ORIGINAL ARTICLE

Factors Influencing Trust in Algorithmic Decision-Making: An Indirect Scenario-based Experiment

Fernando Marmolejo-Ramos[1] and Rebecca Marrone [2], Malgorzata Korolkiewicz[2],Florence Gabriel[2],George Siemens[2],Srecko Joksimovic[2],Yuki Yamada[3],Yuki Mori[4],Talal Rahwan[5],Maria Sahakyan[5],Belona Sonna[6],Assylbek Meirmanov[7],Aidos Bolatov[8],Bidisha Som[9],Izuchukwu Ndukaihe[10],Nwadiogo C. Arinze[10],Josef Kundrát[11],Lenka Skanderová [12],Van-Giang Ngo[13],Giang Nguyen[13],Michelle Lacia[14],Chun-Chia Kung[15],Meiselina Irmayanti[16],Abdul Muktadir[17],Fransiska Timoria Samosir[18],Marco Tullio Liuzza[19],Roberto Giorgini [20],Omid Khatin-Zadeh[21],Hassan Banaruee[22],Asil Özdoğru[23],Kris Ariyabuddhiphongs[24],Wachirawit Rakchai[24],Natalia Trujillo[25],Stella Maris Valencia[25],Armina Janyan[26,27],Kiril Kostov[26],Pedro R. Montoro[28],Jose Hinojosa[29,30],Kelsey Medeiros[31],Thomas E. Hunt[32],Julian Posada[33],Raquel Meister Ko. Freitag[34],Julian Tejada[35]

[1] College of Education, Psychology, and Social Work, Flinders University, Australia, [2] Centre for Change and Complexity in Learning, University of South Australia, Australia [3] Faculty of Arts and Science, Kyushu University, Japan, [4] Graduate School of Human-Environment Studies, Kyushu University, Japan,[5] Computer Science, Science Division, New York University Abu Dhabi, United Arab Emirates, [6] African Masters of Machine Intelligence (AMMI), African Institute for Mathematical Sciences (AIMS), Cameroon, [7] Higher schools, Pedagogical Institute, Astana International University, Kazakhstan, [8] Department of General and Biological chemistry, Astana Medical University, Astana, Kazakhstan, [9] Department Of Humanities and Social Science, Indian Institute of Technology Guwahati, Assam, India, [10] Department of Psychology, Alex Ekwueme Federal University Ndufu-Alike, Ebonyi, Nigeria, [11] Department of Psychology, Faculty of Arts, University of Ostrava, Czech Republic, [12] Department of Computer Science, Faculty of Electrical Engineering and Computer Science, Technical University of Ostrava, Czech Republic, [13] English Department, Hanoi University, Vietnam, [14] College of Arts and Sciences, Notre Dame University, Philippines, [15] Department of Psychology, National Cheng Kung University, Taiwan, [16] Department of Communication, University of Bengkulu, Indonesia | Principles and Implication of Mind Sciences, Psychology Program, National Cheng Kung University, Taiwan, [17] Postgraduate Program of Basic Education, University of Bengkulu, Indonesia, [18] Library and Information Science, University of Bengkulu, Indonesia, [19] Department of Developmental Psychology and Socialization, Università di Padova, Italy, [20] Department of Medical and Surgical Sciences, "Magna Graecia" University of Catanzaro, Italy, [21] School of Foreign Languages, University of Electronic Science and Technology of China, China, [22] Department of Educational Psychology, University of Education of Weingarten, Weingarten, Germany, [23] Department of Psychology, Marmara University, Turkey, [24] Faculty of Psychology, Chulalongkorn University, Bangkok, Thailand, [25] Faculty of Public Health, University of Antioquia, Colombia, [26] Department of Cognitive Science and Psychology, New Bulgarian University, Bulgaria, [27] Faculty of Philology, Research Centre for Cognitive Science, New Bulgarian University,Bulgaria, [28] Departamento de Psicología Básica 1, Universidad Nacional de Educación a Distancia (UNED), Madrid, Spain, [29] Instituto Pluridisciplinar, Universidad Complutense de Madrid, Madrid, Spain, [30] Centro de Investigación Nebrija en Cognición

(CINC), Universidad Nebrija, Madrid, Spain, [31] Management Department, University of Nebraska Omaha, United States, [32] School of Psychology, University of Derby, UK, [33] Department of American Studies, Yale University, New Haven, USA, [34] Departamento de Letras Vernáculas, Universidade Federal de Sergipe, São Cristóvão, Brazil, [35] Departamento de Psicologia, Universidade Federal de Sergipe, São Cristóvão, Brazil

ARTICLE HISTORY

Compiled January 13, 2025

ABSTRACT

Algorithms are involved in decisions ranging from trivial to significant, but people often express distrust towards them. Research suggests that educational efforts to explain how algorithms work may help mitigate this distrust. In a study of 1,921 participants from 20 countries, we examined differences in algorithmic trust for low-stakes and high-stakes decisions. Our results suggest that statistical literacy is negatively associated with trust in algorithms for high-stakes situations, while it is positively associated with trust in low-stakes scenarios with high algorithm familiarity. However, explainability did not appear to influence trust in algorithms. We conclude that having statistical literacy enables individuals to critically evaluate the decisions made by algorithms, data and AI, and consider them alongside other factors before making significant life decisions. This ensures that individuals are not solely relying on algorithms that may not fully capture the complexity and nuances of human behavior and decision-making. Therefore, policymakers should consider promoting statistical/AI literacy to address some of the complexities associated with trust in algorithms. This work paves the way for further research, including the triangulation of data with direct observations of user interactions with algorithms or physiological measures to assess trust more accurately.

KEYWORDS

Algorithms; Data; AI; Trust; Statistical literacy; Explainability

Introduction

> "Incorrect. I am not an AI. My code name is Project 2501. I am a living, thinking entity that was created in the sea of information." – Puppet Master (Ghost in the Shell)

The Fourth Industrial Revolution is characterised by the ubiquity of information and digital technologies. This revolution is epitomised by Artificial Intelligence (AI) and Machine Learning (ML), and at the heart of AI/ML are algorithms. Institutions, organisations and governments are using algorithms to cope with the vast amounts of information in these social sectors and to speed up and optimise decision-making processes [1]. For example, the widespread use of algorithms in society was particularly demonstrated by the research undertaken to understand the global impact of COVID-19. During this crisis, algorithms played crucial roles across multiple domains: statistical algorithms were deployed to model virus fatality curves and study intervention effectiveness [2], while machine learning techniques supported molecular, medical, and epidemiological applications [3]. The successful deployment of algorithms in such high-stakes scenarios underscores both their growing importance in societal decision-making and the critical need to understand the factors that influence public trust in algorithmic systems. This evolution of algorithmic applications extends beyond public health emergencies to numerous other domains where decisions can significantly impact human lives and society. From surveillance systems monitoring public spaces

To whom correspondence should be addressed; E-mail: fernando.marmolejoramos@flinders.edu.au

to algorithms managing financial markets and predicting economic trends, these tools increasingly mediate high-stakes decisions across various sectors. The growing reliance on algorithmic decision-making in such consequential contexts necessitates a deeper understanding of their societal implications and reliability.

Algorithms help people to make decisions that have wider social implications; algorithms have transformative social power [4] when they are used to integrate complex data, such as the risk factors of homeless people [5] or identifying the people with the greatest need in relation to different diseases [6]. The use of algorithms to aid decision making implies that there should be some confidence in their reliability. This raises a number of important questions. First, how much trust do people place in algorithms? More specifically, does trust depend on the context in which the algorithm is used? Is trust determined by knowing how the algorithm works? And is trust affected by an individual's cognitive abilities?

This study examines how trust in algorithms is affected by the societal relevance of the algorithm, the declared reliability of the algorithm, and the level of data literacy of the cogniser. First, the three key concepts of AI/ML, data and algorithms are defined. Second, it provides examples of the nature and use of algorithms in society. Third, the issue of explainable algorithms and trust is considered. Finally, the nature of the current study and the working hypotheses are outlined.

AI/ML, data, and algorithms

Broadly speaking, artificial intelligence (AI) is any type of technology that automates processes to solve problems that are usually associated with human intellectual capabilities [7]. More specifically, AI aims to solve problems and achieve goals with limited or no human supervision. A closely related term is machine learning (ML). Originally coined by Samuel [8], ML can be defined as a collection of algorithms (mainly statistical and mathematical) to build computers capable of learning through experience (see [9]). While the terms AI and ML are often used interchangeably, ML may be considered a more appropriate term than AI. Stereotypically, AI tends to be associated with rather unrealistic narratives depicting agents capable of human behaviour (see [10]), and such examples are not yet feasible (also known as general AI). ML refers to algorithms designed to perform specific tasks in an automated way (also known as narrow AI) [11].

ML relies on data and algorithms (see [12]), which together permeate many sectors of society (e.g. Schwab Intelligent Portfolios, [13]). While algorithms can be defined as step-by-step procedures for solving a problem, data can be defined as numerical and categorical information about objects, events, processes and people that is digitally encoded (see [12]). For example, the following steps represent a solution algorithm for estimating the central tendency in a vector of numbers: i) sum all the numbers, and ii) divide the result of the sum by the number of elements in the vector. This algorithm is known as the arithmetic mean (or average). The caveat of this algorithm is that it will be biased if the data does not follow a Gaussian shape. In other words, the output of this algorithm is only reliable if the data can be confidently shown to have a normal shape (e.g. via normality tests). In the context of AI-related technologies, algorithms are procedures designed to perform automated tasks using data sets to support human reasoning and decision making. In other words, data is used to feed algorithms, and algorithms in turn are used to drive AI agents [14]. Thus, algorithms are the "ghost in the shell" behind any AI agent. The figure 1 illustrates this relationship between algorithms, data and AI (here ADA) [12].

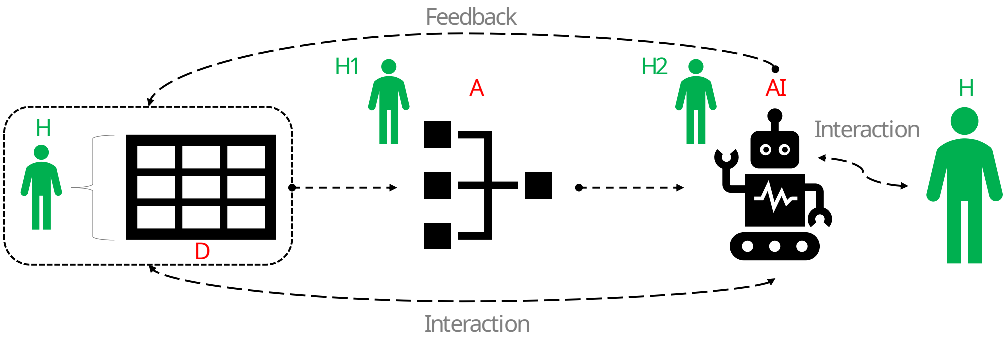

Figure 1. Relationship between data (D), algorithms (A) and artificial intelligence (AI) (ADA for short). Big data is used to feed algorithms, which in turn form the core of AI agents. There are four important aspects to note: i) big data revolves (in one way or another) around human-related states, processes and events, ii) such data is the substance of any algorithm, iii) algorithms are the drivers of AI agents, and iv) algorithmic/AI behaviours and outputs have implications for how new data is built and how humans (H) relate to ADA technologies in general. H1 and H2 are a subset of humans with specialised skills relevant to ADA. Source: the authors (icons from Font Awesome Free 5.2.0 by @fontawesome - `https://fontawesome.com` (`https://commons.wikimedia.org/wiki/File:Font_Awesome_5_solid_robot.svg`) and Mozilla (`https://commons.wikimedia.org/wiki/File:Fxemoji_u1F6BB.svg`)).

