatically evaluate digital data. The evaluation of large volumes of data represents a learning process for AI, from which it continuously processes new information and thus recognizes increasingly precise patterns over time. Based on this analysis, facts can be established and future developments can be predicted. Systems with

artificial intelligence can suggest courses of action to humans or make decisions autonomously and also execute them directly. Over time, AI is also becoming increasingly important and relevant at universities.

Please put yourself in the following situation, about which we would like to ask you a few questions on the next page: A university offers its students a range of support measures but, due to limited resources, cannot allow all students to participate.

An AI application systematically sifts through and analyzes all available student performance data (e.g., exam results). On this basis, the AI generates a performance prediction for all students. Against this background, the university makes the following decision for the distribution of support measures:”

Text Vignette HDM - English Translation

“Please put yourself in the following situation, about which we would like to ask you a few questions on the next page: A university offers its students a range of support measures but, due to limited resources, cannot enable all students to participate.

University employees systematically review and analyze all available student performance data (e.g., exam results). Based on this data, the university employees create performance predictions for the students.

With this in mind, the university makes the following decision for the distribution of funding:”

Vignette Addendum - Group Equality - English Translation

“All students are eligible to apply for remedial measures, regardless of the outcome of the AI’s prediction.” (emphasis in original)

Vignette Addendum - Group Equity - English Translation

“Based on the prediction of the AI, only students who received a good prediction of performance can register for remedial services.” (emphasis in original)

Vignette Addendum - Group Need - English Translation

“Based on the AI’s prediction, only students who received a poor performance prediction can enroll in the range of remedial services.” (emphasis in original)

Acknowledgements The authors thank Albina Maxhuni for her support and feedback on the first version of the paper, as well as Cara Schön Müller and the students of the teaching research project *Predicting student performance through artificial intelligence: what are the perceptions and attitudes of the affected students?* at Heinrich Heine University in the winter term 2021/2022.

Funding Open Access funding enabled and organized by Projekt DEAL. This study was conducted as part of the project Responsible Academic Performance Prediction (RAPP) project. The project is funded by the German Federal Ministry of Education and Research [grant number 16DHB4020].

Data Availability The data and code for data analysis can be accessed via the project’s Open Science Foundation repository: (<https://osf.io/djt39/>).

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Adams, J. S. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology*, 2, 267–299. [https://doi.org/10.1016/S0065-2601\(08\)60108-2](https://doi.org/10.1016/S0065-2601(08)60108-2)
- Adams Becker, S., Cummins, M., Davis, A., Freeman, A., Hall Giesinger, C., & Ananthanarayanan, V. (2017). NMC Horizon Report: 2017 Higher Education Edition. Austin, Texas: The New Media Consortium. <https://www.sconul.ac.uk/sites/default/files/documents/2017-nmc-horizon-report-he-EN.pdf>.
- Adams, Richard, & McIntyre, Niamh. (2020). England a-Level Downgrades Hit Pupils from Disadvantaged Areas Hardest: Analysis Also Shows Pupils at Private Schools Benefited Most from Algorithm. *The Guardian*. <https://www.theguardian.com/education/2020/aug/13/england-a-level-downgrades-hit-pupils-from-disadvantaged-areas-hardest#>.
- Ahmed, SA., & Khan, SI. (2019). A machine learning approach to predict the engineering students at risk of dropout and factors behind: bangladesh perspective. In *10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1–6). <https://doi.org/10.1109/ICCCNT.2019.8944511>.
- Allison, S. T., & Messick, D. M. (1990). Social decision heuristics in the use of shared resources. *Journal of Behavioral Decision Making*, 3(3), 195–204. <https://doi.org/10.1002/bdm.3960030304>
- Allison, S. T., McQueen, L. R., & Schaeffl, L. M. (1992). Social decision making processes and the equal partitionment of shared resources. *Journal of Experimental Social Psychology*, 28(1), 23–42. [https://doi.org/10.1016/0022-1031\(92\)90030-N](https://doi.org/10.1016/0022-1031(92)90030-N)
- Alturki, S., Hulpuş, I., & Stuckenschmidt, H. (2022). Predicting academic outcomes: A survey from 2007 till 2018. *Technology, Knowledge and Learning*, 27(1), 275–307. <https://doi.org/10.1007/s10758-020-09476-0>
- Alyahyan, E., & Düstegör, D. (2020). Predicting academic success in higher education: Literature review and best practices: Literature review and best Practices. *International Journal of Educational Technology in Higher Education*, 17(1), 1–21. <https://doi.org/10.1186/s41239-020-0177-7>
- Ananny, M., & Crawford, K. (2018). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *New Media & Society*, 20(3), 973–989. <https://doi.org/10.1177/1461444816676645>
- Angwin, J., Larson, J., Mattu, S., & Kirchner, L. (2016). Machine bias: There's software used across the country to predict future criminals. And It's Biased Against Blacks. ProPublica. <https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing>.
- Araujo, T., Helberger, N., Kruijkemeier, S., & de Vreese, C. H. (2020). Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability. *AI & Society*, 35(3), 611–23. <https://doi.org/10.1007/s00146-019-00931-w>
- Arcidiacono, P., Hotz, V. J., & Kang, S. (2012). Modeling college major choices using elicited measures of expectations and counterfactuals. *Journal of Econometrics*, 166(1), 3–16. <https://doi.org/10.3386/w15729>
- Armatas, C., Kwong, T., Chun, C., Spratt, C., Chan, D., & Kwan, J. (2022). Learning analytics for programme review: Evidence, analysis, and action to improve student learning outcomes. *Technology, Knowledge and Learning*, 27(2), 461–478. <https://doi.org/10.1007/s10758-021-09559-6>
- Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. In S. Dawson, C. Haythornthwaite, S. Buckingham Shum, D. Gašević, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK 2012)* (pp. 267–270). <https://doi.org/10.1145/2330601.2330666>.

