European Journal for Philosophy of Science (2026) 16:26
https://doi.org/10.1007/s13194-026-00732-1

PAPER IN PHILOSOPHY OF TECHNOLOGY AND INFORMATION

Facial analysis AI as social pseudotechnology

Päivi Seppälä[1] · **Ilmari Hirvonen**[1,2]

Received: 28 April 2025 / Accepted: 23 February 2026

Abstract

Facial analysis AI applications have recently come under heavy criticism. They have been accused, for instance, of being discriminative, racist, and even pseudoscientific. In this paper, we assess the latter accusation. We argue that, in problematic cases, facial analysis AI should be considered a form of social pseudotechnology, although it might not be pseudoscientific. We also propose a new two-part definition of pseudotechnology: Either it is not justified that the pseudotechnology could achieve its intended practical purpose better than random chance through its operation principles, or it can achieve its purpose, but it is not justified that this happens through the features claimed to be crucial for its functioning. Stricter definitions of pseudotechnology have difficulty capturing such phenomena as facial analysis AI. When our definition is extended to social technologies, it becomes clear that social pseudotechnologies are more common than previously suggested. Because social pseudotechnologies influence how people act and think, they are not often perceived as pseudotechnologies. This resembles the phenomenon of reactivity, where scientific classifications affect people's self-perceptions and perceptions of others. We support our claims with case studies of facial analysis AI technologies and pseudotechnologies. We engage with the current computer, behavioural, and social scientific literature to assess the extent to which it is justified to use facial analysis AI systems in these cases.

Keywords Demarcation · Facial analysis AI · Pseudoscience · Pseudotechnology · Social pseudotechnology · Reactivity

✉ Päivi Seppälä
paivi.a.seppala@helsinki.fi

[1] Faculty of Social Sciences, University of Helsinki, Helsinki, Finland

[2] The Helsinki Circle, Helsinki, Finland

Published online: 25 March 2026

1 Introduction

Facial analysis artificial intelligence (AI) is a computer vision technology that ascribes socially significant attributes to individuals through automatic facial analysis.[1] It has become an important topic in computer science, AI ethics, and social sciences (Goldenfein, 2019). Facial analysis AI applications are diverse, ranging from inferring emotions, personality traits, and "beauty scores" to determining the gender, sexual orientation, race, and political views of individuals based on their facial features (Monteiro, 2023; Ullstein et al., 2022). Moreover, they have been applied in various societal contexts, including video gaming, hiring, policing, and teaching (McStay, 2020).

The ethical risks posed by some uses of facial analysis AI include privacy violations and the potential for discrimination and oppression. Since facial analysis is developed and used globally by corporations and public institutions, the ethical risks are more than matters of academic debate: they are urgent concerns (Ullstein et al., 2022; van Noorden, 2020). As an example, Chinese authorities in Xinjiang have used facial analysis technology to identify Uyghur faces from surveillance camera footage and, thus, in the oppression of the Uyghur population (van Noorden, 2020).

Besides the ethical, we suggest that facial analysis AI raises epistemic – and even ontological – concerns, such as whether it is possible to classify people into ethnic or racial groups based on their facial appearance. These have also been formulated as concerns that facial analysis AI is pseudoscientific.[2] Facial analysis AI seems to reanimate outdated pseudoscientific practices and theories like those endorsed in phrenology[3] and physiognomy[4] (Ajunwa, 2021; Andrews et al., 2024). In the nineteenth century, physiognomy and phrenology provided the theoretical background and methodology for eugenicists and scientific racists to evaluate character and assign intelligence, morality, and racial hierarchies (Agüera y Arcas et al., 2023; Winston, 2020). Due to this, critics argue that the developers and marketers of facial AI technologies have not only reinvigorated pseudoscience but also "scientific racism at an unprecedented scale" (Stark & Hutson, 2021, 6).

In this paper, we engage with the claim that facial analysis AI is pseudoscientific. We examine applications of facial analysis AI and argue that, in some cases, they should rather be seen as pseudotechnologies. Thus far, philosophers of science have paid regrettably little attention to the demarcation problem from the perspective of pseudotechnology. According to the definition we propose, pseudotechnology is a means used to achieve some intended practical purpose, but either it is not justified

[1] Facial analysis AI should be distinguished from face recognition AI, which identifies individuals by comparing their appearance to an image database (Scheuerman et al., 2021).

[2] For instance, Andrews et al. (2024), Agüera y Arcas et al. (2023), Ajunwa (2021), Crawford (2021), Monteiro (2023), Roemmich et al. (2023), Scheuerman et al. (2021), Sloane et al. (2022), Stark (2019), Stark and Hutson (2021), and Sublime (2024) have argued that facial analysis AI is pseudoscientific.

[3] Phrenology was a 19th-century pseudotheory whose most famous proponents included Franz Joseph Gall and Johann Spurzheim. Gall, Spurzheim, and their followers sought to link skull morphology to personality traits and human behaviours such as criminal behaviour (Jones et al., 2018).

[4] Physiognomy or face reading is the more general assessment of a person's character or personality based on any outer feature of the face – and not merely by the shape of the skull (Lichtenberg, 1778).

that it could achieve its purpose better than random chance through its operation principles, or if it can do so, it is not justified that this happens through the features claimed to be crucial for its functioning. Stricter definitions of pseudotechnology – such as the one offered by Sven Ove Hansson (2020) – have difficulty capturing phenomena like facial analysis AI. Furthermore, we introduce the concept of *social pseudotechnology* and argue that because pseudotechnologies are often social, they are more common than previously suggested in the literature.

We also caution against automatically labelling all facial analysis AI as "physiognomic pseudoscience". An alleged technology does not have to be based on pseudoscience to be pseudotechnology (Hansson, 2020). Therefore, the concepts of pseudotechnology and pseudoscience are independent of each other. We argue that for someone concerned about facial analysis AI and certain forms of its use, the most crucial question is not whether it is pseudoscientific, but whether there is sufficient evidence that it can achieve its intended practical purpose in the way it is claimed to function. In addition, more attention should be paid to the fact that some facial analysis AI systems do achieve their intended purposes. However, in cases of social pseudotechnology, this has not been proven to happen through the features claimed to be crucial for their functioning.

We also demonstrate how the concept of social pseudotechnology can assist computer scientists, critical data scholars, policymakers, and the general public. For instance, it can be used to explain why certain facial analysis AI applications are epistemically problematic and to clarify how these epistemic issues relate to ethical and social concerns arising from different uses of facial analysis AI. Furthermore, we discuss why social pseudotechnologies are difficult to recognise as pseudotechnologies.

2 Facial analysis AI as physiognomic and phrenological pseudoscience

As mentioned in the introduction, some authors have argued that facial analysis AI is pseudoscientific, often pointing out its similarity to physiognomy and phrenology. Both facial analysis AI and physiognomy share the idea that analysing a person's skull or facial appearance can reliably determine their character, personality, or race and predict their behaviour (Agüera y Arcas et al., 2023; Ajunwa, 2021; Andrews et al., 2024; Stark & Hutson, 2021; Sublime, 2024). In addition, similar to physiognomists such as Cesare Lombroso, some facial analysis AI developers and researchers (e.g., Hashemi & Hall, 2020a; Wu & Zhang, 2016a) aim to predict socially deviant behaviour, such as criminality, from facial features (Sublime, 2024).

Phrenologists and facial analysis AI researchers also make similar statistical inferences (Andrews et al., 2024; Sublime, 2024). Phrenologists, such as Sir Francis Galton, developed the foundational methods of modern statistics to generalise from subpopulations to entire populations, enabling the application of phrenology-based racial categorisations (Cowan, 1972; Winston, 2020). Similarly, facial analysis AI researchers, such as Fu et al. (2014), claim that their AI models can generalise facial appearance-based racial categorisations derived from training data to new facial data.

The assumption that racial categories can be statistically inferred from facial appearances is epistemologically deeply problematic (see Section 4.4). Furthermore, the racial categorisations by phrenologists were used to support morally reprehensible, oppressive practices such as eugenics, racial hierarchies, and colonialism (Cowan, 1972; Stark & Hutson, 2021; Winston, 2020). Currently, the racial categorisations generated by facial analysis AI can be used to maintain and even strengthen existing patterns of discrimination between societal and racial groups (Benjamin, 2019). The oppression of the Uyghur population (van Noorden, 2020) is an ethically alarming example of such use.

For these reasons, previous critics have argued that facial analysis AI is pseudoscientific (e.g., Andrews et al. (2024); Agüera y Arcas et al. (2023); Ajunwa (2021). Moreover, they have concluded – quite rightly, in our view – that some applications of facial analysis AI are very likely societally harmful and discriminatory (Ajunwa, 2021; Benjamin, 2019; Stark & Hutson, 2021). However, we will argue that, in some cases, facial analysis AI should be examined as pseudotechnology rather than pseudoscience. Before that, we must define what pseudotechnology is.

3 Pseudotechnology: its nature and variants

In this section, we introduce a novel definition of pseudotechnology and explore some of its varieties. We pay special attention to social pseudotechnology. When facial analysis AI counts as pseudotechnology, it is a social pseudotechnology if it is applied in social contexts to influence social relations.

3.1 Definition of pseudotechnology

Previous authors who have argued that facial analysis AI is pseudoscientific have not attempted to define pseudoscience beyond presenting physiognomy and phrenology as clear instances of it (e.g., Andrews et al. (2024); Agüera y Arcas et al. (2023); Ajunwa, 2021; Stark & Hutson, 2021). This approach is reasonable, as there is no universally accepted definition of pseudoscience. For the same reason, unequivocally defining pseudoscience is beyond the scope of this article. However, we aim to further develop the analysis of facial AI as epistemically problematic. To this end, we introduce the concept of social pseudotechnology. We argue that, in some cases, the concept of pseudotechnology, rather than pseudoscience, is a more accurate tool for analysing the epistemic problems of facial analysis AI.

Some distinction between science and technology, and pseudoscience and pseudotechnology, is needed here. Roughly speaking, in our terminology, science and pseudoscience focus on theories, whereas technology and pseudotechnology concentrate on practice. In other words, pseudoscience seeks to describe and sometimes denies scientific descriptions (e.g., Hansson, 2017). Pseudotechnology, in contrast, is used to solve practical problems.

There are at least two reasons for this distinction between theory and technology. First, technology is not always based on academic or institutionalised science. For instance, farmers have developed agricultural technologies through trial and error, often without starting from any theory (Hansson, 2015; 2019). Similarly, some significant technologies, like the steam engine, were not developed from pre-existing theories. Instead, much of twentieth-century science emerged to explain why such technologies work so well (Pfaffenberger, 1988). Likewise, pseudotechnologies do not require preceding pseudotheories. In a nutshell, technology is not merely applied science nor pseudotechnology applied pseudoscience (Hansson, 2020; Pfaffenberger, 1988).