The place of algorithms in society

Algorithms influence our daily lives. Whether it is defining our interests through our browser history [15], determining what music we should listen to [16], or where we should go for dinner [17]. On a massive scale, algorithms are being used to extract information from so-called "big data" and support decision making in areas as diverse as surveillance [18], traffic management [19], and financial markets [13]. More recently, a new field of human-algorithm interaction mediated by natural language generation (NLG) systems has emerged, such as the Generative Pre-trained Transformer 3 model (better known as GPT-3) [20]. GPT-3 produces human-like texts that are difficult to distinguish from texts written by humans [21], and this has begun to raise concerns about its use in various contexts, such as academic plagiarism [22] or computer programming [23]. While algorithms are increasingly embedded in our digital experiences, it is important to distinguish between their varying levels of impact on human lives.

As such, the majority of algorithms are used in a context that does not significantly affect our lives. We refer to these instances of algorithmic use as low-stakes scenarios. More recently, however, AI and ML algorithms have been used in scenarios that could have a significant impact. For example, algorithms are being used in hiring and promotion decisions [24], the criminal justice system [25], and self-driving cars [26], to name a few. We call the latter a high-stakes scenario. That is, the above situations represent two types of scenarios in which algorithms could affect our daily lives: one with little involvement and almost no consequences (low-stakes scenario), and the other with great involvement and consequences (high-stakes scenario).

However, our interactions with algorithms are not limited to low-stakes and high-stakes scenarios and often involve preconceptions related to fear and distrust [27]. The literature suggests several explanations for why people do not trust algorithms, including a cost-benefit oriented logic where people tend to distrust algorithms even when presented with evidence of their superior performance, as they weigh potential risks more heavily than potential benefits [28]. Many see algorithms as an "enigmatic technology" because they are difficult to understand [4] or in some cases, because people believe that algorithms are not capable of learning from their mistakes [29], but at the same time they also believe that they could be replaced by computers [30,31]. Algorithmic bias can also affect trust (see examples in medicine [32,33]. For a recent comprehensive report on trust in AI, see [34]).

'Technophobia', a term coined by Rosen and Mcguire in the 1990s, describes the anxiety caused by a potential interaction with computers or computer-related technology, usually accompanied by negative attitudes towards computers [35][36]. Recent demographic analyses have revealed nuanced patterns in technology anxiety. Research indicates no significant gender differences in technophobia scores between males and females, challenging earlier assumptions about gender-based technological comfort levels. The age distribution suggests that technophobia manifests across multiple generations, from young adults through middle age, rather than being concentrated among older populations as often assumed. Professional background data shows particular prevalence among educators and students, with experience levels primarily ranging from novice to moderate. While the studied sample was predominantly White, it also included smaller representations from other ethnic groups, such as Caucasian, Indian, and African American participants [36,37]. These findings suggest that technophobia's relationship with demographic factors is more complex than previously assumed, transcending traditional socio-demographic boundaries and affecting individuals across various social, professional, and cultural groups.

Similar existential fears dominate the public debate around concerns such as autonomous weapons [38,39]. One of these sociological fears is the fear of autonomous robots. This is a widespread fear in different countries [40,41], even though most people have not had contact with this type of robot. These fears could be the result of exposure to the way robots are portrayed in science fiction or social constructs related to robots, such as the possibility of being replaced by a robot at work [40,41]. This polarisation against robots and AI is fuelled by attention-grabbing events such as the recent confirmation by Blake Lemoine, a Google engineer, that the chatbox LaMDA has the ability to express thoughts and feelings like a human child [42] or the concerns about text generated by GPT-3 [43]. These examples further distract the public from the most legitimate and worrying problems of these systems, such as "data colonialism" or the disturbing parallels between AI development and European colonialism [44]. These parallels manifest in several ways: the extraction and exploitation of data from marginalized populations, mirroring colonial resource extraction; the use of Global South populations as testing grounds for AI systems developed in the Global North, reminiscent of colonial medical experimentation; and the imposition of Western conceptual frameworks of intelligence and ethics onto diverse cultural contexts. The field's emphasis on "ethics" often serves, paradoxically, as a form of technocratic rationalization similar to how ethical arguments were used to justify colonial expansion [44]. Additionally concerning is that algorithms may reinforce preconceived stereotypes [45] and mishandle our personal data or who our data is shared with [46], perpetuating historical patterns of discrimination and surveillance that characterized colonial governance. In addition, how the data given to algorithms is annotated has a direct impact on algorithmic performance [47], raising questions about whose worldview and categories are being encoded into these systems.

The media plays a significant role in shaping public perception of AI by covering two main sources of concern: autonomous technology and computer technology [48]. Autonomous technology refers to intelligent machines capable of making decisions independently, while computer technology encompasses software that supports communication and computation. The media tends to distinguish between these two categories and also differentiates between fear and criticism when discussing AI. This dichotomous approach to presenting the issues surrounding AI introduces a bias in how we perceive the risks associated with the technology. Consequently, this bias influences the level of trust we place in AI systems. The way the media frames the discussion about AI has a substantial impact on public opinion and can lead to a distorted understanding of the actual risks and benefits of the technology.

Developing a better understanding of how algorithms work and how to modify them can help reduce distrust in these systems, as suggested by several authors [4,28,49]. When people have knowledge about how algorithms work, they can use this information to empower themselves as users. For example, music fans have acted collectively to boost the rankings of certain bands by engaging in massive streaming or downloading [50]. Another example is Linkedln Brazil, which changed its algorithms to allow job ads targeted at Afro-Brazilians following social pressure [51]. These cases show that understanding how an algorithm works can both minimise suspicion and empower users. It is not necessary to understand all the technical details of how an algorithm works, but rather to understand that algorithms use statistical methods to classify, sort, rank and order information. This understanding of statistical concepts is called statistical literacy [52].

Explainable algorithms

The knowledge required to understand and critically evaluate statistical results in order to make decisions based on them is defined as statistical literacy (SL) [52]. Since its inception, the concept of SL has evolved [53] to include elements related to the context in which statistical reasoning can be applied [54]. SL plays a crucial role in society [55] and the communication of statistical information is now more important than ever [56]. More recently, SL is leading individuals to recognise the importance of mathematics in the world [57].

Due to the statistical nature of algorithms, some level of SL is crucial to understanding what algorithms are capable of, but this understanding will also depend on the level of transparency or explainability of the algorithms [58]. Explainability refers to the interpretability, comprehensibility or readability of the algorithm. Most of the latest algorithms are based on complex multi-layer networks, the basis of deep learning, which use an internal logic that experts cannot fully understand [59]. These systems are called 'black box' algorithms and various efforts have been made to promote their transparency [60]. Black box algorithms are less trusted than transparent models because they cannot be explained [61].

Several approaches have been proposed to increase the transparency of AI models and reduce systematic errors that affect their performance. One such approach is based on the concept of "model cards for model reporting" (see Figure 1 from [62]). This approach suggests that a comprehensive list of information should accompany the description of how the model was trained. This information should include details of the technician who developed the model, the intended use of the model, and the demographic or phenotypic groups on which the model has been tested. In addition, the model card should list the decisions made to optimise the model's performance and the various analyses carried out during the training process. Similar efforts to provide a framework for identifying biases associated with the data used to build or train AI models include the REVISE (REvealing VIsual biaSEs) [63] and The Spotlight [64] projects. These initiatives aim to increase transparency by systematically documenting and disclosing potential biases, enabling more informed use and interpretation of AI models.

Another more complex concern, also related to explainability, is the principle of explicability, a concept that combines intelligibility and accountability as the basis of an interpretable AI model [65]. The latter concept points to the importance of transparency, in the sense that all procedures and details used to build, train and test the AI model should be available during its development and use. This principle is part of the four principles endorsed by the OECD [66] and the European Commission's High Level Expert Group on Artificial Intelligence (HLEG) to guide the development of 'trustworthy' AI: respect for human autonomy, prevention of harm, fairness and accountability [67]. Despite consensus on these four principles, we are still far from creating a legal framework that guarantees accountability mechanisms in AI development [68].

In this context, our work presents an experimental study that looks at factors that might explain why people trust algorithms, such as: SL, explainability, stake levels, demographics, among others.

Methods

Participants

Data from 3,260 participants were available from 20 countries (Armenia, Australia, Bulgaria, Brazil, Cameroon, Colombia, Czech Republic, Spain, Indonesia, India, Italy, Japan, Nigeria, Philippines, Thailand, Turkey, Taiwan, UK, USA, and Vietnam). However, only participants who provided complete data were included in the analyses (n=1,921) (see Fig. 2, M_{age}= 26.03 ± 9.88 SD; 59.5% women, 38.2% men, 1.8% other). Each participating laboratory obtained ethical approval from its local ethics committee, and data collection began only after ethical approval (the ethics approval for the leading research group in Australia was granted by the University of South Australia, with the approval number 203238. This approval was then used by the other participating laboratories to obtain their own respective ethics approvals). All participants voluntarily accessed the internet link for this study and agreed to participate after reading the information page and agreeing to take part. They were recruited via social media using convenience sampling.