- Asaria, M., Costa-Font, J., & Cowell, F. (2023). How does exposure to COVID-19 influence health and income inequality aversion? *Social Choice and Welfare*, 61, 625–647. <https://doi.org/10.1007/s00355-023-01460-8>
- Attaran, M., Stark, J., & Stotler, D. (2018). Opportunities and challenges for big data analytics in US higher education. *Industry and Higher Education*, 32(3), 169–182. <https://doi.org/10.1177/0950422218770937>
- Bai, B., Dai, H., Zhang, D., Zhang, F., & Hu, H. (2021). The impacts of algorithmic work assignment on fairness perceptions and productivity: Evidence from field experiments. In *Academy of Management Proceedings: 81st Annual Meeting of the Academy of Management 2021: Bringing the Manager Back in Management, AoM 2021*, edited by Sonia Taneja, 2021:12335. <https://doi.org/10.5465/AMBPP.2021.175>
- Bai, B., Dai, H., Zhang, D. J., Zhang, F., & Hu, H. (2022). The Impacts of Algorithmic Work Assignment on Fairness Perceptions and Productivity: Evidence from Field Experiments. *Manufacturing & Service Operations Management*, 24(6), 3060–3078. <https://doi.org/10.1287/msom.2022.1120>
- Baker, R. S., & Hawn, A. (2022). Algorithmic Bias in education. *International Journal of Artificial Intelligence in Education*, 32(4), 1052–1092. <https://doi.org/10.1007/s40593-021-00285-9>
- Berens, J., Schneider, K., Gortz, S., Oster, S., & Burghoff, J. (2019). Early detection of students at risk - predicting student dropouts using administrative student data from German universities and machine learning methods. *Journal of Educational Data Mining*, 11(3), 1–41. <https://doi.org/10.5281/zenodo.3594771>
- Bettinger, E. P., & Baker, Rachel B. (2014). The effects of student coaching: An evaluation of a randomized experiment in student advising. *Educational Evaluation and Policy Analysis*, 36(1), 3–19. <https://doi.org/10.3102/0162373713500523>
- Boyd, D., & Crawford, K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, 15(5), 662–679. <https://doi.org/10.1080/1369118X.2012.678878>
- Brade, R., Himmeler, O., & Jäckle, R. (2018). Normatively framed relative performance feedback – Field experiment and replication. Munich: MPRA Paper No. 88830. <https://mpra.ub.uni-muenchen.de/88830/>
- Burton, J. W., Stein, M.-K., & Blegind Jensen, T. (2020). A systematic review of algorithm aversion in augmented decision making. *Journal of Behavioral Decision Making*, 33(2), 220–239. <https://doi.org/10.1002/bdm.2155>
- Busuioc, M. (2021). Accountable artificial intelligence: Holding algorithms to account: Holding algorithms to account. *Public Administration Review*, 81(5), 825–836. <https://doi.org/10.1111/puar.13293>
- Cheng, H-F, Stapleton, L., Wang, R., Bullock, P., Chouldechova, A., Wu, ZSS., & Zhu, H. (2021). Soliciting stakeholders' fairness notions in child maltreatment predictive systems. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*, edited by Yoshifumi Kitamura, Aaron Quigley, Katherine Isbister, Takeo Igarashi, Pernille Bjørn, and Steven Drucker, pp. 1–17. <https://doi.org/10.1145/3411764.3445308>
- Coenen, J., Borghans, L., & Diris, R. (2021). Personality traits, preferences and educational choices: A focus on STEM. *Journal of Economic Psychology*, 84, 1–16. <https://doi.org/10.1016/j.joep.2021.102361>
- Cropanzano, R. S., & Ambrose, M. L. (2015). The Oxford handbook of justice in the workplace. *Oxford University Press*. <https://doi.org/10.1093/oxfordhb/9780199981410.001.0001>
- Cropanzano, R., Rupp, D. E., Mohler, C. J., & Schminke, M. (2001). Three Roads to Organizational Justice. In G. R. Ferris (Ed.), *Research in Personnel and Human Resources Management* (Vol. 20, pp. 1–113). Emerald Group Publishing Limited. [https://doi.org/10.1016/S0742-7301\(01\)20001-2](https://doi.org/10.1016/S0742-7301(01)20001-2)
- Dai, J., Fazelpour, S., & Lipton, Z. (2021). “Fair Machine Learning Under Partial Compliance.” In M. Fourcade, B. Kuipers, S. Lazar, and D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AI/ES '21)* (pp. 55–65). <https://doi.org/10.1145/3461702.3462521>
- Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, 46(5), 904–920. <https://doi.org/10.1111/bjet.12230>
- de Cremer, D., & McGuire, J. (2022). Human-algorithm collaboration works best if humans lead (Because It Is Fair!). *Social Justice Research*, 35(1), 53–55. <https://doi.org/10.1007/s11211-021-00382-z>
- de Fine Licht, K., & de Fine Licht, J. (2020). Artificial intelligence, transparency, and public decision-making: Why explanations are key when trying to produce perceived legitimacy. *AI & Society*, 35(4), 917–926. <https://doi.org/10.1007/s00146-020-00960-w>
- Gomes de Sousa, W., Pereira de Melo, E. R., de Souza Bermejo, P. H., Sousa Farias, R. A., & Oliveira Gomes, A. (2019). How and where is artificial intelligence in the public sector going? A literature review and research agenda. *Government Information Quarterly*, 36(4), 1–14. <https://doi.org/10.1016/j.giq.2019.07.004>