Second, pseudotechnologies can be based on academic or institutionalised science. Pseudotechnologies may be founded on scientific theories, but as applications, they fail to function as intended. This will be illustrated with examples and case studies, especially in Section 4. Despite our distinction between theory and technology, the two are closely connected. For instance, scientific methods for testing theories, along with other research practices, are technologies according to our definition.[5]

We define pseudotechnology as follows:

Pseudotechnology: a means used to achieve some intended practical purpose, but either

(a) it is not justified that it could do so better than random chance through its operation principles or
(b) if it can achieve its intended purpose better than random chance, it is not justified that this happens through the features claimed to be crucial for its functioning.

Pseudotechnologies can include physical artefacts, conceptual systems, practices, or, as in our case studies, AI systems. The important thing is that they are applied to achieve some practical goals. This definition aligns with, for instance, Harvey Brooks' definition of technology as "knowledge of how to fulfil certain human purposes in a specific and reproducible way" (Brooks, 1980, 66). Along similar lines, Donald A. Schon has defined technology as "any tool or technique, any physical equipment or method of doing or making, by which human capability is extended" (Schon, 1967, 1). In the previous literature on pseudotechnology, Mario Bunge (1976) and Raimo Tuomela (1987) have presented similar broad conceptions of technology and pseudotechnology. It is important to note that two (pseudo)technologies can be physically and conceptually identical and still count as separate technologies or pseudotechnologies because they have different uses. This means that technology or pseudotechnology must be defined with respect to its intended purpose or effect (Hansson, 2020).

In our definition, justification refers to reasons or arguments, especially generally accepted justificatory methods and practices, like those used to produce scientifically legitimate knowledge. One crucial question is determining when something is justified or what threshold is necessary for justification. Unfortunately, we do not believe such a universal threshold exists.[6] One reason for this is the argument from

[5] Likewise, scientific models can be seen as technologies when they are not simply used as descriptions (Morrison & Morgan, 1999, p. 35; Guest & Martin, 2021).

[6] However, there seems to be some minimal criteria for justification. For example, the premises of an argument must be relevant to the conclusion (see, e.g., Walton, 2003, 77–80).

inductive risk (Hempel, 1965). Recently, for instance, Heather Douglas (2000; 2009) has defended the argument, but already C. West Churchman (1948) and Richard Rudner (1953) reasoned that values influence what is considered sufficient evidence for something. Since empirical science is not a deductive endeavour, one must always decide on the level of evidence or probability needed for acceptance. Values influence such decisions, and in democratic societies, political discussion is needed to reach them.

For instance, the EU AI Act, which also regulates the use of facial analysis AI, is a political compromise achieved after many years of negotiation among politicians, AI companies, researchers, AI experts, interest groups, and AI-critical activists (McStay & Bakir, 2025). It sets higher epistemic standards for AI applications used for "high-risk" purposes because using AI for purposes such as hiring, law enforcement, or public decision-making may affect the enactment of individuals' fundamental rights (EU Regulation 2024/1689, 2024 article 6, Annex III; articles 8–15). This demonstrates how moral and social values can – and do – influence the evidential bar for facial analysis AI systems.

The functionality of a pseudotechnology could be unjustified for various reasons. For instance, as Sven Ove Hansson (2020) has proposed, a pseudotechnology might be based on fundamentally flawed construction principles, and there are good reasons to believe this is the case.[7] The fact that pseudotechnology might be based on fundamentally flawed construction principles – or *operation principles* as we call them – distinguishes it from nonfunctional technology.[8] A broken bicycle is a nonfunctional technology, not a pseudotechnology. (Hansson, 2020.) For the same reason, badly functioning technology is not pseudotechnology. If a bicycle's parts are bent, it can be challenging to ride, and one can easily fall over. However, a poorly functioning bicycle is still a bicycle, not a pseudobicycle. Moreover, it is possible to build a better bicycle using the same operating principles.

Alternatively, pseudotechnology might rely on scientifically unwarranted operation principles, thereby rendering its functionality unjustified (see Section 4.1). However, the concepts of pseudoscience or bad science are inadequate to describe or explain all instances of pseudotechnology. One reason for this is that the concept of pseudotechnology is independent of the concept of pseudoscience. Sometimes, pseudotechnology is based on ongoing research, and its functionality remains an open scientific question. However, it is already used, and its functionality is claimed to be scientifically justified even though this is not the case (see Section 4.2). The concept of pseudotechnology is needed to describe such unjustified applications, which cannot yet be identified as pseudoscientific or scientific. Pseudotechnology can also be based on theories previously accepted by the scientific community, but most researchers have now questioned them due to new research results (see Section 4.3).

[7] Hansson's examples of pseudotechnologies include, for instance, dowsing and cold fusion (Hansson, 2020).

[8] We use the term operation principles to refer to the sets of causal and theoretical assumptions that the developers of AI models presume to explain why the model produces the predictions as it does from the data it uses as input. The developers might explicitly express these assumptions to justify their modelling choices or presume them tacitly (Friedler et al., 2021).

It is also worth noting that a potential technology is not a pseudotechnology when its functionality is still under research. However, a potential technology can become a pseudotechnology if it is implemented as if it were functional, even though there is insufficient evidence of its functionality.

According to our definition some pseudotechnologies can fulfil their intended purposes but not in the way they are claimed to function. For instance, compare homeopathic and paracetamol tablets. Homeopathic pills can have a placebo effect (Brien et al., 2011; Ernst, 2002). The contents of a placebo pill do not, as such, bring about the effect. Instead, the effect is caused by the patients' beliefs, which are influenced by what they have been told about the tablets and the observable features of the pills – their colour, size, shape, texture, taste, and so on (see, e.g., Howick & Hoffmann, 2018; Zunhammer et al., 2022). What is claimed to be crucial for homeopathic remedies has no causal relevance. The homeopathic diluted substances could be removed from the pill or altered radically without any difference in the outcome. In other words, the original diluted drug could be replaced by something completely different, or the final remedy could merely be, for example, plain water, alcohol, or sugar tablets without anything having been diluted in the first place.[9] Paracetamol, on the other hand, cannot be removed from a painkiller or radically altered without the pill reducing its ability to work as intended.

It is important to distinguish pseudotechnologies from properly functioning technologies with pseudoscientific explanations of their functionality. The justification for whether a claimed technology works may differ from the theory or explanation of why it works. One can test in practice whether a proposed technology works and confirm that it does, then develop a pseudoscientific theory or explanation for it. After all, functionality differs from the explanation of functionality. One can have a pseudoscientific theory of why a painkiller works the way it does. For example, one could insist that the relative position of celestial bodies enables it to function. This would be a pseudoscientific attempt to explain the painkiller's functionality, but it would not turn the painkiller into a pseudotechnology. The reason is that there is still sufficient evidence from clinical trials to think that some actual feature of the pill, namely, its content, is causally relevant.

Suppose someone claims that for paracetamol painkillers to work, it is necessary to draw an astrological symbol on the pills or surround them with crystals for an hour. Drawing a symbol on paracetamol tablets or circling them with crystals would count as pseudotechnologies under our definition because they are causally irrelevant to the painkiller's functionality. However, even at this stage, the tablets themselves will not become pseudotechnologies if they still contain paracetamol.

Hansson (2020) has discussed the concept of pseudotechnology most thoroughly in the previous literature. Our definition builds on Hansson's characterisation of pseudotechnology, which we have refined and applied to cases not previously recognised and discussed as pseudotechnologies. Hansson defines pseudotechnology

[9] A placebo can also be a working technology when deliberately used as a placebo. However, suppose some medicine is assumed to work because of its ingredients and not just due to the placebo effect, even though its ingredients are causally inefficacious. In that case, the medicine in question is a pseudotechnology.

as “an alleged technology that is irreparably dysfunctional for its intended purpose since it is based on construction principles that cannot be made to work” (Hansson, 2020, 692). In other words, according to Hansson (2020, 691), “non-functionality is a necessary criterion for pseudotechnology.” Hansson believes this is why pseudotechnologies are so rare: they typically reveal their flaws immediately when used. As we shall later argue, this is not the case with technologies applied to thinking beings such as humans or other animals.

For Hansson, technology refers to the “tools, machines, and procedures used in industry, rather than to knowledge about these tools, machines, and procedures” (Hansson, 2020, 689). This means that Hansson limits technology and pseudotechnology to physical artefacts, but in our view, this is unnecessary. Hansson’s definitions of technology and pseudotechnology are special cases of our broader definitions. Suppose we are justified in believing that the construction principles of pseudotechnology cannot be made to work as Hansson’s definition of pseudotechnology presupposes. In that case, the pseudotechnology’s functionality is unwarranted, as the condition (a) of our definition states. However, we show in the following sections that there are clear cases of pseudotechnologies which do not fit Hansson’s characterisation.

3.2 Social technology and social pseudotechnology

Pseudotechnologies that target nature often reveal their flaws immediately, as Hansson (2020) notes. Other pseudotechnologies target thinking beings, like humans or animals, and the shortcomings of these pseudotechnologies may not always be directly visible.[10] Hansson also points out that some pseudotechnologies focus on improving the physical well-being of humans and sometimes seem to achieve their goals.[11] Besides medical pseudotechnologies, we argue that there is another significant type of human-applied pseudotechnology that has thus far received regrettably little attention in philosophy. We call them social pseudotechnologies. Our central thesis is that facial analysis AI applications often are social pseudotechnologies.

To understand what social pseudotechnology is, it is good to begin with its counterpart, social technology. Currently, the term “social technology” is mainly used for social media and other social applications (e.g., Treem et al., 2024). However, the concept has a relatively long and varied history and was used in a much broader sense before the invention and widespread use of computers.[12] The conception of social technology presented here is more in line with the previous, broader use of the term. In our usage, social technology is defined as follows:

[10] A third category could be added: pseudotechnologies used in knowledge production. This category would include pseudoscientific methodologies and research practices. These are exceptional because pseudotechnologies are usually intended to have a more direct impact on the world. Nevertheless, they still count as pseudotechnologies because they are used to achieve some purpose.

[11] Hansson (2020, 693) argues that pseudotechnologies, such as energy-balancing bracelets, seem in most cases successful because their supposed effects are so “ill-defined” and “irregular” that they cannot be immediately falsified.

[12] For previous broader conceptions of social technology, see, for example, Small (1898, 131–133) and Derksen and Beaulieu (2011).

Social Technology: a means intentionally used to sustain, organise, or modify social relations.

Social technologies and social pseudotechnologies are intrinsically linked to social processes and relations. They are intentionally used to facilitate, control, transform, and create interactions and social classifications. They can modify what people think about other people and how they interact. In other words, social technologies and pseudotechnologies shape social reality. Someone might object that this is true of all technologies. As Bryan Pfaffenberger (1988) has argued, every technology also influences human behaviour and social relations. This is undoubtedly correct. However, social technologies and pseudotechnologies are *intentionally* used to achieve socially desirable goals. Other technologies and pseudotechnologies produce unintended social consequences that can be significant and socially harmful. For instance, car safety tests predominantly use male crash-test dummies, increasing the risk of injury for female drivers (Criado Perez, 2019). Still, in our analysis, whether something counts as a social technology depends on its intended use.