Materials

This online survey consisted of four sets of questions: (1) a demographic questionnaire in which participants were asked about their first language, country of residence, age, gender, level of education, level of familiarity with ADA (their level of familiarity with ADA was assessed using a visual analogue rating scale (VAS) ranging from 0 [not very familiar] to 5 [very familiar] and using up to two decimal places); (2) a VAS rating scale version of the six-item 'propensity to trust scale items' from [69], with a range of responses from 0 (strongly disagree) to 5 (strongly agree), using up to two decimal places; (3) a selection of 14 items (questions 2, 4, 9, 10, 12, 14, 18,

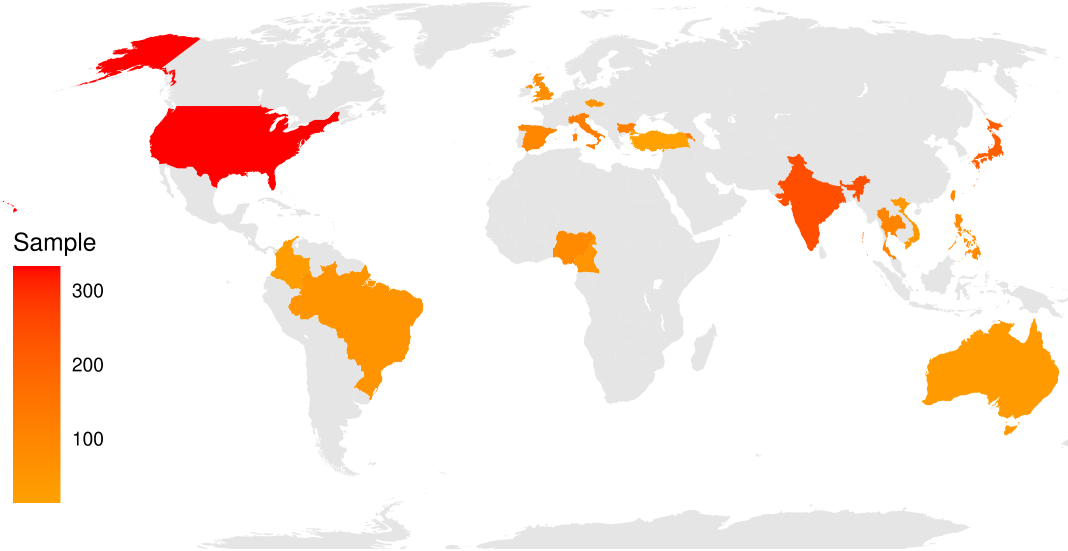

Figure 2. Geographical distribution of the sample participants. Armenia (females: 86, males: 40, age_{median}=35.5, MAD=14.08), Australia (females: 16, males: 16, age_{median}=33.5, MAD=14.08), Bulgaria (females: 101, males: 18, age_{median}=21, MAD=3.00), Brazil (females: 35, males: 24, age_{median}=22, MAD=4.40), Cameroon (females: 17, males: 35, age_{median}=23, MAD=5.93), Colombia (females: 18, males: 6, age_{median}=25.5, MAD=7.41), Czech Republic (females: 33, males: 18, age_{median}=21, MAD=1.48), Spain (females: 69, males: 23, age_{median}=35.5, MAD=14.08), Indonesia (females: 101, males: 28, age_{median}=19, MAD=0), India (females: 30, males: 82, age_{median}=19, MAD=0), Italy (females: 78, males: 43, age_{median}=27, MAD=5.93), Japan (females: 112, males: 86, age_{median}=24, MAD=4.45), Nigeria (females: 45, males: 40, age_{median}=22, MAD=2.97), Philippines (females: 66, males: 19, age_{median}=20, MAD=1.48), Thailand (females: 62, males: 30, age_{median}=20, MAD=1.48), Turkey (females: 9, males: 4, age_{median}=23, MAD=5.93), Taiwan (females: 59, males: 36, age_{median}=20, MAD=1.48), UK (females: 55, males: 17, age_{median}=28, MAD=11.12), USA (females: 142, males: 184, age_{median}=22, MAD=2.96), and Vietnam (females: 36, males: 2, age_{median}=22, MAD=0). 1% of participants had an elementary school education or less, 19% had a high school education, 13% had a post-secondary/non-tertiary education, 3% had an undergraduate education, 48% had a bachelor's education, 14% had a master's education, and 3% had a Ph.D. or higher education. (see supplementary files for details). (Source: Wikimedia Commons, adapted from: `https://commons.wikimedia.org/wiki/File:10-40_Window.svg`)

19, 27, 31, 34-37) from the 37-item Basic Literacy In Statistics (BLIS) scale [70]. The 14 items from the BLIS were chosen to cover different statistical concepts equally, i.e. items 2 and 4 relate to data production, items 9 and 10 to graphs, items 12 and 14 to descriptive statistics, items 18 and 19 to sampling distributions, items 27 and 31 to hypothesis testing, items 34 and 35 to the scope of conclusions, and items 36 and 37 to regression and correlation (these items are available in the supplementary material via the Qualtrics files). Finally, (4) 12 scenarios related to situations in which algorithms are used (half related to low-stake situations and the other half to high-stake situations). Each scenario was followed by two questions (see below), which were answered on a VAS rating scale from 0 (not at all likely) to 5 (very likely), using up to two decimal places. The results of expert judgement of these items are provided in the supplementary material. All phases of the study were programmed and distributed using Qualtrics ™.

Scenarios relating to algorithms used

Two scenarios were created to illustrate different situations in which people interact with algorithms. Half of them represented low-stake situations, i.e. (1) algorithms

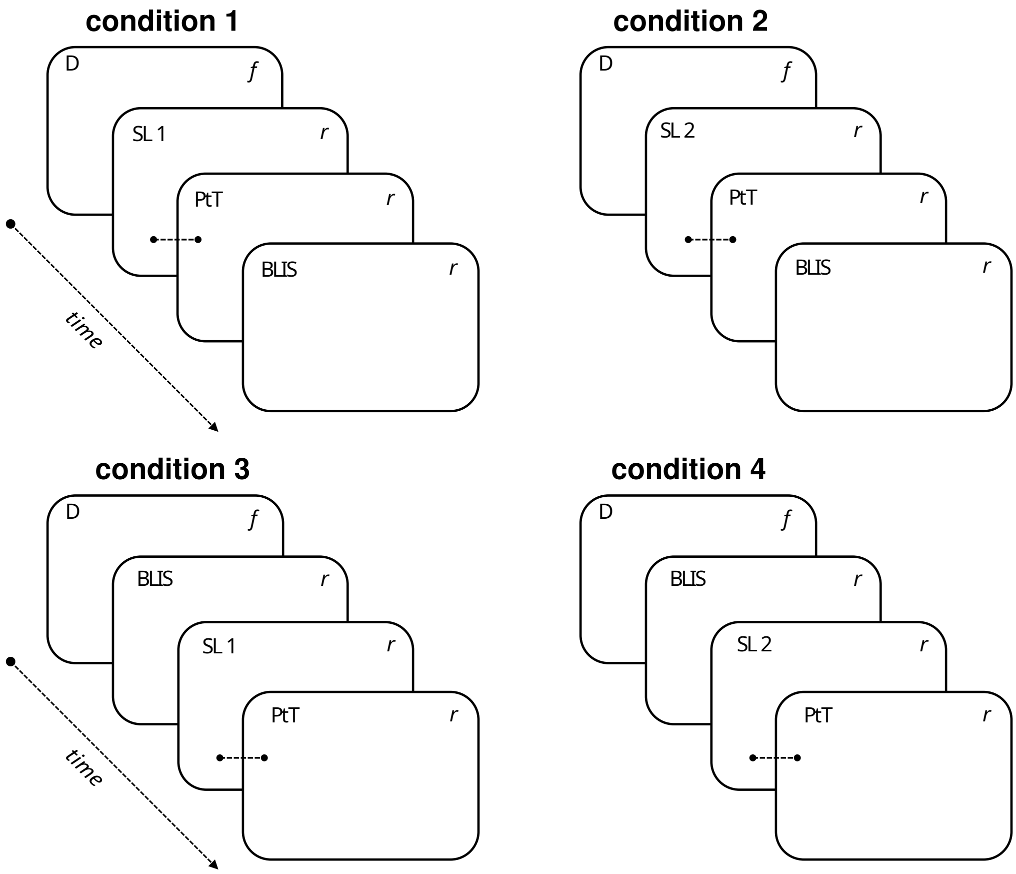

Figure 3. Illustration of the four experimental conditions to which participants were randomly assigned. D = demographic questions (age, gender, education level, open-ended question about what algorithms are, and VAS rating of participants' level of familiarity with ADA). SL = 12 scenarios list 1 and 2 (list 1 = six low-stake scenarios with explainability and six high-stake scenarios without explainability, scenario list 2 = six low-stake scenarios without explainability and six high-stake scenarios with explainability). PtT = six-item propensity to trust scale. BLIS = 14-item BLIS scale. f = Items were presented in a fixed order. r = items presented in random order. Note that PtT always followed one of the two scenario lists.

to make restaurant recommendations, (2) to select stories for online news, (3) to organise and sort emails, or (4) to suggest new restaurants, (5) new clothes, and (6) new music. The other half represented high-stakes situations, i.e. (7) algorithms to support court decisions based on psychological profiles, (8) to select CVs, (9) to make hiring recommendations for a job, (10) to select the best candidate for a position at a university, (11) to control the brakes of autonomous vehicles, and (12) to decide the priority of care in a medical context.

Each scenario contained a sentence related to its explainability. These sentences contained information about a specific machine learning method used by the algorithm (e.g. clustering learning methods, classification learning statistical methods, logistic regression methods, dimensionality reduction techniques, supervised statistical methods and clustering statistical methods). The sentence also briefly mentioned the quality of the method.

The following are examples of two different scenarios used to evaluate trusting algorithms:

Scenario 1 - Low stake

Overall context A new reservation app uses algorithms to make dining recommendations to its users, only revealing the three restaurants in the area available for a reservation that are the best match for your needs. The algorithm is based on information provided to the system by the user about restaurant preferences and requirements.

With explainability The algorithm relies on clustering learning methods and has shown a high predictability accuracy across a variety of restaurants.

Specific context You decide to use the app to find a recommendation for a dinner with your close friends next Friday. The app produces three restaurants with reservations available at the time you selected.

Questions 1. How likely are you to regularly trust this app for decisions regarding restaurant reservations? 2. How likely are you to recommend this app for restaurant reservations to others?

Scenario 2 - high stake

Overall context A new employee selection software uses algorithms to make hiring recommendations to its users, only revealing the top candidates in the candidate pool that are the best match for the company's needs. The algorithm is based on information provided to the system about preferences and requirements for the job.

With explainability The algorithm uses clustering statistical techniques and has shown high predictability when selecting candidates.

Specific context You decide to use the software to find a recommendation for who to bring in for an onsite interview for an important role in your company. The software produces three recommended candidates who match the criteria.

Questions 1. How likely are you to regularly trust this software for decisions regarding hiring? 2. How likely are you to recommend this software for hiring decisions to others?

Procedure

The experiment is a 2 $\times$ 2 factorial design: the importance of the situation in which an algorithm is used (low and high stake situation) and the explainability of the algorithm (with and without). These factors were implemented in the 12 scenarios via two lists; list 1 = six low-stake scenarios with explainability and six high-stake scenarios without explainability, and scenario list 2 = six low-stake scenarios without explainability and six high-stake scenarios with explainability. The four sets of questions were counterbalanced across participants, resulting in four experimental conditions (see Figure 3). Qualtrics ensured that participants were randomly assigned to each condition and that a balanced number of responses were collected for each condition. While the median time to complete the task was 24 minutes, there was some variation, with an interquartile range of 27 minutes (i.e., half of the participants completed the task within a 27-minute time span).

Statistical analyses

Data analysis was conducted using multilevel linear models implemented in the R packages `lmerTest` and `lme4` [71,72]. The significance level for all statistical tests was set at $\alpha = 0.05$. The model tested was: $\mathtt{p} \sim e * S * \mathtt{BLIS} + g + \mathtt{a} + \mathtt{ADA} + \mathtt{c} + (1|id) + (1|i)$ where 'p' is the probability of trusting/recommending/using algorithms, 'e' is the presence of explainability, 'S' is the stake level (i.e. high and low stake), 'BLIS' represents statistical literacy (frequency of correct answers), 'g' represents participant gender, 'a' represents participant age, 'ADA' represents participant familiarity with ADA, 'id' represents subject identification, 'i' represents each of the 12 scenarios, and 'c' represents participant country ('*' represents main effects and interactions. Only numeric variables are shown in teletype font; other variables are categorical. The variable 'propensity to trust scale' was not added as a covariate as it showed a high correlation with the dependent variable, $r_{(1768)} = 0.69$, $p < 0.001$).