- Destatis (2021). Bildung Und Kultur: Studierende an Hochschule. Statistisches Bundesamt (Ed.), (Fachserie 11, Reihe 4.1). <https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Bildung-Forschung-Kultur/Hochschulen/Publikationen/Downloads-Hochschulen/studierende-hochschulen-endg-2110410217004.html>.
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues*, 31(3), 137–149. <https://doi.org/10.1111/j.1540-4560.1975.tb01000.x>
- Diakopoulos, N. (2016). Accountability in algorithmic decision making. *Communications of the ACM*, 59(2), 56–62. <https://doi.org/10.1145/2844110>
- Dietvorst, B. J., & Bharti, S. (2020). People reject algorithms in uncertain decision domains because they have diminishing sensitivity to forecasting error. *Psychological Science*, 31(10), 1302–1314. <https://doi.org/10.1177/0956797620948841>
- Dietvorst, B. J., Simmons, J. P., & Massey, C. (2015). Algorithm aversion: People erroneously avoid algorithms after seeing them Err. *Journal of Experimental Psychology: General*, 144(1), 114–126. <https://doi.org/10.1037/xge0000033>
- Dwork, C., Hardt, M., Pitassi, T., Reingold, O., & Zemel, R. (2012). Fairness through awareness. In S. Goldwasser (Ed.), *Proceedings of the 3rd Innovations in Theoretical Computer Science Conference on - ITCS '12* (pp. 214–226). <https://doi.org/10.1145/2090236.2090255>.
- Edwards, C. (2021). Let the algorithm decide? *Communications of the ACM*, 64(6), 21–22. <https://doi.org/10.1145/3460216>
- Ekowo, M., & Palmer, I. (2016). The promise and peril of predictive analytics in higher education: A landscape analysis.” <https://www.newamerica.org/education-policy/policy-papers/promise-and-peril-predictive-analytics-higher-education/>.
- European Commission. (2019). Ethics guidelines for trustworthy AI. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai-trustworthy-ai>.
- Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using g*power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, 41(4), 1149–1160. <https://doi.org/10.3758/BRM.41.4.1149>
- Fazelpour, S., & Danks, D. (2021). Algorithmic bias: Senses, sources, solutions. *Philosophy Compass*, 16(8), 1–16. <https://doi.org/10.1111/phc3.12760>
- Folger, R., & Cropanzano, R. (2001). Fairness theory: Justice as accountability. In J. Greenberg & R. Cropanzano (Eds.), *Advances in Organizational Justice* (pp. 3–55). Stanford: Stanford University Press.
- Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2021). The (Im)possibility of fairness. *Communications of the ACM*, 64(4), 136–143. <https://doi.org/10.1145/3433949>
- Gibson, J. P., & Mourad, T. (2018). The growing importance of data literacy in life science education. *American Journal of Botany*, 105(12), 1953–1956. <https://doi.org/10.1002/ajb2.1195>
- Greenberg, J. (1987). A taxonomy of organizational justice theories. *Academy of Management Review*, 12(1), 9–22. <https://doi.org/10.5465/amr.1987.4306437>
- Greenberg, J. (1990). Organizational justice: Yesterday, today, and tomorrow. *Journal of Management*, 16(2), 399–432. <https://doi.org/10.1177/014920639001600208>
- Greenberg, J. (1993). The social side of fairness: Interpersonal and informational classes of organizational justice. In R. Cropanzano (Ed.), *Justice in the Workplace* (pp. 79–103). Lawrence Erlbaum Associates Inc.
- Hagendorff, T., & Wezel, K. (2020). 15 challenges for AI: Or What AI (Currently) Can't Do. *AI & Society*, 35(2), 355–365. <https://doi.org/10.1007/s00146-019-00886-y>
- Hannan, J., Chen, H.-Y. W., & Joseph, K. (2021). Who gets what, according to whom? An analysis of fairness perceptions in service allocation. In M. Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.) *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES '21)* (pp. 555–565). <https://doi.org/10.1145/3461702.3462568>.
- Harrison, G., Hanson, J., Jacinto, C., Ramirez, J., & Ur, B. (2020). An Empirical Study on the Perceived Fairness of Realistic, Imperfect Machine Learning Models. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor, & G. Zanfir-Fortuna (Eds.) *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)*, (pp. 392–402). <https://doi.org/10.1145/3351095.3372831>.
- Hausmann, L. R. M., Ye, F., Schofield, J. W., & Woods, R. L. (2009). Sense of belonging and persistence in white and African American first-year students. *Research in Higher Education*, 50(7), 649–669. <https://doi.org/10.1007/s11162-009-9137-8>
- Helberger, N., Araujo, T., & de Vreese, C. H. (2020). Who is the fairest of them all? Public attitudes and expectations regarding automated decision-making: Public attitudes and expectations regarding