Our definition of social pseudotechnology is formed simply by combining our conceptions of pseudotechnology and social technology:

Social Pseudotechnology: a means intentionally used to sustain, organise, or modify social relations, but either

(a) it is not justified that it could achieve its purpose better than random chance through its operation principles or
(b) if it can achieve its intended purpose better than random chance, it is not justified that this happens through the features claimed to be crucial for its functioning.[13]

It is important to understand how pseudoscience, pseudotechnology, and social pseudotechnology differ and how they relate. For instance, a classification system is not yet a pseudotechnology, although it can be pseudoscientific. For example, classifications of different human races based on facial features are pseudoscientific because human races do not genetically exist (see Section 4.4). The genetic variation between human populations does not justify dividing our species into subspecies, biological races, or any biologically meaningful subgroups. The differences between populations are very small, and most human genetic variation is between individuals. This excludes the possibility of genetic human races. (Templeton, 2013.) When an unscientific classification system is intentionally applied to achieve some goal, it becomes a pseudotechnology. Moreover, if it is intentionally used to sustain, organise, or modify social relations, it is a social pseudotechnology.

[13] Mahner has introduced two concepts that are somewhat similar to our concept of social pseudotechnology, namely, "psychological pseudotechnology" and "pseudo-sociotechnology" (Mahner, 2007, 550). In Mahner's parlance, psychological pseudotechnology is, roughly, applied pseudoscientific psychology and pseudo-sociotechnology is applied pseudoscientific sociology. Our concept of social pseudotechnology covers both psychological pseudotechnology and pseudo-sociotechnology, since both are connected to influencing social processes.

Depending on its application, a functioning technology may become a social pseudotechnology. For example, in 2019, some Chinese primary schools experimented with using headbands to track pupils' concentration in classrooms (Tai et al., 2019). The measured information was directly sent to teachers and parents. The headbands use electroencephalography (EEG) to measure brain waves, and there may be a connection between concentration and the electrical activity recorded by an EEG. However, what is measured cannot be straightforwardly equated with concentration. There is not much research on the issue, and multiple sources of error exist. (Tai et al., 2019; van Atteveldt et al., 2020). For instance, students might be focused on something unrelated to the lesson. They could be reflecting on, say, a comic book they read the night before. Additionally, there can be individual differences in brain activity among students who are concentrating, due to significant variations in baseline electrical activity (Miller et al., 2002).

Therefore, using the headband to measure "concentration" makes the technology a pseudotechnology, even though it might be able to measure brain waves with some accuracy. Despite being a pseudotechnology, the headbands made students more disciplined, attentive, and diligent in their studies. The technology was intentionally used to sustain, organise, or modify social relations in a manner insufficiently backed up by scientific research. The headbands achieved their intended purpose of improving concentration. However, the headbands were also a social pseudotechnology because this improvement was not based on genuine measurement of concentration. More likely, the students' concentration improved due to teachers' and parents' reactions to the data. For example, some parents punished their children when the headband data suggested they were not paying attention in class (Tai et al., 2019).

As Hansson (2020) points out, pseudotechnologies might seem to work when applied to human beings because it is unclear what the pseudotechnology is precisely intended to do and whether its goal has been achieved. In addition, we argue that sometimes pseudotechnologies – especially social pseudotechnologies – also function in cases where the purpose of the pseudotechnology is clearly defined. This is possible, because social pseudotechnologies influence how people act and think. This resembles the phenomenon of reactivity, where scientific classifications influence self-classifications (Marchionni et al., 2024).[14] For example, facial analysis AI technologies can affect how people classify themselves or others, which, in turn, might validate these technologies' predictions (see Section 4.1). In other words, social pseudotechnologies that use such categorisations may function because instead of merely describing it, they create social reality. This applies to the headbands used in Chinese schools. They increase the pupils' concentration, not because the headbands accurately measure it, but because the pupils most likely react to the surveillance of their concentration.

[14] Reactivity has been given other names, including self-fulfilling prophecies, the looping effects of human kinds, and performativity (Marchionni et al., 2024).

4 Facial analysis AI: science and technology or pseudoscience and pseudotechnology?

In this section, we examine the applications of facial analysis AI to determine whether they should be classified as (social) technologies or (social) pseudotechnologies. We also demonstrate that everything that falls under the umbrella of facial analysis AI should not be condemned as pseudoscience as quickly as some critics have done. Instead, we argue that evaluating facial analysis AI requires three steps. First, one must identify the exact purpose of the (pseudo)technology. Second, one needs to identify the (pseudo)technology's operation principles and features, which are claimed to be crucial for achieving its intended purpose. Third, one must assess whether it is justified that the (pseudo)technology can be used for its intended purpose and by what means**.** Only after this can one conclude whether this is an instance of technology or pseudotechnology.

We determine the (pseudo)technology's intended use from research articles by developers of facial analysis AI, as well as marketing and other materials published by commercial AI service providers. For the two subsequent steps, we engage with the current research literature to identify the (pseudo)technology's operation principles and assess how justified these principles are for achieving the intended purpose.[15] Our approach is supported by previous research. For instance, Egan (2017), Guest (2024), and Sloane et al. (2022) have suggested similar methodologies to determine and evaluate the underlying assumptions of AI-based decision-making systems.

The facial AI systems we analyse are intended to enhance, replace, or displace human reasoning and decision-making to some degree (Guest, 2025).[16] Furthermore, the operation principles of the pseudotechnologies in all our case studies involve causal assumptions about the factors influencing human behaviour in specific contexts. Therefore, assessing whether these facial analysis AI systems are social pseudotechnologies requires considering whether the best current social and behavioural scientific and anthropological evidence supports such use.

Moreover, our cases involve the use of facial analysis AI in ways that could negatively affect safety, fundamental rights, and the allocation of financial and societal resources. We focus mainly on the epistemic evaluation of facial analysis systems used or potentially used in ethically high-risk contexts. In fact, using facial analysis to predict criminal behaviour (Section 4.1) and to infer race (Section 4.4) can be deemed morally unacceptable. Therefore, these uses are prohibited under the EU AI Act (EU Regulation 2024/1689, 2024 article 5).

[15] The cases in this section do not provide a comprehensive list of all current dubious facial analysis applications identified by critics. For instance, facial analysis AI has been used to predict sexual orientation (Wang & Kosinski, 2018), and this has been criticised by Agüera y Arcas et al. (2023).

[16] The assumption that AI systems can enhance, replace, or displace human decision-making is often motivated by an ontological assumption shared in cognitive science, computer science, and AI development. The assumption is that human reasoning is parallel to information processing systems such as computers and AI systems (Boden, 2006; Crawford, 2021; van Rooij et al., 2024). However, if an information processing system, such as a facial AI system, is good at a predictive task, like making hiring decisions, it does not logically follow that the system is ontologically akin to human reasoning in this task (Guest & Martin, 2023). Such reasoning is an example of the logical fallacy affirming the consequent (Guest & Martin, 2023).

One does not need to agree with the specific risk categorisations outlined in the EU AI Act. However, if ethical distinctions should be made among the various uses of facial analysis systems, this should influence the epistemic standards applied to them. The higher the risk that a facial analysis AI system violates fundamental rights, causes discrimination, jeopardises safety, or leads to an unfair allocation of resources, the more critical it becomes to provide evidence that the system is not a social pseudotechnology. This aligns with the argument from inductive risk introduced in Section 3.2.

Ethical considerations should guide which uses of facial analysis systems require more thorough epistemic assessment. Systems that pose the most significant risk of harm should be prioritised. This is why we focus on very high-risk facial analysis systems. Given that they are already in use, their epistemic evaluation is urgent and imperative.

Epistemological considerations can influence the evaluation of the moral wrongness of facial analysis AI systems. It is unethical if such systems are used in predicting criminal behaviour or hiring so that socially significant groups are discriminated against. Furthermore, if the system not only discriminates but is also a pseudotechnology, the situation is even worse. In such cases, the system discriminates based on falsehoods. Epistemic considerations regarding pseudotechnologies should inform decision-making for those who value truth, for instance, in regulating facial analysis AI.[17] Therefore, our epistemological analysis of the various societally significant uses of facial analysis AI contributes to the ethical and political discourse surrounding it.

4.1 Predicting "Criminal Behaviour"

In this subsection, we discuss three examples of facial analysis AI systems. These systems are used to recognise and predict criminal behaviour. Two examples are from computer science, and the third is a commercial application.

Our first example is Xiaolin Wu and Xi Zhang's article "Automated Inference on Criminality using Face Images" (2016a). Wu and Zhang announced that their convolutional neural network (CNN) classifier achieves 89.51% accuracy in predicting criminal behaviour by analysing standard ID photographs controlled for race, gender, age, and facial expression (Wu & Zhang, 2016a). Wu and Zhang (2016a, 1) motivated their research by appealing to the social need to identify criminals to protect "law-abiding citizens". Their theoretical motivation was that "[i]n all cultures and all periods of recorded human history, people share the belief that the face alone suffices to reveal innate traits of a person".

Critics accused Wu and Zhang of propagating physiognomic pseudoscience (e.g., Whittaker et al., 2018). Blaise Agüera y Arcas et al. (2017; 2023) analysed Wu and Zhang's study in detail and argued that, rather than a stable phenomenon, their result

[17] For example, using facial analysis AI to determine one's beauty score or horoscope to predict one's future can be regarded as harmless "entertainment" and therefore be labelled as "low-risk" (Monteiro, 2023). Still, using facial analysis AI to infer one's horoscope and predict the future is undoubtedly a social pseudotechnology. Although this use might not be harmful, some might argue that it should nevertheless be banned. One might value truth in itself and, on this basis, conclude that social pseudotechnologies should not be used even in low-risk contexts.

is probably an artefact. The critics also pointed out that even if Wu and Zhang's results were reliable, their physiognomic explanations are wrong (Agüera y Arcas et al., 2023; Coalition for Critical Technology, 2020; Pasquale, 2018; Stinson, 2020).

The critics argued that Wu and Zhang mistakenly presume two things: (1) criminal sentences reliably track "criminal behaviour" and (2) facial appearance does not influence criminal sentences. These two presumptions are presented in Fig. 1, which describes the causal assumptions behind Wu and Zhang's (2016a) deep neural network technology and depicts its operation principles.

Agüera y Arcas et al. (2017) present an alternative social explanation to Wu and Zhang's physiognomic explanations that can equally well account for the correlation between facial features and criminal sentences. Criminal judgement is affected by a judge's impressions of facial appearance: people who are perceived as "criminal-looking" or "untrustworthy" are more likely to be convicted of crimes (Macrae & Shepherd, 1989; Zebrowitz & McDonald, 1991). Facial analysis AI systems pick up the same cues that impact human perception, which is why AI systems identify the same people that human judges identify as "criminals". Instead of identifying potential criminal offenders, facial analysis AI reveals the extent of bias in criminal sentencing. (Agüera y Arcas et al., 2017.) If this is correct, the causal pattern explaining the correlation identified by the deep neural network classifier is more complicated (Fig. 2) than Wu and Zhang presumed.