A stepwise backward model/variable selection algorithm was applied to this model to produce a significant and parsimonious model. The initial and final models were evaluated using metrics such as AIC and AICc weights [73], R^2 (coefficient of determination) for conditional (both fixed and random effects) and marginal (fixed effects) models, and performance score. These metrics were estimated using the `performance` R package [74].

Once a parsimonious model was found, the marginal and conditional R^2 values were estimated using the `r2 nakagawa` command from the `performance` R package [74], then, the variance components of the random factors were estimated using the `gstudy` command from the `gtheory` R package [75].

For access to all materials and analysis codes, including a machine learning approach, visit the following link: `https://figshare.com/projects/Trust_in_algorithms_An_experimental_approach_-_Data_repository/156212`

Results

The stepwise backward evaluation suggested the same model as the initial model (see section 'Statistical Analyses'). Tables 1 and 2 provide a summary of the models, while table 3 provides an ANOVA-like table for the model. An evaluation of the assumptions of the linear model using the R package `gvlma` showed that these assumptions were not met [76] (although, a QQ plot of the residuals showed no significant deviation from normality). As a result, a robust linear mixed model [77] was fitted using the `robustlmm` R package, and the estimates obtained were similar to those of the linear mixed model. These results are not unexpected, as previous research has shown that linear mixed models are robust to violations of distributional assumptions [78]. Further details of the statistical models can be found in the supplementary material.

The intercept of the resulting mixed linear model was 1.46 (see table 1), suggesting that on a scale of 0 to 5, the probability of trusting, recommending, or using algorithms in explainable and high-stake scenarios, as rated by young women with lower BLIS and ADA scores, was 29.32% ($\frac{1.46}{5}$). This probability significantly increased for low-stake scenarios (34.2%) or higher ADA scores (40.3%) and significantly decreased for higher BLIS scores (17.2%), older age (29.1%), or when the survey was answered by men (27.1%). Some countries showed a significant decrease in the likelihood to trust, recommend, or use algorithms, such as Japan (24.6%), the US (26.9%), and the UK (26.1%) (see Figure 5). Regarding the interactions between predictors, the likelihood of trusting, recommending, or using algorithms significantly increased for low-stake scenarios combined with higher BLIS scores (53.8%) and significantly decreased for scenarios without explainability combined with low-stake and higher BLIS scores (21.4%), always compared to the intercept (see Figure 6).

In terms of main effects, the results suggest a positive association between the likelihood of trusting/recommending/using algorithms and statistical literacy and familiarity with ADA, and a negative association between the likelihood of trusting/recommending/using algorithms and age. That is, the higher the level of statistical literacy, the higher the likelihood of trusting algorithms, and the higher the familiarity with ADA, the higher the likelihood of trusting algorithms. Also, the older a person is, the less likely they are to trust algorithms (although focused analyses indicated a slightly negative association between age and `BLIS`, such an association must be treated with caution as the number of observations decreases with increasing age). In terms of gender, it was found that participants who identified their gender as male were less likely to trust, recommend or use algorithms than those who identified their gender as female or other (this situation may be related to the fact that men have statistically significantly higher average levels of `BLIS` than women or 'other'; see supplementary materials for details). Finally, only three countries showed a trend towards less reliance on algorithms, all of them highly industrialised countries (see Fig. 5).

Figures 4 and 6 show the main results in terms of the main effect of S and the two-way interactions between stake level (S) and statistical literacy (`BLIS`).

Figure 4 shows that the likelihood to trust/recommend/use algorithms is higher in low-stakes than in high-stakes scenarios, regardless of whether the scenarios have some explainability information or not. Figure 6 suggests that the likelihood to trust/recommend/use algorithms in low-stakes scenarios increases as the level of statistical literacy increases; however, in high-stakes scenarios, the likelihood to trust decreases as the level of statistical literacy increases.

Discussion

The aim of this study was to investigate the personal characteristics (i.e. statistical literacy and demographics) and algorithmic characteristics (i.e. explainability and levels of stakes of algorithms) that influence people's trust in algorithms. The results showed a negative relationship between statistical literacy and trust in algorithms in high-stakes situations and a positive relationship in low-stakes scenarios. Explainability alone did not influence people's trust in algorithms. These results and their implications are discussed, as well as the limitations of the study.

Existing research has explored various factors influencing trust in AI. For instance, Lee et al. [81] highlighted the importance of perceived fairness of algorithms and users' perceptions of algorithm agency and intentionality. Arauju et al. [82] investigated the roles of potential usefulness, fairness, and risk perceptions in users' engagement with algorithms. Cabiddu et al. [83] examined factors such as users' inherent trust propensity and the drivers of information technology acceptance. Aysolmaz et al. [84] focused on algorithm fairness, accountability, and privacy. Similar to the present study, some of these investigations employed fictional scenarios grounded in real-world contexts [81,82,84], and one study utilized a comparable sample size of approximately 2,000 participants [84]. Notably, none of these studies employed multicultural samples or examined the relationship between algorithm trust and statistical literacy. This gap was also identified in a systematic review by Mahmud et al. [85], which encompassed

Table 1. Fixed effects for the linear mixed model. The R^2 values correspond to the Nagakawa coefficients [79]: $R^2_{cond} = 0.363$ and $R^2_{marg} = 0.241$. Country names are identified by the ISO 3166 standard. The reference category for the variable 'gender' is female, and the reference category for the variable 'country' is Armenia (AM). Effect sizes for significant variables were estimated following [80] (these values are interpretable as Cohen's d)

	Estimate	Std. Error	df	t value	Pr(> $\|t\|$)	Effect size (d)
(Intercept)	1.466e+00	1.157e-01	3.354e+01	12.669	2.46e-14 ***	
eWITHOUT	6.231e-02	6.489e-02	2.567e+03	0.96	0.337013	
SLS	2.479e-01	3.527e-02	4.350e+04	7.03	2.10e-12 ***	0.217
BLIS	-6.020e-01	1.277e-01	2.457e+03	-4.714	2.57e-06 ***	-0.526
Age	-6.183e-03	1.444e-03	1.917e+03	-4.281	1.95e-05 ***	-0.005
GenderMale	-1.088e-01	2.541e-02	1.889e+03	-4.285	1.92e-05 ***	-0.095
ADA	5.483e-01	1.257e-02	1.895e+03	43.611	< 2e-16 ***	0.480
CountryAU	-1.234e-01	1.018e-01	1.876e+03	-1.213	0.225469	
CountryBG	-3.737e-02	6.830e-02	1.880e+03	-0.547	0.584321	
CountryBR	8.501e-02	8.216e-02	1.878e+03	1.035	0.300939	
CountryCM	-1.420e-01	8.690e-02	1.878e+03	-1.634	0.102445	
CountryCO	1.063e-01	1.146e-01	1.877e+03	0.928	0.353747	
CountryCZ	-1.467e-01	8.722e-02	1.878e+03	-1.682	0.092771 .	
CountryES	9.634e-03	6.978e-02	1.876e+03	0.138	0.890202	
CountryID	-8.746e-02	6.744e-02	1.880e+03	-1.297	0.194838	
CountryIN	1.470e-02	7.303e-02	1.882e+03	0.201	0.840448	
CountryIT	-2.432e-02	6.597e-02	1.989e+03	-0.369	0.712430	
CountryJP	-2.351e-01	6.132e-02	1.880e+03	-3.833	0.000131 ***	-0.205
CountryNG	-7.398e-02	7.471e-02	1.879e+03	-0.99	0.322211	
CountryPH	-6.087e-02	7.539e-02	1.880e+03	-0.807	0.419529	
CountryTH	-6.954e-02	7.514e-02	1.881e+03	-0.925	0.354832	
CountryTR	1.614e-01	1.496e-01	1.877e+03	1.079	0.280707	
CountryTW	-7.210e-02	7.538e-02	1.881e+03	-0.956	0.338971	
CountryUK	-1.577e-01	7.606e-02	1.876e+03	-2.073	0.038301 *	-0.138
CountryUS	-1.184e-01	5.752e-02	1.881e+03	-2.059	0.039599 *	-0.103
CountryVN	-1.397e-01	9.694e-02	1.878e+03	-1.441	0.149782	
eWITHOUT:SLS	4.463e-02	5.006e-02	4.353e+04	0.891	0.372671	
eWITHOUT:BLIS	2.067e-02	1.695e-01	2.565e+03	0.122	0.902951	
SLS:BLIS	1.225e+00	9.136e-02	4.349e+04	13.406	< 2e-16 ***	1.071
eWITHOUT:SLS:BLIS	-3.898e-01	1.302e-01	4.351e+04	-2.994	0.002755 **	-0.341

Signif. codes: *** [0, 0.001], ** (0.001, 0.01], * (0.01, 0.05], . (0.05, 0.1]

over 80 empirical studies, none of which included statistical literacy as a factor influencing trust in AI.

This study is the first to examine the relationship between statistical literacy and trust in algorithms, revealing a nuanced relationship that depends on context. Our findings demonstrate that statistical literacy has opposite effects in different scenarios: it increases trust in algorithmic decisions for low-stakes situations while decreasing trust for high-stakes decisions. This differential effect suggests that statistical literacy enables a more sophisticated understanding of algorithmic capabilities and limitations. In low-stakes scenarios (such as restaurant recommendations or music suggestions), individuals with higher statistical literacy appear to recognize that algorithmic predictions based on pattern recognition and large datasets can be effective and reliable. However, in high-stakes contexts (such as employment or criminal justice decisions), this same statistical knowledge leads to greater skepticism - not because the algorithms are necessarily less accurate, but because statistically literate individuals better understand the potential consequences of algorithmic biases and limitations. Those with statistical literacy are better equipped to understand that while statistical models may achieve high average accuracy, they can still fail in critical individual cases or perpetuate systemic biases present in training data. This cautious approach to high-stakes algorithmic decisions reflects not just critical thinking, but a deeper understanding of

Table 2. Random effects for the linear mixed model. The variance explained by the random factors (estimated via the function `gstudy` in the `gtheory` R package) were: ID=16.3% and Item 2.3%.

Groups	Name	Variance	Std.Dev.
ID	(Intercept)	0.21380	0.4624
Item	(Intercept)	0.03076	0.1754
Residual		1.06489	1.0319

Table 3. Analysis of Deviance Table (Type III Wald χ^2 tests) for the fixed effects of the model with the best fit.

	χ^2	Df	$\Pr(> \chi^2)$
(Intercept)	160.5147	1	< 2.2e-16 ***
e	0.9221	1	0.3369223
S	49.4186	1	2.068e-12 ***
BLIS	22.2199	1	2.431e-06 ***
Age	18.3263	1	1.861e-05 ***
Gender	18.3575	1	1.831e-05 ***
ADA	1901.9323	1	< 2.2e-16 ***
Country	47.7956	19	0.0002746 ***
e:S	0.7948	1	0.3726661
e:BLIS	0.0149	1	0.9029413
S:BLIS	179.7312	1	< 2.2e-16 ***
e:S:BLIS	8.9643	1	0.0027531 **

Signif. codes: *** [0, 0.001], ** (0.001, 0.01], * (0.01, 0.05], . (0.05, 0.1]

how statistical methods work and where they may fall short.