- automated decision-making. *Computer Law & Security Review*, 39, 1–16. <https://doi.org/10.1016/j.clsr.2020.105456>
- Holzinger, A. (2016). Interactive machine learning for health informatics: When do we need the human-in-the-loop? *Brain Informatics*, 3(2), 119–131. <https://doi.org/10.1007/s40708-016-0042-6>
- Howell, J. A., Roberts, L. D., Seaman, K., & Gibson, D. C. (2018). Are we on our way to becoming a ‘helicopter university’? Academics’ views on learning analytics. *Technology, Knowledge and Learning*, 23(1), 1–20. <https://doi.org/10.1007/s10758-017-9329-9>
- Hsu, S., Li, T.W., Zhang, Z., Fowler, M., Zilles, C., & Karahalios, K. (2021). Attitudes surrounding an imperfect AI autograder. In Y. Kitamura, A. Quigley, K. Isbister, T. Igarashi, P. Björn, & S. Drucker (Eds.), *Proceedings of the 2021 chi conference on human factors in computing systems (CHI ’21)* (pp. 1–15). <https://doi.org/10.1145/3411764.3445424>
- Humburg, M. (2017). Personality and field of study choice in university. *Education Economics*, 25(4), 366–378. <https://doi.org/10.1080/09645292.2017.1282426>
- Huppert, E., Cowell, J. M., Cheng, Y., Contreras-Ibáñez, C., Gomez-Sicard, N., Gonzalez-Gadea, M. L., Huepe, D., et al. (2019). The development of children’s preferences for equality and equity across 13 individualistic and collectivist cultures. *Developmental Science*, 22(2), 1–16. <https://doi.org/10.1111/desc.12729>
- Jang, Y., Choi, S., & Kim, H. (2022). Development and validation of an instrument to measure undergraduate students’ attitudes toward the ethics of artificial intelligence (AT-EAI) and analysis of its difference by gender and experience of AI education. *Education and Information Technologies*, 27(8), 11635–11667. <https://doi.org/10.1007/s10639-022-11086-5>
- Jiang, W., & Pardos, Z. A. (2021). Towards equity and algorithmic fairness in student grade prediction. In M. Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society (AIES ’21)* (pp. 608–617). <https://doi.org/10.1145/3461702.3462623>
- Jones, M., Molly, D. A., & Lahiri, A. (2020). Shape the future: How the social sciences, humanities and the arts can SHAPE a positive, post-pandemic future for peoples, economies and environments. *Journal of the British Academy*, 8, 167–266. <https://doi.org/10.5871/jba/008.167>
- Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2019). SemTools: Useful tools for structural equation modeling. R Package Version 0.5-2.” <https://CRAN.R-project.org/package=semTools>.
- Kaibel, C., Koch-Bayram, I., Biemann, T., & Mühlenbock, M. (2019). Applicant perceptions of hiring algorithms - Uniqueness and discrimination experiences as moderators. *Academy of Management Proceedings*, 2019(1), 18172. <https://doi.org/10.5465/AMBPP.2019.210>
- Kasinidou, M., Kleanthous, S., Barlas, P., & Otterbacher, J. (2021). I agree with the decision, but they didn’t deserve this: Future developers’ perception of fairness in algorithmic decisions. In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 690–700). <https://doi.org/10.1145/3442188.3445931>
- Keller, B., Lünich, M., & Marcinkowski, F. (2022). How is socially responsible academic performance prediction possible? Insights from a Concept of Perceived AI Fairness. In F. Almaraz-Menéndez, A. Maz-Machado, C. López-Esteban, & C. Almaraz-López (Eds.), *Strategy, Policy, Practice, and Governance for AI in Higher Education Institutions* (pp. 126–155). IGI Global. <https://doi.org/10.4018/978-1-7998-9247-2.ch006>
- Kemper, L., Vorhoff, G., & Wigger, B. U. (2020). Predicting student dropout: A machine learning approach. *European Journal of Higher Education*, 10(1), 28–47. <https://doi.org/10.1080/21568235.2020.1718520>
- Knox, J., Hoel, T., & Yuan, L. (2022). From principles to processes: Lessons from higher education from the development of AI ethics. In F. Almaraz-Menéndez, A. Maz-Machado, C. López-Esteban, & C. Almaraz-López (Eds.), *Strategy, Policy, Practice, and Governance for AI in Higher Education Institutions* (pp. 101–125). IGI Global. <https://doi.org/10.4018/978-1-7998-9247-2.ch005>
- König, P. D., & Wenzelburger, G. (2022). Between technochauvinism and human-centrism: Can algorithms improve decision-making in democratic politics? *European Political Science*, 21(1), 132–149. <https://doi.org/10.1057/s41304-020-00298-3>
- Chiteng Kot, F. (2014). The impact of centralized advising on first-year academic performance and second-year enrollment behavior. *Research in Higher Education*, 55(6), 527–563. <https://doi.org/10.1007/s11162-013-9325-4>
- Kotsiantis, S. B. (2012). Use of machine learning techniques for educational proposes: A decision support system for forecasting students’ grades. *Artificial Intelligence Review*, 37(4), 331–344. <https://doi.org/10.1007/s10462-011-9234-x>