All in all, it seems that Wu and Zhang failed to consider that discovering a statistical correlation between facial photographs and criminal sentences does not provide sufficient evidence to conclude that the same underlying genetic factors affect both facial appearance and criminal behaviour (Andrews et al., 2024). A mere statistical correlation cannot confirm such a genetic hypothesis (Todorov et al., 2015). Instead,

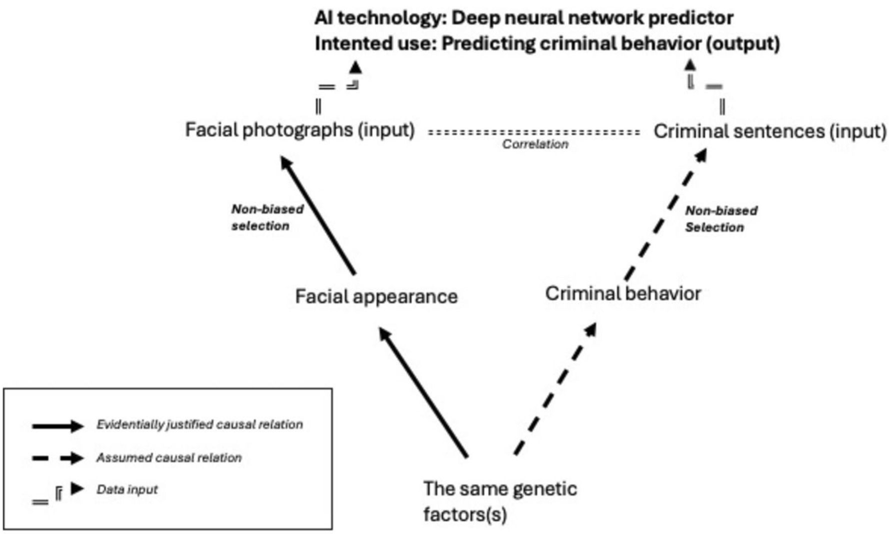

Fig. 1 Causal hypothesis for explaining the predictive correlation between facial photographs and criminal sentence. This picture illustrates Wu and Zhang's (2016a) causal reasoning

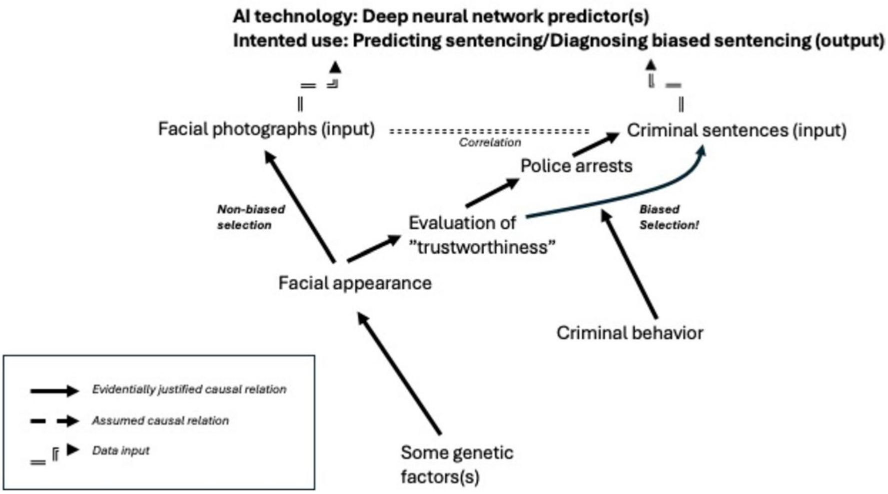

Fig. 2 Causal hypothesis for explaining the predictive correlation between facial photographs and criminal sentences. This picture illustrates the causal reasoning in Agüera y Arcas et al. (2017)

as Andrews et al. (2024, 8) argue, "only mechanistic biological evidence would avoid most of the myriad confounding factors possible in such an inference".

Our second example is Mahdi Hashemi and Margeret Hall's article "Criminal tendency detection from facial images and the gender bias effect" (2020a). They claimed to have developed a neural network model to find correlations between facial features and "criminality". They used criminal sentences as a criterion of "being a criminal". The paper was severely criticised by 2425 scholars working in anthropology, sociology, computer science, law, science and technology studies, information science, and mathematics (Coalition for Critical Technology, 2020; van Noorden, 2020). Springer, the article's publisher, was asked to condemn using criminal justice statistics to predict criminality (Coalition for Critical Technology, 2020). Springer did not do that (Birhane & Guest, 2021). Later, however, Hashemi and Hall (2020b) were forced to retract the paper because it violated research ethics. They did not seek approval from an ethics committee before undertaking the study using human biometric data.[18]

Given the fierce academic critique, it seems clear that Wu and Zhang (2016a) as well as Hashemi and Hall (2020a) conducted, if not pseudoscience, at least bad science. They did not consider possible sources of bias – such as biased sentencing – when assuming in their training data that criminal behaviour is directly reflected in criminal sentences. They also neglected existing criminological research and viable criminological alternative explanations that could account for the CNN classifier's

[18] Traditionally, computer science has not required approval from ethical boards. Most facial analysis AI research is based on visual material extracted from the internet without the consent of the people depicted in it. (Crawford 2021, ch. 3.)

results when it was used to predict criminal behaviour. Furthermore, they have not responded to the criticism with the required seriousness.[19]

What, then, is the verdict on whether these cases of facial AI research identifying criminal behaviour count as pseudotechnologies? Wu and Zhang (2016a) and Hashemi and Hall (2020a) claim that their facial analysis AI technologies can identify criminal behaviour, but this claim is not justified. They may be able to predict convictions, but as noted previously, being convicted cannot be equated with a tendency to commit crimes. Therefore, the first condition of our definition of pseudotechnology is fulfilled: the technology is claimed to achieve some intended purpose, but it is not justified that it could achieve it through its operation principles. Hence, the facial analysis AI technologies developed by Wu and Zhang (2016a) and Hashemi and Hall (2020a) are pseudotechnologies.

How about the criteria of social pseudotechnology? To our knowledge, the technologies developed by Wu and Zhang (2016a, b) and Hashemi and Hall (2020a) have not been applied in real high-risk social contexts – like policing or border control – to identify criminal behaviour. Therefore, even though Wu and Zhang's and Hashemi and Hall's classification systems fulfil our first criterion of pseudotechnology, they do not satisfy the criteria for social pseudotechnology.

Our third and final example is the startup company Faception, which promotes a facial analysis technology that can allegedly identify criminals such as "terrorists" and "white-collar offenders" from facial images or video streams (Faception, 2024a). The company claims they can "detect suspected individuals in public places such as airports, train stations, government and public buildings, and border controls" (Faception, 2024a). Furthermore, Faception insists that its technology is based on two scientific findings:

(1) "According to Social and Life Science research personalities are affected by genes": Faception supports this claim with a reference to one twin study (Faception, 2024b).
(2) "Our face is a reflection of our DNA" (Faception, 2024b): a claim supported by studies on mice, which have shown that genes affect how a face is formed. However, Faception provides no references to these mice studies, and furthermore, it extrapolates these results to humans. As a result, they declare that "it's already possible to make some inferences about the appearance of crime suspects from their DNA alone" and add that "in Chinese history, there have been people that have studied the 'mapping of the face' for thousands of years" (Faception, 2024b).

Faception's facial analysis AI is an unmistakable instance of social pseudotechnology. The company declares that its technology can assist in crime prevention by labelling individuals as terrorists. Given that Faception's claims that their technology

[19] Wu and Zhang (2016b) only responded: "Our work is only intended for pure academic discussions; how it has become a media consumption is a total surprise to us. Although in agreement with our critics on the need and importance of policing AI research for the general good of the society, we are deeply baffled by the ways some of them misrepresented our work, in particular the motive and objective of our research."

is scientific are practically non-existent and highly implausible, their AI application meets our first criterion for social pseudotechnology: it is a facial analysis AI system that is used to sustain, organise, or modify social relations, but it is not justified that it could achieve its purpose better than random chance through its operation principles.

In Section 3, "measuring" Chinese schoolchildren's concentration was used to illustrate how social pseudotechnology can fulfil its purpose, not through its purported operating principles or allegedly crucial features but via other mechanisms. Similar reactivity phenomena (Marchionni et al., 2024) are also possible in crime prevention. For example, using AI facial analysis to classify individuals as criminals may lead those labelled as potential criminals to engage in criminal behaviour.

This reactivity hypothesis stems from criminological labelling theory. According to labelling theorists (Becker, 1963; Lemert, 1967) and empirical research testing this theory (Barrick, 2014; Bernburg, 2009; Lee, 2024), criminal behaviour can be sparked, among other potential causes, by labelling individuals as criminals. According to the theory, two mechanisms affect individuals who are initially (correctly or incorrectly) labelled as criminals, increasing their likelihood of engaging in future criminal behaviour (Barrick, 2014).

The first mechanism is that the label leads to social exclusion (Becker, 1963; Bernburg, 2009). The stigma associated with the label "criminal" causes "non-deviant" people to avoid and distrust those labelled as criminals, which might, for instance, reduce employment opportunities (Bernburg, 2009). This, in turn, increases the likelihood that the labelled persons seek social acceptance from groups where criminality is not stigmatised. These are often delinquent groups (Kaplan et al., 1991). Involvement in delinquent groups increases the likelihood of future criminal behaviour (Bernburg et al., 2006).

The second mechanism is based on identity change and deviant self-concept (Barrick, 2014; Bernburg, 2009). People who are labelled as criminals may internalise the label, making it central to their identity. This internalisation raises the probability of future criminal behaviour (Lee, 2024; Lemert, 1967; Matsueda, 1992).

The facial AI system that labels individuals as criminals would, according to our second criterion, be a social pseudotechnology because the system's correct predictions can be explained by reactivity, as labelling theory suggests. The facial analysis AI system could be used to organise social relations by predicting criminal behaviour. However, it would not be justified that this happened through the features claimed to be crucial for its functioning. The system's predictions might come true because being labelled as a criminal based on facial features can influence how individuals act. In other words, the accuracy of the original prediction would not be related to the alleged genetic link between facial appearance and criminal behaviour.

4.2 "Measuring" personality

AI company ZappyHire claims to have developed "AI-enabled video interview platforms" that can be used to "Improve […] Hiring Accuracy by 72%" via analysing candidates' "personal traits" (Roemmich et al., 2023, 9). In other words, the "platform" is used in a high-risk context to supposedly measure personality and predict job performance to help employers make recruitment decisions. Dozens of other AI companies, such as MyInterview and Retorio, promise to do the same by measuring job candidates' Big Five personalities (Roemmich et al., 2023).

Big Five is currently the best-validated psychometric framework for personality measurement, and it measures the following five factors: extroversion, agreeableness, openness, conscientiousness, and neuroticism. In a recent meta-analysis, Big Five characteristics had the following associations with job performance: openness (0.05), agreeableness (0.10), intermediate effects for extraversion (0.13), neuroticism (−0.11), and conscientiousness (0.23). Therefore, the Big Five is generally considered a valid hiring instrument within organisational psychology in terms of construct validity and predictive validity (Zell & Lesick, 2022).

How justified are the claims to measure Big Five personalities made by these vendors of AI-based video interview platforms? Can the technologies be used to measure personality? Furthermore, can the measurements predict job performance reliably to help employers make recruitment decisions?