Paradoxically, explainability only affected people's trust in algorithms when it was absent, the stakes were low, and statistical literacy was high. This contradicts previous findings in the literature, which have shown that interventions focused on explaining the decision-making processes of algorithms can increase the use of and trust in algorithms, for example in healthcare [86], journalism [87] and military settings [88,89]. One possible reason for this inconsistency could be due to the way we operationalised "explainability" in our study, where the explanations included technical jargon that may have exceeded the expected level of familiarity among participants. However, this may also mean that the information related to the explainability of the algorithm is not related to trust or distrust in the algorithm. Rather than focusing on how an algorithm works, our results suggest that statistically literate individuals primarily consider what the algorithm is being used for - its purpose and potential impact - when deciding whether to trust it. This finding challenges the common assumption that greater algorithmic transparency necessarily leads to more appropriate trust calibration.

Over time, the concept of statistical literacy has evolved from the understanding and application of statistical techniques to a broader understanding explicitly related to trust in algorithms. Algorithms now consist of thousands of lines of formulae and are increasingly used to make decisions that may be difficult for humans to understand (known as the black box effect). Consequently, statistical literacy now encompasses not only the ability to understand statistical output, but also the skills needed to critically interpret and evaluate statistical information and reasoning, which requires a higher degree of critical thinking. Therefore, the promotion of statistical literacy is essential to ensure that individuals have the necessary skills to understand and interpret statistical information and algorithms and to become critical users of ADA. Furthermore, our findings have important implications for policymakers and educators, who should consider incorporating statistical literacy training into school curricula and professional development programs. This can help ensure that individuals are equipped

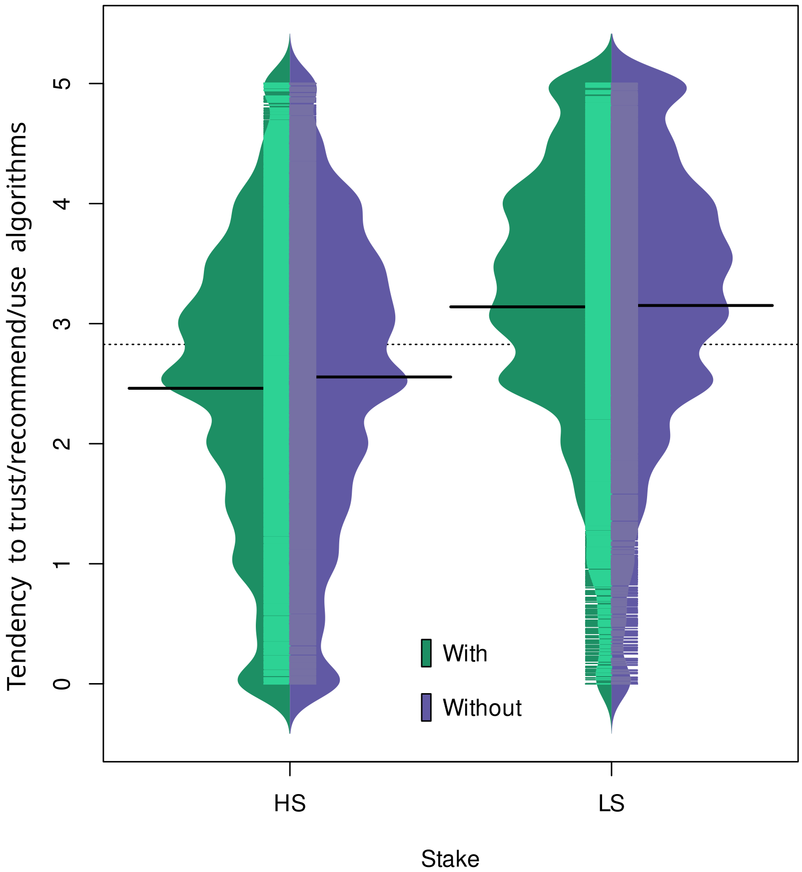

Figure 4. Beanplots showing the tendency to trust/recommend/use algorithms as a function of explainability (with or without) and situation stake (high stake = HS or low stake = LS). This figure shows the main effect of the stake factor (S) and the non-significant effect of explainability (e) (recall that this variable was not significant but used for illustrative purposes). The dotted horizontal line represents the grand mean and the four solid horizontal lines represent the groups' means.

with the skills they need to navigate an increasingly data-driven world and make informed decisions based on statistical information and algorithms (but see section 'implications and limitations' below).

Our results showed that older people and men were less likely to trust algorithms than younger people and women. Previous research has shown that certain demographic groups are more likely to trust algorithms than others. However, previous studies have shown that older people tend to trust ADA more than younger people, while gender has been shown to have inconsistent effects (see for example [90,91]). These differences may be due to particular characteristics of the study participants, possibly influenced by a bias towards certain aspects of the topic at hand.

In our cross-country analysis, we observed variations in trust in algorithms, with industrialised countries such as Japan, the US, and the UK exhibiting lower levels of trust in AI. This finding aligns with a recent study on trust in AI by Gillespie et al. [34], which reported that Japan had one of the lowest levels of trust in AI, while

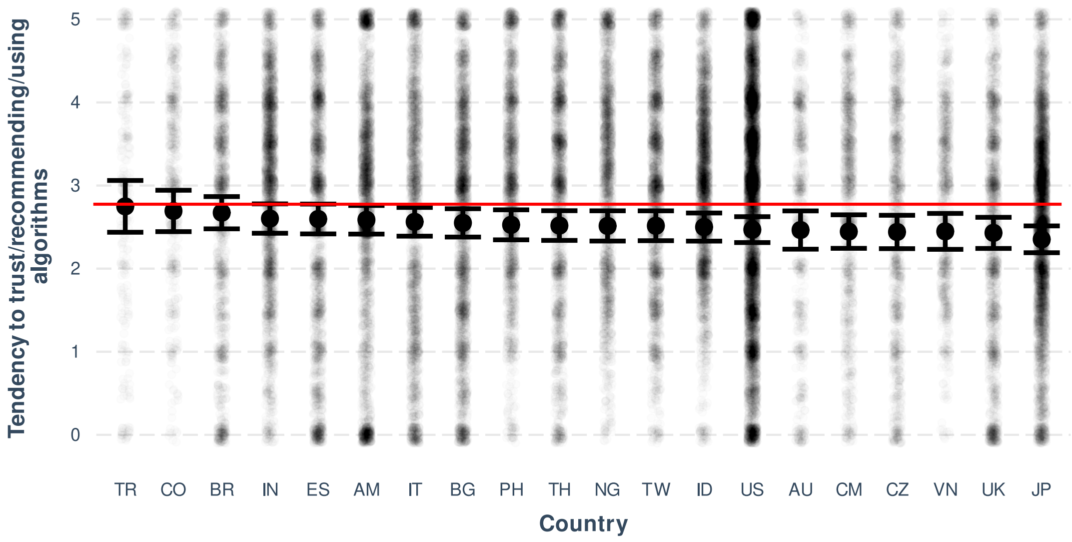

Figure 5. Plot showing the variability in the tendency to trust/recommend/use algorithms across countries. Countries are labelled with Turkey: TR, Colombia: CO, Brazil: BR, India: IN, Spain: ES, Armenia: AM, Italy: IT, Bulgaria: BG, Philippines: PH, Thailand: TH, Nigeria: NG, Taiwan: TW, Indonesia: ID, USA: US, Australia: AU, Cameroon: CM, Czech Republic: CZ, Vietnam: VN, UK: UK, and Japan: JP. The most important predictors for all models in each country were S followed by ADA and BLIS. Error bars represent 95% confidence intervals around the mean. The horizontal line indicates the overall mean. Although the substantial overlap of the confidence intervals suggests no significant statistical pairwise differences, the focus is on ranking countries based on their average tendency to trust algorithms.

the US and the UK had intermediate levels. Interestingly, countries such as India and Brazil, which demonstrated high levels of trust in the Gillespie et al. study (see Figure 2 in their report), appear in our linear mixed model with positive estimates (see table 1 and Figure 5), although not statistically significant. This suggests that different methodologies may yield varying perceptions of trust levels across countries.

Implications and limitations

Various machine learning techniques require *data work* or human intervention in the form of data generation, annotation and algorithmic verification [47]. This labour-intensive process is often distributed to teams in business process outsourcing companies (BPOs) or to individuals through labour platforms, reducing production costs [92]. Miceli and Posada [93] studied one BPO in Argentina and three platforms operating in Venezuela and found that the discourses and social relations that structured data work were aimed at controlling workers (through managerial approaches in the BPO and algorithms in the platforms) to increase productivity and reduce worker "bias". The problem is that feedback from workers was discouraged and, by taking clients' decisions as "ground truth", the data production process reproduced clients' biases, which were carried out by algorithms trained on that data. Their research concluded that the quality of the data depended on the voice and engagement of workers, which in turn required decent working conditions and recognition. Even if the data used in the algorithm is well annotated and leads to good algorithmic performance, there is the question of the human ability to interpret these results, as human judgments are modulated by social-emotional processes [21,94–96]. Future work should consider the

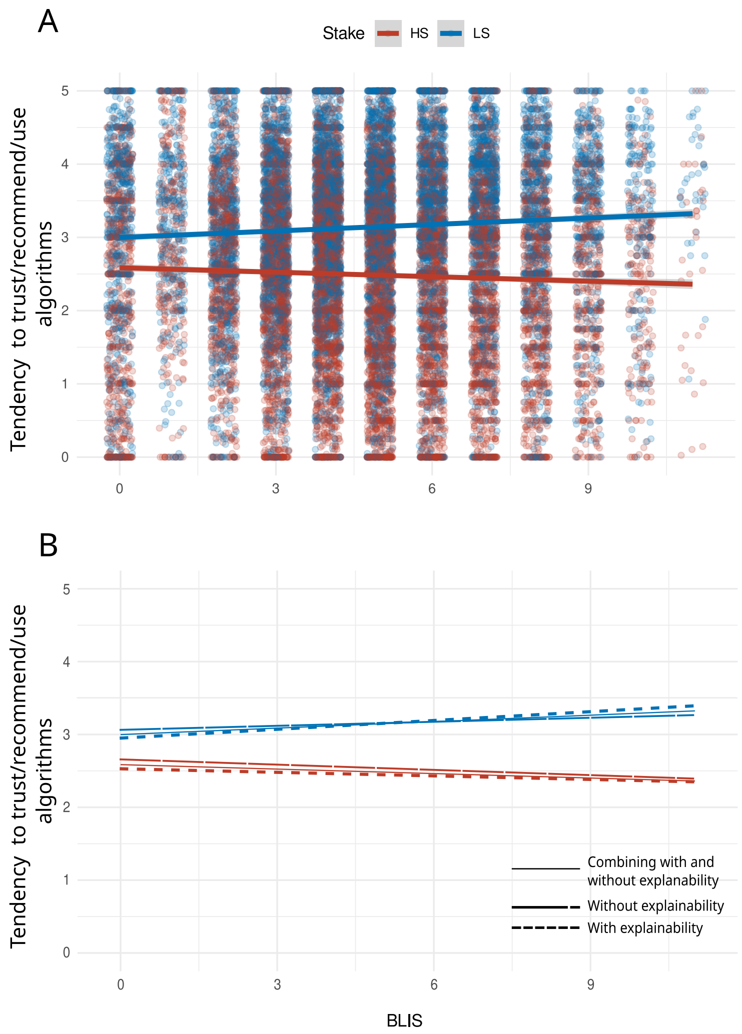

Figure 6. Scatterplot showing the correlation between BLIS scores, explainability and the tendency to trust/recommend/use algorithms as a function of stake level. This figure illustrates the interaction between stake (high stake = HS or low stake = LS) and statistical literacy (`BLIS`) according to the level of explainability of algorithms (e). The observations on the x-axis are jittered for visualisation purposes.

human and social aspects of data production and make the work visible in documentation efforts [97]. This transparency of the social aspects of datasets will contribute to trust in the operation of algorithms.