- Kusner, M., Loftus, J., Russell, C., & Silva, R. (2017). Counterfactual Fairness. In U. von Luxburg, I. Guyon, S. Bengio, H. Wallach, R. Fergus, S. V. N. Vishwanathan, & R. Garnett (Eds.), *NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 1–11).
- Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. *Big Data & Society*, 5(1), 1–16. <https://doi.org/10.1177/2053951718756684>
- Lee, M. K., & Baykal, S. (2017). Algorithmic mediation in group decisions: Fairness perceptions of algorithmically mediated Vs. discussion-based social division. In *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing* (pp. 1035–1048). <https://doi.org/10.1145/2998181.2998230>.
- Lee, M. K., & Rich, K. (2021). Who is included in human perceptions of AI?: Trust and perceived fairness around healthcare AI and cultural mistrust: Trust and perceived fairness around healthcare AI and cultural mistrust. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). <https://doi.org/10.1145/3411764.3445570>.
- Lee, M. K., Kim, J. T., & Lizarondo, L. (2017). A human-centered approach to algorithmic services: Considerations for fair and motivating smart community service management that allocates donations to non-profit organizations. In G. Mark, S. Fussell, C. Lampe, M. C. Schraefel, J. P. Hourcade, C. Appert, & D. Wigdor (Eds.), *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (pp. 3365–3376). <https://doi.org/10.1145/3025453.3025884>.
- Lee, M. K., Jain, A., Cha, H. J., Ojha, S., & Kusbit, D. (2019). Procedural justice in algorithmic fairness: Leveraging transparency and outcome control for fair algorithmic mediation. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–26. <https://doi.org/10.1145/3359284>
- Lee, M. K., Kusbit, D., Kahng, A., Kim, J. T., Yuan, X., Chan, A., See, D., et al. (2019). WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, 3, 1–35. <https://doi.org/10.1145/3359283>
- Legault, L., Green-Demers, I., & Pelletier, L. (2006). Why do high school students lack motivation in the classroom? Toward an understanding of academic amotivation and the role of social support. *Journal of Educational Psychology*, 98(3), 567–582. <https://doi.org/10.1037/0022-0663.98.3.567>
- Little, T. D., Slegers, D. W., & Card, N. A. (2006). A non-arbitrary method of identifying and scaling latent variables in SEM and MACS models. *Structural Equation Modeling: A Multidisciplinary Journal*, 13(1), 59–72. https://doi.org/10.1207/s15328007sem1301_3
- Logg, J. M., Minson, J. A., & Moore, D. A. (2019). Algorithm appreciation: People prefer algorithmic to human judgment. *Organizational Behavior and Human Decision Processes*, 151, 90–103. <https://doi.org/10.1016/j.obhdp.2018.12.005>
- Lünich, M., & Kieslich, K. (2022). Exploring the roles of trust and social group preference on the legitimacy of algorithmic decision-making Vs. Human decision-making for allocating COVID-19 vaccinations. *AI & SOCIETY*, 1–19. <https://doi.org/10.1007/s00146-022-01412-3>
- Mah, D.-K. (2016). Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning*, 21(3), 285–305. <https://doi.org/10.1007/s10758-016-9286-8>
- Makhlouf, K., Zhioua, S., & Palamidessi, C. (2021). Machine learning fairness notions: Bridging the gap with real-world applications. *Information Processing & Management*, 58(5), 1–32. <https://doi.org/10.1016/j.ipm.2021.102642>
- Mäntymäki, M., Minkinen, M., Birkstedt, T., & Viljanen, M. (2022). Defining organizational AI governance. *AI and Ethics*, 2(4), 603–609. <https://doi.org/10.1007/s43681-022-00143-x>
- Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). Implications of AI (Un-)fairness in Higher Education Admissions: The Effects of Perceived AI (Un-)Fairness on Exit, Voice and Organizational Reputation. In M. Hildebrandt, C. Castillo, E. Celis, S. Ruggieri, L. Taylor, & G. Zanfir-Fortuna (Eds.), *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency (FAT* '20)* (pp. 122–130). <https://doi.org/10.1145/3351095.3372867>.
- Martinez Neda, B., Zeng, Y., & Gago-Masague, S. (2021). Using machine learning in admissions: Reducing human and algorithmic bias in the selection process. In M. Sherrieff, L. D. Merkle, P. Cutter, A. Monge, & J. Sheard (Eds.), *Proceedings of the 52nd ACM Technical Symposium on Computer Science Education (SIGCSE '21)* (p. 1323). <https://doi.org/10.1145/3408877.3439664>.
- Morand, D., Merriman, K., & Deckop, J. (2020). Making sense of egalitarian pay: Reconceptualising its treatment in organisations. *Human Resource Management Journal*, 30(2), 180–197. <https://doi.org/10.1111/1748-8583.12264>
- Morosanu, L., Handley, K., & O'Donovan, B. (2010). Seeking support: Researching first-year students' experiences of coping with academic life. *Higher Education Research & Development*, 29(6), 665–678. <https://doi.org/10.1080/07294360.2010.487200>