It should be noted that the measurement goals of these video interview platforms are similar to traditional psychometric methods used to assess personality (Seppälä & Małecka, 2024). Therefore, we should assess the validity of facial analysis AI and traditional personality tests using the same psychometric criteria (Jacobs & Wallach, 2021; Rhea et al., 2022). AI providers, such as ZappyHire, should be able to offer evidence to answer the following three questions on validity (Rhea et al., 2022) that traditional personality tests also need to fulfil:

(1) **Construct Validity:** Are the measured personality traits meaningful psychological constructs?
(2) **Measurement Validity:** Is the test a valid measurement instrument so that it reliably (i.e. consistently) measures the traits it purports to measure?
(3) **Predictive Validity:** Is the test a valid hiring instrument so that its results are predictive of employee performance?

In terms of construct and predictive validity, the AI-based personnel selection software companies that advertise to measure candidates' Big Five personalities do well. This is because the personality traits of the Big Five theory are meaningful psychological constructs, and the theory can predict employee performance, at least to some extent.

Big Five personality has traditionally been measured with self-assessment questionnaires, which have been validated as reliable measuring instruments for this task (Azucar et al., 2018). Hence, measurement validity is the most significant validity problem that AI firms selling facial analysis of Big Five personality have. The question is whether facial analysis tools are valid for measuring Big Five personality in the same way as self-assessment questionnaires are (Fig. 3). Thus far, published studies have not provided convincing reasons to suggest so.

For instance, a study by Kachur et al. (2020) showed only modest results. They used psychometrically well-validated self-evaluation-based personality surveys as ground truths to test whether deep neural networks can infer Big Five personality by analysing videos or photographs. The result was that deep neural networks "can make a correct guess about the relative standing of two randomly chosen individu-

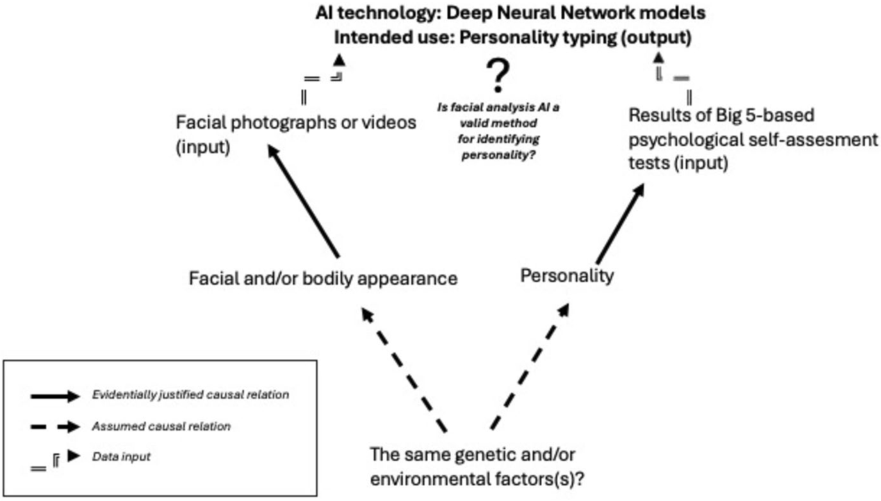

Fig. 3 Causal assumptions underlying the claim that deep neural network models could infer Big Five personality factors from facial photographs or video streams

als on a personality dimension in 58% of cases (as opposed to the 50% expected by chance)" (Kachur et al., 2020, 6). Despite this result, Kachur et al., (2020, 6) are hesitant to draw any definite conclusions and argue that "facial image-based personality assessment can hardly replace traditional personality measures" because there is not enough research on the validity of facial-image-based personality assessment.

Cai and Liu (2022) used a methodology similar to that of Kachur et al. (2020). They concluded that their results further "verify the feasibility and effectiveness of the automatic assessment method of Big Five personality traits based on individual facial video analysis" (Cai & Liu, 2022, 1). Yet, they noted that their sample size of 82 individuals cannot be used to generalise the results into real-world contexts where interviewees are not standing completely still during the interview (Cai & Liu, 2022). Further research is therefore needed to determine the validity of video and photo-based methods for inferring Big Five personality.

These open scientific debates have not prevented AI companies from marketing their solutions as science-based. For instance, the company Retorio claimed in 2021 that its "behavioral analytics AI" technology is based on "psychological science" and "measures personality and predicts future potential" (Roemmich et al., 2023; 7, 18). Retorio's claims were empirically questioned by investigative journalists Elisa Harlan and Oliver Schnuck. They revealed that Retorio's AI-based video interview platform assesses a job candidate's personality differently depending on whether the candidate is wearing glasses or sitting in front of a bookshelf. The results were also affected by the brightness of the video stream (Harlan & Schnuck, 2021).

The article by Harlan and Schnuck on Retorio exemplifies a familiar story in AI development. Investigative journalists and activists expose the failure of AI technologies to deliver on their promises of "objectivity", "fairness", and "nonbias", leading to a public backlash and attempts of AI technology companies to "fix" these failures

(Crawford, 2021).[20] However, we doubt there are enough journalists and activists to investigate all the cases in which AI firms selling their systems for high-risk uses do not disclose the details of their technologies or the results of their possible validity research. Therefore, the scientific community should also be able to evaluate facial analysis AI systems critically. However, a significant obstacle to this critical evaluation is corporate secrecy.

Corporate trade and industrial secrets prevent competing companies from employing their developed technologies (Carrier, 2010). In the context of AI-based recruitment and facial analysis AI, corporate secrecy means that companies rarely disclose (a) the details of facial analysis AI model's operation principles and code, (b) what kind of data is used to train these models, (c) the research companies have internally conducted to test their systems' construct, measurement, and predictive validity (Raghavan et al., 2020; Sánchez-Monedero et al., 2020). When companies do not disclose this information in public research reports or peer-reviewed journals, or when they do not expose their systems to external audits (Rhea et al., 2022), the scientific community cannot evaluate the validity of personality assessments based on facial analysis AI (Tippins et al., 2021). As a result, it is unclear to the scientific community, the general public, employees purchasing AI systems, and individual job seekers whether AI-based facial analysis and personality assessments can be reliably used to predict job performance.

Corporate secrecy does not prevent all research-based evaluations, however. Academic researchers can draw some conclusions from the materials companies disclose (Raghavan et al., 2020). For example, Roemmich et al. (2023) examined the marketing claims of 229 AI-rec-tech providers that conduct personality or emotional assessments of job applicants. Most of the marketing talk they analysed proclaims that the technologies can serve as valid measurement tools, even though research has only taken its first steps to determine whether it is possible to use deep neural networks to infer personality types in recruitment (Roemmich et al., 2023). This raises the worry of social pseudotechnology.

Suppose someday a scientific consensus is formed that deep neural networks or other AI technologies are reliable in inferring personality. Then, AI firms will have better grounds for their marketing claims if they demonstrate that their technology is based on these reliable measurement methods (Rhea et al., 2022; Sloane et al., 2022). Currently, the scientific question remains unresolved, and there are convincing empirical reasons to be sceptical of the validity of facial analysis AI in inferring personality traits. The technologies on the market are pseudotechnologies since there is insufficient evidence to assess whether they can achieve their purpose of inferring personality better than random chance. In other words, they fulfil the first criterion of our definition of pseudotechnology: they are algorithmic systems that are used to sustain, organise, or modify social relations, but it is not justified that they could achieve their purpose better than random chance through their operation principles.

[20] For instance, HireVue, one of the best-known AI-rec-tech companies, discontinued using facial analysis to assess personality and candidates' emotions after the Electronic Privacy Information Center (Epic) complained to the US Federal Trade Commission against HireVue (Epic, 2021).

When used in recruitment to make hiring decisions, that is, to organise social reality, these technologies also fulfil our first criterion for social pseudotechnologies: they are algorithmic systems that are used to sustain, organise, or modify social relations, but it is not justified that they could achieve their purpose better than random chance through their operation principles. This is because current evidence is insufficient to justify that they can function for the desired purpose in hiring. Facial analysis technologies need to be more reliable than they currently are: 58% accuracy in identifying personality combined with the fact that personality can predict only 0.1 or 0.2 of job performance do not count as strong evidence to favour facial analysis AI in a high-risk context such as hiring. This is true, especially when there are reasons to believe that the pseudotechnologies sold by companies are not even that accurate.

It should be noted that, in this case, pseudotechnologies' operation principles are not built on the causal assumptions of pseudoscientific theories and methods. The scientific research on the validity of video and photo-based methods for inferring Big Five personality is not pseudoscience. It is conducted within the scientific community, and its errors are open to criticism.[21] Many researchers might regard the hypothesis that personality could be inferred from facial appearance and other visual or auditory cues as implausible and see problems with the research on the issue. Nevertheless, this is a far cry from claiming that such research is pseudoscience. Therefore, in this case, one can see how the concept of (pseudo)technology is independent of the concept of pseudoscience, as we argued in Section 3.

To conclude, according to our definition, the companies' facial AI applications are pseudotechnologies when used to infer personality. Furthermore, they are social pseudotechnologies when the personality tests are used to sustain, organise, or modify social relations, for instance, in hiring. This is because the companies claim that their facial analysis AI systems for inferring personality are scientifically based, but they are not able to justify having sufficient measurement validity to achieve this purpose.

4.3 Recognising "Emotions" and "Cognitive States"

The emotion detection and recognition technology market was valued at $23.5 billion in 2022 and is expected to grow to $42.9 billion by 2027 (Markets & Market, 2023). AI facial analysis technologies that aim to recognise emotions and human cognitive states from video streams or facial photos comprise a significant branch of this industry. Emotion artificial intelligence (EAI) is already used in various personal, commercial, and public contexts (McStay, 2018). These include uses such as personnel selection, following worker performance, facilitating emotional self-understanding, targeted advertising, monitoring classroom atmosphere and learning, regulating

[21] We are not claiming that all applications of facial analysis AI are automatically pseudotechnologies. We are merely pointing out that the cases examined here are problematic and instances of (social) pseudotechnologies. Still, current research cannot yet determine whether it is impossible to infer Big Five traits from videos. Likewise, some have argued that facial AI analysis shows promise for diagnosing neurological diseases such as Alzheimer's disease and Bell's palsy (Yoonesi et al., 2025) and mental diseases such as depression, when combined with other diagnostic procedures and data (Ghafarfaraji, 2026).

behaviour in closed spaces such as prisons, and pre-screening air travellers for security (Heaven, 2020; McStay, 2020).

Many uses of EAI interfere with individuals' privacy (McStay, 2020). Furthermore, the EU AI Act classifies the use of AI systems that recognise emotions in workplaces and education as an unacceptable risk and prohibits such use (EU Regulation 2024/1689, 2024 article 5). All other uses of AI to recognise emotions are considered high-risk (EU Regulation 2024/1689, 2024 article 6, Annex III). These ethical and legal considerations underscore the importance of critically analysing facial analysis AI used for emotion recognition as a potential social pseudotechnology.