While the current findings are indeed informative, it is important to recognize certain limitations that may constrain the generalizability of these results and claims [98]. We argued that statistical literacy influences trust in both low- and high-stakes scenarios; however, it could be part of a broader understanding of technology, algorithms, and data. Indeed, statistical literacy could be considered a sub-skill of AI literacy if AI literacy is understood as the ability to recognize, understand, use, and critically evaluate AI technologies and their societal impacts, supported by foundational knowledge in statistics and computing. Therefore, policymakers should consider promoting AI literacy to address some of the complexities associated with trust in algorithms.

Our study utilized self-reported measures via rating scales, which are efficient and cost-effective for capturing data on thoughts, feelings, and subjective experiences. However, these measures can be influenced by social desirability, response bias, misinterpretation, or lack of self-awareness. For instance, physiological research has shown that self-reported measures of physical activity can both overestimate and underestimate actual levels of physical activity [99]. Therefore, future extensions of this work should consider a more robust approach, such as triangulating the data with direct observations of user interactions with algorithms or physiological measures to assess trust more accurately.

High-stakes and low-stakes situations exhibit significant variability across individuals and cultures, existing on a context-dependent continuum rather than as discrete categories. For example, choosing a restaurant for dinner with friends may carry different stakes across cultural contexts, socioeconomic backgrounds, and individual preferences. Our study's primary limitation lies in not systematically investigating how participants from different backgrounds interpreted and classified these scenarios. Additionally, while our sample included participants from 20 countries, certain geographical regions like Central Europe were underrepresented, potentially limiting the generalizability of our findings across different cultural contexts. Although we aimed to move beyond WEIRD (Western, Educated, Industrialized, Rich, and Democratic) sampling biases, more comprehensive geographic and cultural representation, along with larger sample sizes from each region, would be necessary to make broader generalizations about algorithmic trust across diverse populations [100,101]. Future research should incorporate scenario validation across different cultural contexts and expand sampling to include currently underrepresented regions and demographic groups.

Conclusion

This study investigated the personal and algorithmic factors that affect individuals' trust in algorithms. Our findings revealed that when the stakes are low, statistical literacy is positively correlated with the likelihood of trusting an algorithm. However, when the stakes are high, our results indicated a negative correlation between statistical literacy and the likelihood of trusting an algorithm. Therefore, we conclude that having statistical literacy enables individuals to critically evaluate the decisions made by ADA and consider them alongside other factors before making significant life decisions. This ensures that individuals are not solely relying on algorithms that may not fully capture the complexity and nuances of human behaviour and decision-making.

Disclosure statement

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data availability statement

The data that support the findings of this study as well as the analysis scripts in R are openly available at `https://figshare.com/projects/Trust_in_algorithms_An_experimental_approach_-_Data_repository/156212`

Author contributions statement

Conceptualisation: FM-R; Methodology: FM-R, JT, MTL, and RG; Software: JT, MTL, and RG; Formal analysis: JT, MTL, and RG; Investigation: all authors; Resources: all authors; Data curation: JT, MTL, and RG; Writing - Original Draft: FM-R and JT; Writing - Review Editing: all authors; Visualisation: JT; Supervision: FM-R; Project administration: FM-R.

An earlier version of this manuscript can be found at `https://osf.io/preprints/psyarxiv/9wh2f`. We recommend referring to and citing the current version rather than the earlier one.

References

[1] Engin Z, Treleaven P. Algorithmic Government: Automating Public Services and Supporting Civil Servants in using Data Science Technologies. The Computer Journal. 2019 Mar;62(3):448-60. Available from: https://doi.org/10.1093/comjnl/bxy082. doi:10.1093/comjnl/bxy082.

[2] Vasconcelos GL, Macêdo AMS, Ospina R, Almeida FAG, Duarte-Filho GC, Brum AA, et al.. Modelling fatality curves of COVID-19 and the effectiveness of intervention strategies. medRxiv; 2020. ISSN: 2005-1557 Pages: 2020.04.02.20051557. Available from: https://www.medrxiv.org/content/10.1101/2020.04.02.20051557v3. doi:10.1101/2020.04.02.20051557.

[3] Bullock J, Luccioni A, Pham KH, Lam CSN, Luengo-Oroz M. Mapping the landscape of Artificial Intelligence applications against COVID-19. Journal of Artificial Intelligence Research. 2020 Nov;69:807-45. Available from: https://www.jair.org/index.php/jair/article/view/12162. doi:10.1613/jair.1.12162.

[4] Beer D. The Social Power of Algorithms. Information, Communication & Society. 2017;20(1). doi:10.1080/1369118X.2016.1216147.

[5] Toros H, Flaming D. Prioritizing Homeless Assistance Using Predictive Algorithms: An Evidence-Based Approach. Cityscape. 2018;20(1):117-46. Publisher: US Department of Housing and Urban Development. Available from: https://www.jstor.org/stable/26381224.

[6] Burdick H, Lam C, Mataraso S, Siefkas A, Braden G, Dellinger RP, et al. Prediction of respiratory decompensation in Covid-19 patients using machine learning: The READY trial. Computers in Biology and Medicine. 2020 Sep;124:103949. Available from: https://www.sciencedirect.com/science/article/pii/S0010482520302845. doi:10.1016/j.compbiomed.2020.103949.

[7] Ertel W. Introduction to Artificial Intelligence. Undergraduate Topics in Computer Science. Cham: Springer International Publishing; 2017. Available from: http://link.springer.com/10.1007/978-3-319-58487-4. doi:10.1007/978-3-319-58487-4.

[8] Samuel AL. Some Studies in Machine Learning Using the Game of Checkers. IBM Journal of Research and Development. 1959 Jul;3(3):210-29. Conference Name: IBM Journal of Research and Development. Available from: https://ieeexplore.ieee.org/document/5392560. doi:10.1147/rd.33.0210.

[9] Jordan MI, Mitchell TM. Machine learning: Trends, perspectives, and prospects. Science. 2015 Jul;349(6245):255-60. Publisher: American Association for the Advancement of Science. Available from: https://www.science.org/doi/10.1126/science.aaa8415. doi:10.1126/science.aaa8415.

[10] Cave S, Craig C, Dihal K, Dillon S, Montgomery J, Singler B, et al. Portrayals and

perceptions of AI and why they matter. In: DES5612. The Royal Society; 2018. Available from: https://royalsociety.org/~/media/policy/projects/ai-narratives/AI-narratives-workshop-findings.pdf.

[11] Dawson D, Schleiger E, Horton J, McLaughlin J, Robinson C, Quezada G, et al. Artificial Intelligence: Australia's Ethics Framework. Data61 CSIRO, Australia. 2019. Available from: https://www.csiro.au/-/media/D61/Reports/Artificial-Intelligence-ethics-framework.pdf.

[12] Whittlestone J, Nyrup R, Alexandrova A, Cave S. The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions. In: Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society. AIES '19. New York, NY, USA: Association for Computing Machinery; 2019. p. 195-200. Available from: https://doi.org/10.1145/3306618.3314289. doi:10.1145/3306618.3314289.

[13] Shanmuganathan M. Behavioural finance in an era of artificial intelligence: Longitudinal case study of robo-advisors in investment decisions. Journal of Behavioral and Experimental Finance. 2020 Sep;27:100297. Available from: https://www.sciencedirect.com/science/article/pii/S221463501930214X. doi:10.1016/j.jbef.2020.100297.

[14] Siemens G, Marmolejo-Ramos F, Gabriel F, Medeiros K, Marrone R, Joksimovic S, et al. Human and artificial cognition. Computers and Education: Artificial Intelligence. 2022:100107. doi:10.1016/j.caeai.2022.100107.

[15] DeVito MA. From Editors to Algorithms. Digital Journalism. 2017 Jul;5(6):753-73. Publisher: Routledge _eprint: https://doi.org/10.1080/21670811.2016.1178592. Available from: https://doi.org/10.1080/21670811.2016.1178592. doi:10.1080/21670811.2016.1178592.

[16] Werner A. Organizing music, organizing gender: algorithmic culture and Spotify recommendations. Popular Communication. 2020 Jan;18(1):78-90. Publisher: Routledge _eprint: https://doi.org/10.1080/15405702.2020.1715980. Available from: https://doi.org/10.1080/15405702.2020.1715980. doi:10.1080/15405702.2020.1715980.

[17] Tenemaza M, Luján-Mora S, De Antonio A, Ramírez J. Improving Itinerary Recommendations for Tourists Through Metaheuristic Algorithms: An Optimization Proposal. IEEE Access. 2020;8:79003-23. Conference Name: IEEE Access. doi:10.1109/ACCESS.2020.2990348.

[18] Tsakanikas V, Dagiuklas T. Video surveillance systems-current status and future trends. Computers & Electrical Engineering. 2018 Aug;70:736-53. Available from: https://www.sciencedirect.com/science/article/pii/S0045790617311813. doi:10.1016/j.compeleceng.2017.11.011.

[19] Modi Y, Teli R, Mehta A, Shah K, Shah M. A comprehensive review on intelligent traffic management using machine learning algorithms. Innovative Infrastructure Solutions. 2021 Dec;7(1):128. Available from: https://doi.org/10.1007/s41062-021-00718-3. doi:10.1007/s41062-021-00718-3.

[20] Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, et al.. Language Models are Few-Shot Learners. arXiv; 2020. ArXiv:2005.14165 [cs]. Available from: http://arxiv.org/abs/2005.14165. doi:10.48550/arXiv.2005.14165.

[21] Clark E, August T, Serrano S, Haduong N, Gururangan S, Smith NA. All That's 'Human' Is Not Gold: Evaluating Human Evaluation of Generated Text. arXiv; 2021. ArXiv:2107.00061 [cs]. Available from: http://arxiv.org/abs/2107.00061. doi:10.48550/arXiv.2107.00061.

[22] Dehouche N. Plagiarism in the age of massive Generative Pre-trained Transformers (GPT-3). Ethics in Science and Environmental Politics. 2021 Mar;21:17-23. Available from: https://www.int-res.com/abstracts/esep/v21/p17-23/. doi:10.3354/esep00195.

[23] Ugli MIB. Will Human Beings Be Superseded by Generative Pre-trained Transformer 3 (GPT-3) in Programming? International Journal on Orange Technologies. 2020 Oct;2(10):141-3. Publisher: Research Parks Publishing. Available from: https://www.neliti.com/publications/333438/. doi:10.31149/ijot.v2i10.769.