- Muñoz, C., Megan, S., & Patil, D. J. (2016). Big data: A report on algorithmic systems, opportunity, and civil rights. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/2016_0504_data_discrimination.pdf.
- Murphy, S. P. (2011). Fairness. In D. K. Chatterjee (Ed.), *Encyclopedia of Global Justice* (pp. 336–340). Springer Netherlands.
- Neugebauer, M., Heublein, U., & Daniel, A. (2019). Studienabbruch in Deutschland: Ausmaß, Ursachen, Folgen, Präventionsmöglichkeiten. *Zeitschrift für Erziehungswissenschaft*, 22(5), 1025–1046. <https://doi.org/10.1007/s11618-019-00904-1>
- Nguyen, M.-T., Tran-Tien, M., Viet, A. P., Vu, H.-T., & Nguyen, V.-H. (2021). Building a chatbot for supporting the admission of universities. In T. Theeramunkong (Ed.), *Proceedings of 2021 13th International Conference Knowledge and Systems Engineering (KSE 2021)* (pp. 1–6). <https://doi.org/10.1109/KSE53942.2021.9648677>.
- Nyarko, J., Goel, S., & Sommers, R. (2021). Breaking taboos in fair machine learning: An experimental study: An experimental study. In *Equity and Access in Algorithms, Mechanisms, and Optimization (EAAMO '21)* (pp. 1–11). <https://doi.org/10.1145/3465416.3483291>.
- Olaya, D., Vázquez, J., Maldonado, S., Miranda, J., & Verbeke, W. (2020). Uplift Modeling for preventing student dropout in higher education. *Decision Support Systems*, 134, 1–11. <https://doi.org/10.1016/j.dss.2020.113320>
- Ornelas, F., & Ordóñez, C. (2017). Predicting student success: A Naïve Bayesian application to community college data. *Technology, Knowledge and Learning*, 22(3), 299–315. <https://doi.org/10.1007/s10758-017-9334-z>
- Orwat, C. (2020). *Risks of Discrimination Through the Use of Algorithms: A Study Compiled with a Grant from the Federal Anti-Discrimination Agency*. Berlin, Germany: Federal Anti-Discrimination Agency.
- Pelletier, K., Brown, M., Brooks, D. C., McCormack, M., Reeves, J., Arbino, N., Bozkurt, A et al. (2021). 2021 EDUCAUSE Horizon Report: Teaching and Learning Edition. Boulder, CO: EDUCAUSE. <https://library.educause.edu/resources/2021/4/2021-educause-horizon-report-teaching-and-learning-edition>.
- Pistilli, M. D., & Arnold, K. E. (2010). Purdue signals: Mining real-time academic data to enhance student success. *About Campus*, 15(3), 22–24. <https://doi.org/10.1002/abc.20025>
- Pozzebón, J. A., Ashton, M. C., & Visser, B. A. (2014). Major changes: Personality, ability, and congruence in the prediction of academic outcomes. *Journal of Career Assessment*, 22(1), 75–88. <https://doi.org/10.1177/1069072713487858>
- Rawls, J. (1999). *A Theory of Justice* (Revised). Harvard University Press.
- Rosseel, Y. (2012). Lavaan: An r package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–26. <https://doi.org/10.18637/jss.v048.i02>
- Abu Saa, A., Al-Emran, M., & Shaalan, K. (2019). Factors affecting students' performance in higher education: A systematic review of predictive data mining techniques. *Technology, Knowledge and Learning*, 24(4), 567–598. <https://doi.org/10.1007/s10758-019-09408-7>
- Saha, D., Schumann, C., McElfresh, D. C., Dickerson, J. P., Mazurek, M. L., & Tschantz, M. C. (2020). Human comprehension of fairness in machine learning. In A. Markham (Ed.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society* (p. 152). <https://doi.org/10.1145/3375627.3375819>.
- Saxena, N. A., Huang, K., DeFilippis, E., Radanovic, G., Parkes, D. C., & Liu, Y. (2020). How do fairness definitions fare? Testing public attitudes towards three algorithmic definitions of fairness in loan allocations. *Artificial Intelligence*, 283, 1–15. <https://doi.org/10.1016/j.artint.2020.103238>
- Schlicker, N., Langer, M., Ötting, S., Baum, K., König, C. J., & Wallach, D. (2021). What to expect from opening up 'black boxes'? Comparing perceptions of justice between human and automated agents. *Computers in Human Behavior*, 122, 1–16. <https://doi.org/10.1016/j.chb.2021.106837>
- Sclater, N., & Mullan, J. (2017). Jisc briefing: Learning analytics and student success - assessing the evidence." Jisc. https://repository.jisc.ac.uk/6560/1/learning-analytics_and_student_success.pdf.
- Sen, A. K. (2009). *The Idea of Justice*. Harvard University Press. <https://doi.org/10.4159/9780674054578>
- Sherrick, M. F., Davenport, C. A., & Colina, T. L. (1971). Flexibility and satisfaction with college major. *Journal of Counseling Psychology*, 18(5), 487–489. <https://doi.org/10.1037/h0031509>
- Shin, D. (2019). Toward fair, accountable, and transparent algorithms: Case studies on algorithm initiatives in Korea and China. *Javnost - The Public*, 26(3), 274–290. <https://doi.org/10.1080/13183222.2019.1589249>
- Shin, D. (2020). User perceptions of algorithmic decisions in the personalized AI system: Perceptual evaluation of fairness, accountability, transparency, and explainability. *Journal of Broadcasting & Electronic Media*, 64(4), 541–565. <https://doi.org/10.1080/08838151.2020.1843357>