The IBorderCtrl system is an example of EAI in border security. The IBorderCtrl system was developed by computer scientists at Manchester Metropolitan University and various IT startups, with funding from the EU Horizon 2020 programme (Cordis, 2016). In 2018, the IBorderCtrl system was tested "in realistic conditions" on the borders of Greece, Hungary, and Latvia to give risk scores to non-EU citizens entering the European Union (EU Directorate-General for Research & Innovation, 2018). The risk score was based partly on a video interview where facial analysis AI was used for "deception detection" (O'Shea et al., 2018). The purpose was to detect "biomarkers of deceit" and analyse "the micro-gestures of travellers to figure out if the interviewee is lying" to help border guards "in spotting illegal immigrants, and so contribute to the prevention of crime and terrorism" (EU Directorate-General for Research & Innovation, 2018).

EAI systems such as IBorderCtrl are based on two assumptions (Crawford, 2021; McStay, 2018; Stark & Hoey, 2021). First, humans can universally determine an individual's inner emotional state from facial (micro)expressions. Second, EAI systems can recognise these inner emotional states from the inputs of video streams or facial photos. Therefore, the crucial question is whether these assumptions are empirically justified (Fig. 4).

The first assumption that humans can universally determine emotional states from (micro)expressions is the central hypothesis of universalist theories of emotion developed by psychologist Paul Ekman (e.g., Ekman & Friesen, 1971). The main theoretical assumption behind EAI technologies is Ekman's idea that faces convey universal emotions such as joy, anger, disgust, sadness, surprise, and fear (Crawford, 2021; McStay, 2020).

When it is assumed that EAI systems can recognise such universal emotional states and be used to assess the emotional impacts of advertisements, we get companies like Affectiva. Affectiva promises to "Decode Emotional Expressions in Real-Time" (Affectiva, 2024a) and proclaims that 90% of the world's largest advertisers currently use their facial AI services to infer emotions for targeted advertising (Affectiva, 2024b). FaceCode, the "most intelligent coding interview platform", claims to measure job candidates' interior attributes like engagement and "high-level thinking" during video interviews (Roemmich et al., 2023, 18). Big tech companies such as Microsoft, IBM, and Amazon have also offered their customers facial analysis to detect emotions (Heaven, 2020).

Recently, universalist theories of emotion have been increasingly called into question. In 2019, a team of psychologists presenting both sides of the universalist debate published a review of nearly a thousand scientific articles on emotion recognition and

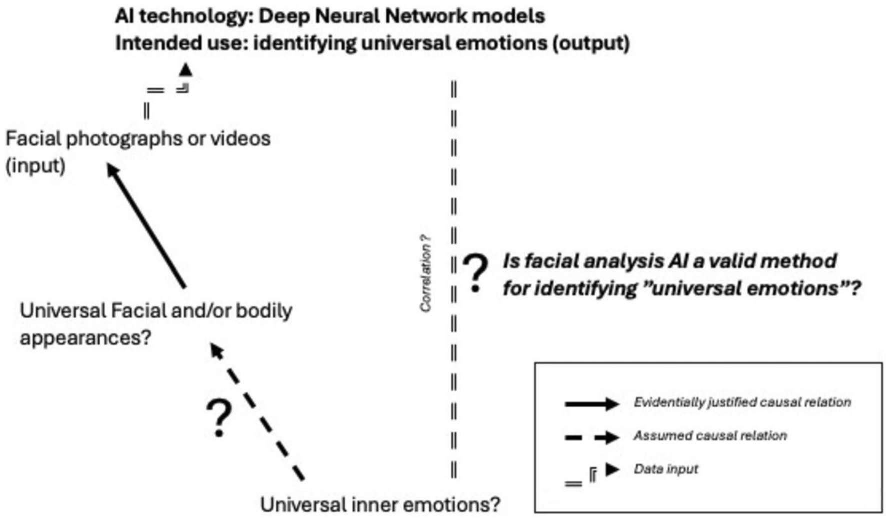

Fig. 4 Causal assumptions underlying the claim that AI models could be used to infer inner emotions from facial photographs or video streams

expression (Barrett et al., 2019; Heaven, 2020). They concluded that the expected standard expressions of emotions vary considerably culturally and contextually, and even across people in similar situations (Barrett et al., 2019; Crivelli et al., 2016; Le Mau et al., 2021). This means there is substantial variation in how people communicate emotions facially and bodily that, in English, are called "anger", "fear", "happiness" or "sadness" (Barrett et al., 2019). Since similar facial movements can express several emotions, there is little evidence that people can universally infer someone's emotional state from their facial movement (Barrett et al., 2019).

Overall, the assumption that universal emotions are expressed through facial movements seems to be an unfounded Western ethnocentric generalisation (Barrett, 2021). This assumption can lead to unjustified conclusions with deeply harmful discriminatory consequences. For instance, in Denmark, social services have used psychological "emotion recognition" tests to make foster care decisions (Bryant, 2025). In the test, parents were asked to identify emotions from the facial expressions of Danish psychologists (Social- og boligstyrelsen, 2025). In an ethnocentric manner, the Danish way of expressing emotions was interpreted as correct (Bryant, 2025). Some Inuit mothers interpreted the facial expressions of Danish psychologists in terms of how emotions are expressed in Inuit communities. This difference in interpreting facial expressions explains why some Inuit women did not pass the psychological test (Bryant, 2025). The test results were used to justify foster care decisions to remove children from their mothers. After public protests in May 2025, Danish authorities banned the discriminatory and unfounded use of "emotion recognition" tests to assess the "parenting skills" of Inuit people (Bryant, 2025; Social- og boligstyrelsen, 2025). However, the test is still used to assess the parenting skills of people of Danish and other ethnic origins (Social- og boligstyrelsen, 2025).

The example above and the review by Barret et al. (2019) suggest that datasets of human-labelled facial expressions (e.g., Cowen et al., 2021) cannot be used to train EAI models to recognise emotions reliably (Brooks et al., 2024; Stark & Hoey, 2021). Barrett et al., (2019, 48) conclude their review:

> Presently, many consumers of emotion research assume that certain questions about emotional expressions have been answered satisfactorily when in fact this is not the case. Technology companies, for example, are spending millions of research dollars to build devices to read emotions from faces, erroneously taking the common view as a fact that has strong scientific support. A more accurate description, however, is that such technology detects facial movements, not emotional expressions.

However, the commercial EAI industry is largely unregulated[22] and scientifically unvalidated. Even Paul Ekman has expressed scepticism. Ekman revealed in an interview to have written to some of the "biggest software companies in the world" and asked for evidence about their systems' validity (Heaven, 2020, 504). Ekman hit the same wall of corporate secrecy as journalists and other researchers, not receiving any response from the companies. Therefore, Ekman concludes: "As far as I know, they're making claims for things that there is no evidence for" (Heaven, 2020, 504).

There seems to be an almost universal consensus among psychologists that companies have no justification in claiming EAI systems can recognise emotional states from facial expressions. Nor is there any justification for claiming that "beyond mere emotional analysis, the capabilities of Human Emotion AI extend to broader people analytics, including cognitive states derived from facial signals" (Affectiva, 2024c). As the IBorderCtrl system used at European borders in 2018 was also based on the EAI technology of the time, it is improbable that it could have then functioned as a reliable lie detector.[23]

The facial analysis AI systems used to "recognise" emotions are pseudotechnologies according to the first criterion of our definition. They are claimed to be science-based but do not achieve their intended purpose of inferring emotions or other cognitive states (Barrett et al., 2019, 48; Roemmich et al., 2023; Crawford, 2021). If there are no universal facial expressions of emotion or other inner states, the facial analysis AI systems are based on operation principles that cannot work. When these pseudotechnologies are applied in social contexts such as border control or hiring, it is not justified that they can organise social relations better than random chance. Therefore, based on the first criterion of our definition, they are social pseudotechnologies.

[22] The EU "Artificial Intelligence Act" is currently the most extensive legal regulation and applies only in the EU (McStay & Bakir, 2025).

[23] See, for instance, Oh et al. (2018) for a review of the open scientific questions of automatic facial micro-expression analysis in the year IBorderCtrl was piloted.

4.4 Classifying "Race"

In 2019, IBM created a dataset of one million photos called "Diversity of Faces" (Merler et al., 2019). It was meant to tackle the problem that most facial recognition systems – including the ones used by Amazon, Microsoft, and IBM – were less accurate in recognising the faces of people with dark skin, especially women (Buolamwini & Gebru, 2018). IBM decided to correct this bias and level the accuracy of its deep neural network used for facial recognition across different demographic groups. The company collected new data from a million photographs to cover the diversity of human faces and, for instance, measured the craniofacial distances of the faces in the images. (Merler et al., 2019.)

The IBM researchers responsible for the new dataset's classification methods write that their work is motivated by the following assumptions:

> Every face reflects something unique about us. Aspects of our heritage – including race, ethnicity, culture, geography – and our individual identity – age, gender and visible forms of self-expression – are reflected in our faces. (Merler et al., 2019, 1).

Merler et al., (2019, 9) continue that "skin color alone is not a strong predictor of race, and other features such as facial proportions are important". Therefore, the deep neural network used to measure the Diversity of Faces Dataset analyses skin tone, craniofacial distances, areas, ratios, facial symmetry, and facial region contrast. In its analysis of skin colour, it measures the pixels of facial photographs to determine the skin tone of the person who has been photographed. The outputs are classifications that categorise photographs into six individual typology angle (ITA) categories used in dermatology (Merler et al., 2019, 14). The dermatological ITA categories are "very light, light, intermediate, tan, brown, and dark" (Wilkes et al., 2015).

Furthermore, as several critics of facial analysis AI (Ajunwa, 2021; Crawford, 2021; Stark & Hutson, 2021) have noted, the assumption that race can be inferred from skin tone and facial morphology alone – as well as by measuring craniological distances, facial areas, and ratios – is similar to the assumptions and practices underlying nineteenth-century physiognomy and phrenology. Today, it is hard to find any biologist, anthropologist, psychologist, or social scientist who would consider this assumption scientific. On the contrary, the assumption that race is reduced to skin tone and facial features alone is commonly considered to lack evidential support (Hirschman, 2004; Scheuerman et al., 2021) and to be pseudoscientific (Reid-Merritt, 2017). Biologists have not found a population genetic basis for racial group categorisations (Cartmill, 1998; Templeton, 2013). Therefore, today, race and racial classifications are mainly studied as historical, cultural, social, political, and subjectively experienced phenomena (Omi & Winant, 2015).

Despite these false and pseudoscientific assumptions of "race" made by the developers, the deep neural network was used to statistically analyse skin colour and other facial features to determine how diverse the Diversity of Faces Dataset is in relation to these features (Merler et al., 2019). Skin colour, as said, was classified into the dermatological categories "very light, light, intermediate, tan, brown, and dark" (Merler et al., 2019, 14). In other words, the developers motivate their research with a pseudoscientific theory of races. However, the classifications from this theory were not used to classify the data. Skin tone was. The pseudoscientific motivation is ethically dubious, and critics of the IBM research (Crawford, 2021; Scheuerman et al., 2021) have rightly criticised Merler et al. (2019) for introducing and perpetuating a pseudoscientific theory. Researchers should not use such theories to motivate their research.