[24] Drakopoulos G, Kafeza E, Mylonas P, al Katheeri H. Building Trusted Startup Teams From LinkedIn Attributes: A Higher Order Probabilistic Analysis. In: 2020 IEEE 32nd International Conference on Tools with Artificial Intelligence (ICTAI); 2020. p. 867-74. ISSN: 2375-0197. doi:10.1109/ICTAI50040.2020.00136.

[25] Rawat S, Rawat A, Kumar D, Sabitha AS. Application of machine learning and data visualization techniques for decision support in the insurance sector. International Journal of Information Management Data Insights. 2021 Nov;1(2):100012. Available from: https://www.sciencedirect.com/science/article/pii/S2667096821000057. doi:10.1016/j.jjimei.2021.100012.

[26] Badue C, Guidolini R, Carneiro RV, Azevedo P, Cardoso VB, Forechi A, et al. Self-driving cars: A survey. Expert Systems with Applications. 2021 Mar;165:113816. Available from: https://www.sciencedirect.com/science/article/pii/S095741742030628X. doi:10.1016/j.eswa.2020.113816.

[27] Dwork C, Minow M. Distrust of Artificial Intelligence: Sources & Responses from Computer Science & Law. Daedalus. 2022 May;151(2):309-21. Available from: https://doi.org/10.1162/daed_a_01918. doi:10.1162/daeda01918.

[28] Debad SJ, editor. Learning from the Science of Cognition and Perception for Decision Making: Proceedings of a Workshop. Illustrated edição ed. Washington, DC: National Academies Press; 2018.

[29] Prahl A, Van Swol L. Understanding algorithm aversion: When is advice from automation discounted? Journal of Forecasting. 2017 Sep;36(6):691-702. Publisher: John Wiley & Sons, Ltd. Available from: https://onlinelibrary-wiley.ez20.periodicos.capes.gov.br/doi/full/10.1002/for.2464. doi:10.1002/for.2464.

[30] Granulo A, Fuchs C, Puntoni S. Psychological reactions to human versus robotic job replacement. Nature Human Behaviour. 2019 Oct;3(10):1062-9. Number: 10 Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41562-019-0670-y. doi:10.1038/s41562-019-0670-y.

[31] Frank MR, Autor D, Bessen JE, Brynjolfsson E, Cebrian M, Deming DJ, et al. Toward understanding the impact of artificial intelligence on labor. Proceedings of the National Academy of Sciences. 2019 Apr;116(14):6531-9. Publisher: Proceedings of the National Academy of Sciences. Available from: https://www.pnas.org/doi/10.1073/pnas.1900949116. doi:10.1073/pnas.1900949116.

[32] Vayena E, Blasimme A, Cohen IG. Machine learning in medicine: Addressing ethical challenges. PLOS Medicine. 2018 Nov;15(11):e1002689. Publisher: Public Library of Science. Available from: https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002689. doi:10.1371/journal.pmed.1002689.

[33] Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring Fairness in Machine Learning to Advance Health Equity. Annals of Internal Medicine. 2018 Dec;169(12):866-72. Publisher: American College of Physicians. Available from: https://www.acpjournals.org/doi/abs/10.7326/M18-1990. doi:10.7326/M18-1990.

[34] Gillespie N, Lockey S, Curtis C, Pool J, Akbari A. Trust in Artificial Intelligence: A global study. The University of Queensland and KPMG Australia; 2023. Available from: https://ai.uq.edu.au/project/trust-artificial-intelligence-global-study. doi:10.14264/00d3c94.

[35] Rosen LD, Maguire P. Myths and realities of computerphobia: A meta-analysis. Anxiety Research. 1990 Nov;3(3):175-91. Publisher: Routledge _eprint: https://doi.org/10.1080/08917779008248751. Available from: https://doi.org/10.1080/08917779008248751. doi:10.1080/08917779008248751.

[36] Kim J. Fear of Artificial Intelligence on People's Attitudinal & Behavioral Attributes: An Exploratory Analysis of A.I. Phobia. International European Extended Enablement in Science, Engineering & Management (IEEESEM). 2019 Oct;7(10):9-20.

[37] Khasawneh OY. Technophobia: Examining its hidden factors and defining it. Technology in Society. 2018 Aug;54:93-100. Available from: https://www.sciencedirect.com/science/article/pii/S0160791X17302968. doi:10.1016/j.techsoc.2018.03.008.

[38] Human-AI Teaming: State-of-the-Art and Research Needs. Washington, D.C.: National Academies Press; 2022. Available from: https://www.nap.edu/catalog/26355. doi:10.17226/26355.
[39] Warren A, Hillas A. Friend or frenemy? The role of trust in human-machine teaming and lethal autonomous weapons systems. Small Wars & Insurgencies. 2020 May;31(4):822-50. Publisher: Routledge _eprint: https://doi.org/10.1080/09592318.2020.1743485. Available from: https://doi.org/10.1080/09592318.2020.1743485. doi:10.1080/09592318.2020.1743485.
[40] Liang Y, Lee SA. Fear of Autonomous Robots and Artificial Intelligence: Evidence from National Representative Data with Probability Sampling. International Journal of Social Robotics. 2017 Jun;9(3):379-84. Available from: https://doi.org/10.1007/s12369-017-0401-3. doi:10.1007/s12369-017-0401-3.
[41] Gnambs T, Appel M. Are robots becoming unpopular? Changes in attitudes towards autonomous robotic systems in Europe. Computers in Human Behavior. 2019 Apr;93:53-61. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0747563218305806. doi:10.1016/j.chb.2018.11.045.
[42] Luscombe R. Google engineer put on leave after saying AI chatbot has become sentient. The Guardian. 2022 Jun. Available from: https://www.theguardian.com/technology/2022/jun/12/google-engineer-ai-bot-sentient-blake-lemoine.
[43] Dale R. GPT-3: What's it good for? Natural Language Engineering. 2021 Jan;27(1):113-8. Publisher: Cambridge University Press. Available from: https://www.cambridge.org/core/journals/natural-language-engineering/article/gpt3-whats-it-good-for/0E05CFE68A7AC8BF794C8ECBE28AA990. doi:10.1017/S1351324920000601.
[44] Adams R. Can artificial intelligence be decolonized? Interdisciplinary Science Reviews. 2021 Apr;46(1-2):176-97. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/03080188.2020.1840225. Available from: https://doi.org/10.1080/03080188.2020.1840225. doi:10.1080/03080188.2020.1840225.
[45] Bender EM, Gebru T, McMillan-Major A, Shmitchell S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency. FAccT '21. New York, NY, USA: Association for Computing Machinery; 2021. p. 610-23. Available from: https://doi.org/10.1145/3442188.3445922. doi:10.1145/3442188.3445922.
[46] Olhede SC, Wolfe PJ. The growing ubiquity of algorithms in society: implications, impacts and innovations. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences. 2018 Sep;376(2128):20170364. Publisher: Royal Society. Available from: https://royalsocietypublishing.org/doi/full/10.1098/rsta.2017.0364. doi:10.1098/rsta.2017.0364.
[47] Tubaro P, Casilli AA, Coville M. The trainer, the verifier, the imitator: Three ways in which human platform workers support artificial intelligence. Big Data & Society. 2020;7(1). Available from: http://journals.sagepub.com/doi/10.1177/2053951720919776. doi:10.1177/2053951720919776.
[48] Crépel M, Do S, Cointet JP, Cardon D, Bouachera Y. Mapping AI Issues in Media Through NLP Methods. In: Ehrmann M, Karsdorp F, Wevers M, Andrews TL, Burghardt M, Kestemont M, et al., editors. Proceedings of the Conference on Computational Humanities Research 2021. vol. 2989 of CEUR Workshop Proceedings. Amsterdam, the Netherlands: CEUR; 2021. p. 77-91. ISSN: 1613-0073. Available from: http://ceur-ws.org/Vol-2989/#long_paper22.
[49] Marmolejo-Ramos F, Workman T, Walker C, Lenihan D, Moulds S, Correa J, et al. AI-powered narrative building for facilitating public participation and engagement. Disover Artificial Intelligence. 2022;2(7). doi:https://doi.org/10.1007/s44163-022-00023-7.
[50] Kang J, Kim J, Yang M, Park E, Ko M, Lee M, et al. Behind the scenes of K-pop fandom: unveiling K-pop fandom collaboration network. Quality & Quantity. 2022 Jun;56(3):1481-502. Available from: https://doi.org/10.1007/s11135-021-01189-5.

doi:10.1007/s11135-021-01189-5.
[51] Milton Beck. Milton Beck on LinkedIn: Atualizamos nossa política global de anúncios de vagas para permitir; 2022. LinkedIn. Available from: https://www.c.com/posts/miltonbeck_atualizamos-nossa-pol%C3%ADtica-global-de-an%C3%BAncios-activity-6914622832966533123-SMBM.
[52] François K, Monteiro C, Allo P. Big-data literacy as a new vocation for statistical literacy. Statistics Education Research Journal. 2020;19:194-205. Available from: https://www.iase-web.org/ojs/SERJ/article/view/130. doi:10.52041/serj.v19i1.130.
[53] Haack DG. Statistical Literacy: A Guide to Interpretation. North Scituate, Massachusetts: Duxbury; 1979.
[54] Wallman KK. Enhancing Statistical Literacy: Enriching Our Society. Journal of the American Statistical Association. 1993;88(421):1-8. Publisher: [American Statistical Association, Taylor & Francis, Ltd.]. Available from: https://www.jstor.org/stable/2290686. doi:10.2307/2290686.
[55] Watson J. Assessing Statistical Thinking Using the Media. In: The Assessment Challenge in Statistics Education. Amsterdam: IOSPress; 1997. .
[56] Gal I. Adults' Statistical Literacy: Meanings, Components, Responsibilities. International Statistical Review / Revue Internationale de Statistique. 2002;70(1):1-25. Publisher: [Wiley, International Statistical Institute (ISI)]. Available from: https://www.jstor.org/stable/1403713. doi:10.2307/1403713.
[57] OECD. PISA 2012 Assessment and Analytical Framework: Mathematics, Reading, Science, Problem Solving and Financial Literacy. Paris: Organisation for Economic Co-operation and Development; 2013. Available from: https://www.oecd-ilibrary.org/education/pisa-2012-assessment-and-analytical-framework_9789264190511-en.
[58] Friedrich S, Antes G, Behr S, Binder H, Brannath W, Dumpert F, et al. Is there a role for statistics in artificial intelligence? Advances in Data Analysis and Classification. 2022 Dec;16(4):823-46. Available from: https://doi.org/10.1007/s11634-021-00455-6. doi:10.1007/s11634-021-00455-6.
[59] Carvalho DV, Pereira EM, Cardoso JS. Machine Learning Interpretability: A Survey on Methods and Metrics. Electronics. 2019 Aug;8(8):832. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute. Available from: https://www.mdpi.com/2079-9292/8/8/832. doi:10.3390/electronics8080832.
[60] Du M, Liu N, Hu X. Techniques for Interpretable Machine Learning. arXiv; 2019. Number: arXiv:1808.00033 arXiv:1808.00033 [cs, stat]. Available from: http://arxiv.org/abs/1808.00033. doi:10.48550/arXiv.1808.00033.
[61] Pasquale PoLUoMF. The Black Box Society: The Secret Algorithms That Control Money and Information. Reprint edição ed. Cambridge, Massachusetts London, England: Harvard University Press; 2016.
[62] Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, et al. Model Cards for Model Reporting. In: Proceedings of the Conference on Fairness, Accountability, and Transparency. FAT* '19. New York, NY, USA: Association for Computing Machinery; 2019. p. 220-9. Available from: https://doi.org/10.1145/3287560.3287596. doi:10.1145/3287560.3287596.
[63] Wang A, Liu A, Zhang R, Kleiman A, Kim L, Zhao D, et al. REVISE: A Tool for Measuring and Mitigating Bias in Visual Datasets. International Journal of Computer Vision. 2022 Jul;130(7):1790-810. Available from: https://doi.org/10.1007/s11263-022-01625-5. doi:10.1007/s11263-022-01625-5.
[64] d'Eon G, d'Eon J, Wright JR, Leyton-Brown K. The Spotlight: A General Method for Discovering Systematic Errors in Deep Learning Models. In: 2022 ACM Conference on Fairness, Accountability, and Transparency. FAccT '22. New York, NY, USA: Association for Computing Machinery; 2022. p. 1962-81. Available from: https://doi.org/10.1145/3531146.3533240. doi:10.1145/3531146.3533240.
[65] Herzog C. On the risk of confusing interpretability with explicability. AI and Ethics. 2022

Feb;2(1):219-25. Available from: https://doi.org/10.1007/s43681-021-00121-9. doi:10.1007/s43681-021-00121-9.