- Shin, D. (2021). The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. *International Journal of Human-Computer Studies*, 146, 1–10. <https://doi.org/10.1016/j.ijhcs.2020.102551>
- Shin, D., Zhong, B., & Biocca, F. A. (2020). Beyond user experience: What constitutes algorithmic experiences? *International Journal of Information Management*, 52(3), 1–11. <https://doi.org/10.1016/j.ijinfomgt.2019.102061>
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centered AI systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), 1–31. <https://doi.org/10.1145/3419764>
- Slade, S., & Prinsloo, P. (2013). Learning analytics: Ethical issues and dilemmas. *American Behavioral Scientist*, 57(10), 1510–1529. <https://doi.org/10.1177/0002764213479366>
- Smith, H. (2020). Algorithmic bias: Should students pay the price? *AI & Society*, 35(4), 1077–1078. <https://doi.org/10.1007/s00146-020-01054-3>
- Smith, B. O., Shrader, R., White, D. R., Wooten, J., Dogbey, J., Nath, Steve, O'Hara, Michael, Nan, X., & Rosenman, R. (2019). Improving student performance through loss aversion. *Scholarship of Teaching and Learning in Psychology*, 5(4), 278–288. <https://doi.org/10.1037/stl0000149>
- Srivastava, M., Heidari, H., & Kraus, A. (2019). Mathematical notions Vs. Human perception of fairness: A descriptive approach to fairness for machine learning. In A. Teredesai, V. Kumar, Y. Li, R. Rosales, E. Terzi, & G. Karypis (Eds.), *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)* (pp. 2459–2468). <https://doi.org/10.1145/3292500.3330664>
- Starke, C., & Lünich, M. (2020). Artificial intelligence for political decision-making in the European Union: Effects on Citizens' perceptions of input, throughput, and output legitimacy. *Data & Policy*, 2, 1–17. <https://doi.org/10.1017/dap.2020.19>
- Starke, C., Baleis, J., Keller, B., & Marcinkowski, F. (2022). Fairness perceptions of algorithmic decision-making: A systematic review of the empirical literature. *Big Data & Society*, 9(2), 1–16. <https://doi.org/10.1177/20539517221115189>
- Stegmann, H. (1980). Abiturient Und Studium: Bestimmungsfaktoren für Die Studienaufnahme Und Die Wahl Des Studiengangs. *Mitteilungen Aus Der Arbeitsmarkt- Und Berufsforschung*, 13, 531–542.
- Sun, L., & Tang, Y. (2021). Data-driven discrimination, perceived fairness, and consumer trust-the perspective of consumer attribution. *Frontiers in Psychology*, 12, 1–13. <https://doi.org/10.3389/fpsyg.2021.748765>
- Thurman, N., Moeller, J., Helberger, N., & Trilling, D. (2019). My Friends, Editors, Algorithms, and I: Examining audience attitudes to news selection. *Digital Journalism*, 7(4), 447–469. <https://doi.org/10.1080/21670811.2018.1493936>
- Tinto, V. (1975). Dropout from higher education: A theoretical synthesis of recent research. *Review of Educational Research*, 45(1), 89–125.
- Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, 106(4), 1039–1061. <https://doi.org/10.2307/2937956>
- Uhde, A., Schlicker, N., Wallach, D. P., & Hassenzahl, M. (2020). Fairness and decision-making in collaborative shift scheduling systems. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, et al. (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–13). <https://doi.org/10.1145/3313831.3376656>
- van Hootegeem, A., Abts, K., & Meuleman, B. (2020). Differentiated Distributive Justice Preferences? Configurations of preferences for equality, equity and need in three welfare domains. *Social Justice Research*, 33(3), 257–283. <https://doi.org/10.1007/s11211-020-00354-9>
- Veale, M., & Binns, R. (2017). Fairer machine learning in the real world: Mitigating discrimination without collecting sensitive data: Mitigating discrimination without collecting sensitive data. *Big Data & Society*, 4(2), 1–17. <https://doi.org/10.1177/2053951717743530>
- Vedel, A. (2016). Big five personality group differences across academic majors: A systematic review. *Personality and Individual Differences*, 92, 1–10. <https://doi.org/10.1016/j.paid.2015.12.011>
- Verbree, A.-R., Maas, L., Hornstra, L., & Wijngaards-de Meij, L. (2021). Personality predicts academic achievement in higher education: Differences by academic field of study? *Learning and Individual Differences*, 92, 1–11. <https://doi.org/10.1016/j.lindif.2021.102081>
- Verma, S., & Rubin, J. (2018). Fairness Definitions Explained. In Y. Brun, B. Johnson, & A. Meliou (Eds.), *Proceedings of the 40th International Workshop on Software Fairness (ICSE '18)* (pp. 1–7). <https://doi.org/10.1145/3194770.3194776>
- Vieira, D., Chinchilla, J. R., Liu, B. L., Yerovi, C., & Morales, D. (2022). Beyond the chatbot: How are universities using ai nowadays? In F. Almaraz-Menéndez, A. Maz-Machado, C. López-Esteban, & C.