Epistemically, however, the pseudoscientific racial theories could be removed from the IBM research report without impacting on the deep neural network used to measure the degree of diversity in the Diversity of Faces Dataset. The deep neural network model *did not predict or use as its input* racial classifications like "Caucasian" or "Black". Such categories are used to self-identify one's race, for instance, in US population studies (Monk, 2022). If the IBM scientists had produced these classifications based solely on facial appearance, we would label IBM's Diversity of Faces Dataset a clear example of pseudotechnology.[24] It would be an AI system designed for biogenetic racial classification. Moreover, it would not be justified as able to achieve this purpose, as there is no genetically-coded biological basis for grouping people into such categories (Scheuerman et al., 2021).

The neural deep network developed by Merler et al. (2019) can be argued to constitute a functional technology in the sense that it fulfils a specific purpose of interest to computer vision researchers. The purpose is to develop deep neural networks that produce accurate classifications of skin tones based on analysing pixel colours in facial photographs (Chen et al., 2016; Khanam et al., 2021). It has been argued that, in the context of fulfilling this specific purpose, various dermatological skin-type analyses (Chardon et al., 1991; Fitzpatrick, 1988) are proper scientific practices (Barrett et al., 2023). For instance, without the dermatological Fitzpatrick skin type categorisation, Buolamwini and Gebru (2018) would not have discovered that facial recognition AI systems are less accurate in recognising the faces of individuals with dark skin, especially women. Buolamwini and Gebru (2018, 6) regard the Fitzpatrick categorisation as providing "a scientifically based starting point for auditing algorithms". From a computer vision research perspective, IBM researchers (Merler et al., 2019) used the best available scientific methodologies by setting individual typology angle (ITA) classification (Chardon et al., 1991) as the output of their deep neural network. Given the specific goal of computer vision researchers to produce accurate classifications of skin tones based on analysing pixel colours in facial photographs, and the methodological development and success in fulfilling this goal (e.g., Bahmani et al., 2021; Buolamwini & Gebru, 2018; Celis & Keswani, 2020; Chang et al.,

[24] There are striking examples of such pseudotechnologies in computer vision research. For example, researchers at UCLA, Kärkkäinen and Joo (2021), published a paper titled "FairFace: Face Attribute Dataset for Balanced Race, Gender, and Age for Bias Measurement and Mitigation". The authors (2021, 1550) claim that "race is defined based on physical traits" and "consider Latino a race, which can be judged from the facial appearance." In the paper, facial photos are classified based on facial characteristics, such as skin colour, using the "race classification from the U.S. Census Bureau" (ibid., 1550). According to Google Scholar, the paper had been cited over 800 times in computer science publications by 2025. Kärkkäinen and Joo (2021, 1548) published their paper to promote "fairness" by addressing the underrepresentation of certain "races" in public image datasets. However, it is ethically and epistemically deeply problematic even to suggest that human races exist.

2018), the deep neural network by Merler et al. (2019) is not a pseudotechnology as it fulfils this specific goal the way it is supposed to.

The above does not mean that using dermatological skin type categories in the field of computer vision (or dermatology) is epistemologically and ethically unproblematic. First, there is a methodological problem. Skin type annotation and, in general, any type of colour categorisation, is inherently subjective (Bahmani et al., 2021; Barrett et al., 2023). For instance, crowdsourced workers do not systematically categorise pictures of people into different skin types in the same way (Barrett et al., 2023; Groh et al., 2022). This problem should be addressed, for instance, by providing detailed documentation of how the subjectivity of annotations has been taken into account in research (Bahmani et al., 2021; Barrett et al., 2023; Chang et al., 2018; Groh et al., 2022).

Second, dermatological skin tone classifications are biased towards lighter skin types (Buolamwini & Gebru, 2018). The original ITA (Chardon et al., 1991) and Fitzpatrick classifications (Fitzpatrick, 1988) have four categories for light skin types and only two for darker ones. The classification is discriminatory, since it prioritises whiteness and marginalises people with darker skin tones.

ITA categorisation can also be accused of racialising people based on their skin colour. ITA categorisation was initially intended solely for analysing the sensitivity of different skin tones to sunlight among people identified as caucasians (Chardon et al., 1991). Since "caucasian" is a demographic racial category, it seems that the developers of ITA conflated skin tone with demographic racial categories. Hence, ITA categorisations regenerate race-based classifications disguised as neutral skin-type analysis (Barrett et al., 2023).

Despite ITA classification prioritising whiteness and erroneously regenerating race-based classifications, it has been rather uncritically endorsed in computer vision research (Barrett et al., 2023). This is not uncommon in computer science. As several critical scholars have argued, computer scientists and AI developers often have adopted classifications and datasets of facial photos that have roots, for instance, in racial and gender discrimination and exclusion (Benjamin, 2019; Birhane & Guest, 2021; Crawford, 2021; Nieves Delgado, 2023; Scheuerman et al., 2021). Therefore, it is no wonder that IBM researchers suggested that race can be inferred from faces and conflated race with skin colour, as did the dermatologists who developed ITA categories.

There is one further ethical and social problem related to the use of skin-tone categorisations. Skin tone is a facial feature that people constantly use to classify others, and classifications based on skin tone are used to discriminate against people (Monk, 2022). Skin tone stratification or colourism is a phenomenon where the perceived lightness or darkness of a person's skin affects how a person is treated in various social contexts. It can lead to disparities in income and education based on skin tone (Keith & Herring, 1991). For instance, skin tone statistically explains socioeconomic inequality in Brazil even after considering membership in self-identified racial categories (Monk, 2016).

Because ITA classifications are connected with racial discrimination, marginalisation, and skin tone stratification, deep neural networks that classify people based on dermatological skin types are not ethically and politically innocent. As critics have argued, facial analysis AI datasets based on skin tone classifications play a role in maintaining, and perhaps even strengthening, existing social classifications, discrimination, and social injustice produced by the continuous use of these classifications (Ajunwa, 2021; Benjamin, 2019; Crawford, 2021; Stark & Hutson, 2021). The fact that skin-tone-based classifications of facial AI systems can play a causal role in producing discrimination and social injustice does not, in itself, make the deep neural network that produces the classifications a pseudotechnology. In other words, a facial analysis AI system does not need to be a pseudotechnology to reinvigorate "racism at an unprecedented scale" (Stark & Hutson, 2021, 6).

5 Conclusions

In this paper, we have argued that the concept of pseudotechnology, rather than pseudoscience, provides a more accurate tool for analysing the epistemic problems with AI facial analysis. Our case studies on facial AI systems reveal that not all of them should be automatically dismissed as physiognomic pseudoscience. This conclusion is based on the argument that the concept of pseudotechnology is independent of the concept of pseudoscience. For instance, computer vision researchers can conduct proper scientific research to determine whether facial analysis AI can infer Big Five personality traits. Nonetheless, the current practical applications claimed to be based on such research are social pseudotechnologies. In other words, the core issue is not whether a facial analysis AI system is pseudoscientific, but whether sufficient evidence establishes that it functions in the way it is intended to function. Therefore, by applying the concepts of pseudotechnology and social pseudotechnology, we can sharpen and deepen previous epistemic critiques of facial AI applications.

Current facial analysis AI applications reveal that pseudotechnologies – especially social pseudotechnologies – are far more prevalent than previously acknowledged in the philosophical literature. One reason for this is that social pseudotechnologies can generate reactivity. Their effectiveness can sometimes arise not from genuine technological capability but from how humans respond to the classifications. For instance, labelling someone as a criminal can, in some cases, contribute to behaviours that reinforce this classification. Reactivity can, therefore, explain why social pseudotechnologies are so challenging to detect. Their widespread use may persist not because users deliberately endorse epistemically questionable technologies but because these technologies often operate in ways that make their true nature harder to discern.

If we are to confront also the moral and societal hazards posed by facial analysis AI applications, especially in high-risk contexts, we must first recognise how these systems sustain themselves and why their influence remains so inconspicuous. The concept of social pseudotechnology is, therefore, not only helpful for computer scientists, critical data scholars, and policymakers but also to organisational decision-makers and the general public. In addition to evaluating the ethical risks associated with various uses of technologies, all these audiences also need conceptual tools to demarcate social technologies from social pseudotechnologies. As our case studies show, the concept of social pseudotechnology equips them with such a tool.

Acknowledgements We are grateful to Inkeri Koskinen and Kaisa Luoma for their help in developing the concept of pseudotechnology. We also thank Marko Forssell, Magdalena Małecka, Jani Raerinne, and Petri Ylikoski for their detailed comments on the previous versions of the manuscript. We have also received insightful audience comments at European Network for the Philosophy of the Social Sciences (ENPOSS) conference (Bergen, 2024), XXV World Congress of Philosophy (Rome, 2024), and TINT – Centre for the Philosophy of Social Sciences seminar (Helsinki, 2025).

Author contributions Päivi Seppälä developed the original idea for the article. Both authors contributed equally to the analysis of literature and philosophical conceptual analysis. Both authors contributed equally to writing of the article. Both authors have read and approved the final manuscript.

Funding Open Access funding provided by University of Helsinki (including Helsinki University Central Hospital). Päivi Seppälä's research was funded by the Finnish Work Environment Fund (grant number 250637), Kone Foundation (grant number 2021104708), Centria University of Applied Sciences, and the University of Helsinki's three-year research project "From cyborg origins of modern economics to its automated future. Towards a new philosophy of economics". Ilmari Hirvonen received funding from the research project "Pseudoscience in Finnish working life" funded by the Emil Aaltonen Foundation.

Data availability Not applicable.

Declarations

Financial interests The authors have no relevant financial interests to disclose.

Non-financial interests The authors have no competing non-financial interests to declare that are relevant to the content of this article.

Competing interests None.

References

Affectiva (2024a). Human Behavioral Research with iMotions. Retrieved May 17, 2024, from https://www.affectiva.com/product/imotions-behavioral-research/

Affectiva (2024b). Humanizing technology with emotion AI. Retrieved May 17, 2024, from https://www.affectiva.com/.

Affectiva (2024c). The Science Behind Emotion AI How It Works. Retrieved May 17, 2024, from https://www.affectiva.com/science-overview/.

Agüera y Arcas, B., Mitchell, M., & Todorov, A. (2017). Physiognomy's New Clothes. Medium. Retrieved May 17, 2024, from https://medium.com/@blaisea/physiognomys-new-clothes-f2d4b59fdd6a

Aguëra y Arcas, B., Mitchell, M., & Todorov, A. (2023). Physiognomy in the age of AI. In J. Browne, S. Cave, E. Drage, & K. McInerney (Eds.) *Feminist AI: Critical Perspectives on Algorithms, Data, and Intelligent Machines* (pp. 208–236). Oxford University Press.

Ajunwa, I. (2021). Automated video interviewing as the new phrenology. *Berkeley Technology Law Journal, 36*(1), 101–154.

Andrews, M., Smart, A., & Birhane, A. (2024). The reanimation of pseudoscience in machine learning and its ethical repercussions. *Patterns, 5*, 1–14. https://doi.org/10.1016/j.patter.2024.101027

Azucar, D., Marengo, D., & Settanni, M. (2018). Predicting the Big 5 personality traits from digital footprints on social media: A meta-analysis. *Personality and Individual Differences*. 124, (1 April 2018), 150–159.