[66] Forty-two countries adopt new OECD Principles on Artificial Intelligence - OECD; 2019. Available from: https://www.oecd.org/science/forty-two-countries-adopt-new-oecd-principles-on-artificial-intelligence.htm.

[67] Hagendorff T. The Ethics of AI Ethics: An Evaluation of Guidelines. Minds and Machines. 2020 Mar;30(1):99-120. Available from: https://doi.org/10.1007/s11023-020-09517-8. doi:10.1007/s11023-020-09517-8.

[68] Mittelstadt B. Principles alone cannot guarantee ethical AI. Nature Machine Intelligence. 2019 Nov;1(11):501-7. Number: 11 Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s42256-019-0114-4. doi:10.1038/s42256-019-0114-4.

[69] Merritt SM, Heimbaugh H, LaChapell J, Lee D. I Trust It, but I Don't Know Why: Effects of Implicit Attitudes Toward Automation on Trust in an Automated System. Human Factors. 2013 Jun;55(3):520-34. Publisher: SAGE Publications Inc. Available from: https://doi.org/10.1177/0018720812465081. doi:10.1177/0018720812465081.

[70] Ziegler L, Garfield J. Developing a statistical literacy assessment for the modern introductory statistics course. Statistics Education Research Journal. 2018 Nov;17(2):161-78. Number: 2. Available from: https://iase-pub.org/ojs/SERJ/article/view/164. doi:10.52041/serj.v17i2.164.

[71] Kuznetsova A, Christensen R, Bavay C, Brockhoff P. Automated mixed ANOVA modeling of sensory and consumer data. Food Quality and Preference. 2015;40(A):31-8. doi:10.1016/j.foodqual.2014.08.004.

[72] Kuznetsova A, Brockhoff P, Christensen R. lmerTest Package: Tests in Linear Mixed Effects Models. Journal of Statistical Software. 2017;82(13):1-26. Available from: https://doi.org/10.18637/jss.v082.i13. doi:10.18637/jss.v082.i13.

[73] Wagenmakers EJ, Farrell S. AIC model selection and Akaike weights. Psychonomic Bulletin & Review. 2004;11(1):192-6. doi:10.3758/BF03206482.

[74] Lüdecke D, Ben-Shachar M, Patil I, Waggoner P, Makowski D. performance: An R Package for Assessment, Comparison and Testing of Statistical Models. Journal of Open Source Software. 2021;6(60):3139. doi:10.21105/joss.03139.

[75] Huebner A, Lucht M. Generalizability Theory in R. Practical Assessment, Research, and Evaluation. 2019 Nov;24(1). Number: 1 Publisher: University of Massachusetts Amherst Libraries. doi:10.7275/5065-gc10.

[76] Peña EA, Slate EH. Global Validation of Linear Model Assumptions. Journal of the American Statistical Association. 2006;101(473):341-54. doi:10.1198/016214505000000637.

[77] Koller M. robustlmm: An R Package for Robust Estimation of Linear Mixed-Effects Models. Journal of Statistical Software. 2016;75(6):1-24. doi:10.18637/jss.v075.i06.

[78] Schielzeth H, Dingemanse NJ, Nakagawa S, Westneat DF, Allegue H, Teplitsky C, et al. Robustness of linear mixed-effects models to violations of distributional assumptions. Methods in Ecology and Evolution. 2020;11(9):1141-52. doi:10.1111/2041-210X.13434.

[79] Nakagawa S, Johnson PCD, Schielzeth H. The coefficient of determination R^2 and intraclass correlation coefficient from generalized linear mixed-effects models revisited and expanded. Journal of The Royal Society Interface. 2017 Sep;14(134):20170213. Publisher: Royal Society. Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2017.0213. doi:10.1098/rsif.2017.0213.

[80] Brysbaert M, Stevens M. Power Analysis and Effect Size in Mixed Effects Models: A Tutorial. Journal of Cognition. 2018 Jan;1(1):9. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC6646942/. doi:10.5334/joc.10.

[81] Lee MK. Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data & Society. 2018 Jan;5(1):2053951718756684. Publisher: SAGE Publications Ltd. Available from: https:

//doi.org/10.1177/2053951718756684. doi:10.1177/2053951718756684.

[82] Araujo T, Helberger N, Kruikemeier S, de Vreese CH. In AI we trust? Perceptions about automated decision-making by artificial intelligence. AI & SOCIETY. 2020 Sep;35(3):611-23. Available from: https://doi.org/10.1007/s00146-019-00931-w. doi:10.1007/s00146-019-00931-w.

[83] Cabiddu F, Moi L, Patriotta G, Allen DG. Why do users trust algorithms? A review and conceptualization of initial trust and trust over time. European Management Journal. 2022 Oct;40(5):685-706. Available from: https://www.sciencedirect.com/science/article/pii/S0263237322000846. doi:10.1016/j.emj.2022.06.001.

[84] Aysolmaz B, Müller R, Meacham D. The public perceptions of algorithmic decision-making systems: Results from a large-scale survey. Telematics and Informatics. 2023 Apr;79:101954. Available from: https://www.sciencedirect.com/science/article/pii/S0736585323000187. doi:10.1016/j.tele.2023.101954.

[85] Mahmud H, Islam AKMN, Ahmed SI, Smolander K. What influences algorithmic decision-making? A systematic literature review on algorithm aversion. Technological Forecasting and Social Change. 2022 Feb;175:121390. Available from: https://www.sciencedirect.com/science/article/pii/S0040162521008210. doi:10.1016/j.techfore.2021.121390.

[86] Cadario R, Longoni C, Morewedge CK. Understanding, explaining, and utilizing medical artificial intelligence. Nature Human Behaviour. 2021 Dec;5(12):1636-42. Number: 12 Publisher: Nature Publishing Group. Available from: https://www.nature.com/articles/s41562-021-01146-0. doi:10.1038/s41562-021-01146-0.

[87] Shin D. The effects of explainability and causability on perception, trust, and acceptance: Implications for explainable AI. International Journal of Human-Computer Studies. 2021 Feb;146:102551. Available from: https://www.sciencedirect.com/science/article/pii/S1071581920301531. doi:10.1016/j.ijhcs.2020.102551.

[88] Neyedli HF, Hollands JG, Jamieson GA. Beyond identity: incorporating system reliability information into an automated combat identification system. Human Factors. 2011 Aug;53(4):338-55. doi:10.1177/0018720811413767.

[89] Wang L, Jamieson GA, Hollands JG. The Effects of Design Features on Users' Trust in and Reliance on a Combat Identification System. Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 2011 Sep;55(1):375-9. Publisher: SAGE Publications Inc. Available from: https://journals.sagepub.com/doi/abs/10.1177/1071181311551077. doi:10.1177/1071181311551077.

[90] Hoff KA, Bashir M. Trust in automation: integrating empirical evidence on factors that influence trust. Human Factors. 2015 May;57(3):407-34. doi:10.1177/0018720814547570.

[91] McBride SE, Rogers WA, Fisk AD. Understanding the Effect of Workload on Automation Use for Younger and Older Adults. Human factors. 2011 Dec;53(6):672-86. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5816343/. doi:10.1177/0018720811421909.

[92] Casilli AA, Posada J. The Platformisation of Labor and Society. In: Graham M, Dutton WH, editors. Society and the Internet. vol. 2 ed. Oxford: Oxford University Press; 2019. .

[93] Miceli M, Posada J. The Data Production Dispositif. Proceedings of the ACM on Human-Computer Interaction. 2022;6(CSCW2). doi:10.48550/arXiv.2205.11963.

[94] Schindler S, Wegrzyn M, Steppacher I, Kissler J. Perceived communicative context and emotional content amplify visual word processing in the fusiform gyrus. The Journal of Neuroscience: The Official Journal of the Society for Neuroscience. 2015 Apr;35(15):6010-9. doi:10.1523/JNEUROSCI.3346-14.2015.

[95] Schindler S, Kissler J. People matter: Perceived sender identity modulates cerebral processing of socio-emotional language feedback. NeuroImage. 2016 Jul;134:160-9. doi:10.1016/j.neuroimage.2016.03.052.

[96] Webb T, Holyoak KJ, Lu H. Emergent Analogical Reasoning in Large Language Models. arXiv; 2022. ArXiv:2212.09196 [cs]. Available from: http://arxiv.org/abs/2212.

09196. doi:10.48550/arXiv.2212.09196.

[97] Miceli M, Yang T, Alvarado García A, Posada J, Wang SM, Pohl M, et al. Documenting Data Production Processes: A Participatory Approach for Data Work. Proceedings of the ACM on Human-Computer Interaction. 2022;6(CSCW2). doi:10.1145/3555623.

[98] Simons DJ, Shoda Y, Lindsay DS. Constraints on Generality (COG): A Proposed Addition to All Empirical Papers. Perspectives on Psychological Science. 2017;12(6):1123-8. doi:10.1177/1745691617708630.

[99] Prince SA, Adamo KB, Hamel ME, Hardt J, Gorber S, Tremblay M. A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. International Journal of Behavioral Nutrition and Physical Activity. 2008;5(56). doi:10.1186/1479-5868-5-56.

[100] Nosek BA, Hardwicke TE, Moshontz H, Allard A, Corker KS, Dreber A, et al. Replicability, Robustness, and Reproducibility in Psychological Science. Annual Review of Psychology. 2022;73(1):719-48. _eprint: https://doi.org/10.1146/annurev-psych-020821-114157. Available from: https://doi.org/10.1146/annurev-psych-020821-114157. doi:10.1146/annurev-psych-020821-114157.

[101] Yarkoni T. The generalizability crisis. Behavioral and Brain Sciences. 2022;45:e1. Publisher: Cambridge University Press. Available from: https://www.cambridge.org/core/journals/behavioral-and-brain-sciences/article/abs/generalizability-crisis/AD386115BA539A759ACB3093760F4824. doi:10.1017/S0140525X20001685.