- Almaraz-López (Eds.), *Strategy, policy, practice, and governance for AI in higher education institutions* (pp. 1–22). IGI Global. <https://doi.org/10.4018/978-1-7998-9247-2.ch001>
- Wang, R., Maxwell H. F., & Zhu, H. (2020). Factors influencing perceived fairness in algorithmic decision-making. In R. Bernhaupt, F. Mueller, D. Verweij, J. Andres, J. McGrenere, A. Cockburn, I. Avellino, et al. (Eds.), *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–14). <https://doi.org/10.1145/3313831.3376813>.
- Webb, H., Koene, A., Patel, M., & Vallejos, E. P. (2018). Multi-stakeholder dialogue for policy recommendations on algorithmic fairness. In *Proceedings of the 9th International Conference on Social Media and Society* (pp. 395–399). <https://doi.org/10.1145/3217804.3217952>.
- Wenzelburger, G., & Hartmann, K. (2021). Explaining the rise and the fall of a policy: Investigating the introduction and the dismantling of a French University admission policy through the lens of the multiple streams framework.
- Wenzelburger, G., & Hartmann, K. (2022). Policy formation, termination and the multiple streams framework: the case of introducing and abolishing automated university admission in France. *Policy Studies*, 43(5), 1075–1095. <https://doi.org/10.1080/01442872.2021.1922661>
- Wirtz, BW., & Müller, WM. (2019). An integrated artificial intelligence framework for public management. *Public Management Review*, 21(7), 1076–1100. <https://doi.org/10.1080/14719037.2018.1549268>
- Wirtz, B. W., Weyerer, J. C., & Geyer, C. (2019). Artificial intelligence and the public sector-applications and challenges. *International Journal of Public Administration*, 42(7), 596–615. <https://doi.org/10.1080/01900692.2018.1498103>
- Wong, P.-H. (2020). Democratizing algorithmic fairness. *Philosophy & Technology*, 33(2), 225–244. <https://doi.org/10.1007/s13347-019-00355-w>
- Yilmaz, K., Gizem, F., & Yilmaz, R. (2020). Student opinions about personalized recommendation and feedback based on learning analytics. *Technology, Knowledge and Learning*, 25(4), 753–768. <https://doi.org/10.1007/s10758-020-09460-8>
- Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education - Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), 1–27. <https://doi.org/10.1186/s41239-019-0171-0>
- Zhou, J., Verma, S., Mittal, M., & Chen, F. (2021). Understanding relations between perception of fairness and trust in algorithmic decision making. In *Proceedings of the 8th International Conference on Behavioral and Social Computing (BESC 2021)* (pp. 1–5). <https://doi.org/10.1109/BESC53957.2021.9635182>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Marco Lünich¹  · Birte Keller¹  · Frank Marcinkowski¹

✉ Marco Lünich
marco.luenich@hhu.de

Birte Keller
birte.keller@hhu.de

Frank Marcinkowski
frank.marcinkowski@hhu.de

¹ Department of Social Sciences, Heinrich Heine University, Düsseldorf, Germany