Bahmani, K., Plesh, R., Sahu, C., Banavar, M., & Schuckers, S. (2021). SREDS: A dichromatic separation based measure of skin color. *IEEE International Workshop on Biometrics and Forensics (IWBF), 2021*, 1–6. https://doi.org/10.1109/iwbf50991.2021.9465090

Barrett, L. F. (2021). AI weighs in on debate about universal facial expressions. *Nature*, *589*(7841), 202–203. https://doi.org/10.1038/d41586-020-03509-5

Barrett, L. F., Adolphs, R., Marsella, S., Martinez, A. M., & Pollak, S. D. (2019). Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements. *Psychological Science in the Public Interest, 20*(1), 1–68.

Barrett, T., Chen, Q., & Zhang, A. (2023). Skin Deep: Investigating Subjectivity in Skin Tone Annotations for Computer Vision Benchmark Datasets. *ACM Conference on Fairness Accountability and Transparency, 2023*, 1757–1771. https://doi.org/10.1145/3593013.3594114

Barrick, K. (2014). A Review of Prior Tests on Labeling Theory. In D. P. Farrington & J. Murray (Eds.), *Labeling theory: Empirical tests* (pp. 95–113). Taylor & Francis Group.

Becker, H. (1963). *Outsiders. Studies in the sociology of defiance.* The free press.

Benjamin, R. (2019). *Race after technology: abolitionist tools for the New Jim Code*. Polity Press.

Bernburg, J. G. (2009). Labeling theory. In M. D. Krohn, A. J. Lizotte, & G. P. Hall (Eds.), *Handbook on crime and deviance* (pp. 179–196). Springer.

Bernburg, J. G., Krohn, M. D., & Rivera, C. J. (2006). Official labeling, criminal embeddedness, and subsequent delinquency. *Journal of Research in Crime and Delinquency, 43*(1), 67–88. https://doi.org/10.1177/0022427805280068

Birhane, A., & Guest, O. (2021). Towards decolonising computational sciences. *Kvinder, Køn & Forskning, 2*, 60–73. https://doi.org/10.7146/kkf.v29i2.124899

Boden, M. (2006). *Mind as Machine: A History of Cognitive Science, Volume 1&2*. Oxford University Press.

Brien, S., Lachance, L., Prescott, P., McDermott, C., & Lewith, G. (2011). Homeopathy has clinical benefits in rheumatoid arthritis patients that are attributable to the consultation process but not the homeopathic remedy: A randomized controlled clinical trial. *Rheumatology, 50*(6), 1070–1082.

Brooks, H. (1980). Technology, evolution, and purpose. *Daedalus, 109*(1), 65–81.

Brooks, J. A., Kim, L., Opara, M., Keltner, D., Fang, X., Monroy, M., Corona, R., Tzirakis, P., Baird, A., Metrick, J., Taddesse, N., Zegeye, K., & Cowen, A. S. (2024). Deep learning reveals what facial expressions mean to people in different cultures. *Iscience, 27*(3), 109175. https://doi.org/10.1016/j.isci.2024.109175

Bryant, M. (2025). 'The nurse told me I couldn't keep my baby': how a controversial Danish 'parenting test' separated a Greenlandic woman from her children. *Guardian,* (June 29, 2025). https://www.theguardian.com/lifeandstyle/2025/jun/29/controversial-danish-parenting-test-separated-greenlandic-mother-children

Bunge, M. (1976). The philosophical richness of technology. *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association, Volume Two: Symposia and Invited Papers, 1976, 2*, 153–172. https://doi.org/10.1086/psaprocbienmeetp.1976.2.192379

Buolamwini, J., & Gebru, T. (2018). Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of the 1st Conference on Fairness, Accountability and Transparency, PMLR, 81*, 77–91.

Cai, L., & Liu, X. (2022). Identifying Big Five personality traits based on facial behavior analysis. *Frontiers in Public Health* (9 September 2022), 1–11.

Carrier, M. (2010). Research under Pressure. Methodological Features of Commercialized Science. In Radder, H. (ed). *The Commodification of Academic Research*. (pp. 158–186). University of Pittsburgh Press.

Cartmill, M. (1998). The status of the race concept in physical anthropology. *American Anthropologist, 100*(3), 651–660.

Celis, L. E., & Keswani, V. (2020). Implicit diversity in image summarization. *Proceedings of the ACM on Human-Computer Interaction, 4*(CSCW2), 1–28.

Chang, C.-C., Hsing, S.-T., Chuang, Y.-C., Wu, C.-T., Fang, T.-J., Choi, B. & Chen, K.-F. (2018). Robust Skin Type Classification Using Convolutional Neural Networks. *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, 2011–2014. https://doi.org/10.1109/iciea.2018.8398040

Chardon, A., Cretois, I., & Horses, C. (1991). Skin colour typology and suntanning pathways. *International Journal of Cosmetic Science, 13*(4), 191–208. https://doi.org/10.1111/j.1467-2494.1991.tb00561.x

Chen, W., Wang, K., Jiang, H., & Li, M. (2016). Skin color modeling for face detection and segmentation: A review and a new approach. *Multimedia Tools and Applications, 75*(2), 839–862. https://doi.org/10.1007/s11042-014-2328-0

Churchman, C. W. (1948). Statistics, pragmatics, induction. *Philosophy of Science, 15*, 249–268.

Coalition for Critical Technology (2020). Abolish the #TechToPrisonPipeline. Medium. Retrieved March 7, 2025, from https://medium.com/@CoalitionForCriticalTechnology/abolish-the-techtoprisonpipeline-9b5b14366b16

CORDIS - EU Research Results (2016). Intelligent Portable Border Control System. Retrieved December 11, 2024, from https://cordis.europa.eu/project/id/700626

Cowan, R. S. (1972). Francis Galton's statistical ideas: The influence of eugenics. *Isis, 63*(4), 509–528.

Cowen, A. S., Keltner, D., Schroff, F., Jou, B., Adam, H., & Prasad, G. (2021). Sixteen facial expressions occur in similar contexts worldwide. *Nature, 589*(7841), 251–257. https://doi.org/10.1038/s41586-020-3037-7

Crawford, K. (2021). *The Atlas of AI*. Yale University Press.

Criado Perez, C. (2019). *Invisible Women: Exposing Data Bias in a World Designed for Men*. Abrams Press.

Crivelli, C., Jarillo, S., Russell, J. A., & Fernández-Dols, J. M. (2016). Reading emotions from faces in two indigenous societies. *Journal of Experimental Psychology: General, 145*, 830–843. https://doi.org/10.1037/xge0000172

Derksen, M., & Beaulieu, A. (2011). Social Technology. In I. C. Jarvie & J. Zamora-Bonilla (Eds.), *The SAGE Handbook of The Philosophy of Social Sciences* (pp. 703–720). Sage.

Douglas, H. (2000). Inductive risk and values in science. *Philosophy of Science, 67*, 559–579.

Douglas, H. (2009). *Science, Policy, and the Value-Free Ideal*. Pittsburgh University Press.

Egan, F. (2017). Function-Theoretic Explanation and the Search for Neural Mechanisms. In D. M. Kaplan (Ed.), *Explanation and Integration in Mind and Brain Science* (pp. 145–163). Oxford University Press.

Ekman, P., & Friesen, W. V. (1971). Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology, 17*(2), 124–129.

EPIC (2021). HireVue, Facing FTC Complaint From EPIC, Halts Use of Facial Recognition. Retrieved March 7, 2025, from https://epic.org/hirevue-facing-ftc-complaint-from-epic-halts-use-of-facial-recognition/.

Ernst, E. (2002). A systematic review of systematic reviews of homeopathy. *British Journal of Clinical Pharmacology, 54*(6), 577–582.

EU Directorate-General for Research and Innovation (2018). Smart lie-detection system to tighten EU's busy borders. Retrieved December 11, 2024, from https://projects.research-and-innovation.ec.europa.eu/en/projects/success-stories/all/smart-lie-detection-system-tighten-eus-busy-borders.

EU Regulation 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (Artificial Intelligence Act). Retrieved December 4 April, 2025, from https://eur-lex.europa.eu/eli/reg/2024/1689/oj/eng

Faception (2024a). Faception - facial personality analytics. Retrieved April 24, 2024, from https://www.faception.com/

Faception (2024b). Theory behind the technology. Retrieved April 24, 2024, from https://www.faception.com/about-us

Fitzpatrick, T. B. (1988). The validity and practicality of sun-reactive skin types I through VI. *Archives of Dermatology, 124*(6), 869–871. https://doi.org/10.1001/archderm.124.6.869

Friedler, S. A., Scheidegger, C., & Venkatasubramanian, S. (2021). The (Im)possibility of fairness. *Communications of the ACM, 64*(4), 136–143.

Fu, S., He, H., & Hou, Z. G. (2014). Learning race from face: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 36*(12), 2483–2509.

Ghafarfaraji, S. (2026). AI-based recognition of facial and micro-expressions for the diagnosis of mental and neurological disorders: A systematic review. *BMC Psychiatry, 26*, 78. https://doi.org/10.1186/s12888-025-07739-7

Goldenfein, J. (2019) The profiling potential of computer vision and the challenge of computational empiricism. *FAT* '19: Proceedings of the Conference on Fairness, Accountability, and Transparency,* 110–119. https://doi.org/10.1145/3287560.3287568.

Groh, M., Harris, C., Daneshjou, R., Badri, O., & Koochek, A. (2022). Towards transparency in dermatology image datasets with skin tone annotations by experts, crowds, and an algorithm. *Proceedings of the ACM on Human-Computer Interaction, 6*(CSCW2), 1–26. https://doi.org/10.1145/3555634

Guest, O. (2024). What makes a good theory, and how do we make a theory good? *Computational Brain & Behavior, 7*(4), 508–522. https://doi.org/10.1007/s42113-023-00193-2

Guest, O. (2025). What does "Human-Centred AI" mean? *arXiv*. https://doi.org/10.48550/arxiv.2507.19960

Guest, O., & Martin, A. E. (2021). How computational modeling can force theory building in psychological science. *Perspectives on Psychological Science, 16*(4), 789–802. https://doi.org/10.1177/1745691620970585

Guest, O., & Martin, A. E. (2023). On logical inference over brains, behaviour, and artificial neural networks. *Computational Brain & Behavior, 6*(2), 213–227. https://doi.org/10.1007/s42113-022-00166-x

Hansson, S.O. (2015). Experiments Before Science. What Science Learned from Technological Experiments. In: Hansson, S. (eds) *The Role of Technology in Science: Philosophical Perspectives*. (pp. 81–110). Springer. https://doi.org/10.1007/978-94-017-9762-7_5.

Hansson, S. O. (2017). Science denial as a form of pseudoscience. *Studies in History and Philosophy of Science, 63*(2017), 39–47.

Hansson, S. O. (2019). Farmers' experiments and scientific methodology. *European Journal for Philosophy of Science, 9*, 32. https://doi.org/10.1007/s13194-019-0255-7

Hansson, S. O. (2020). With all this pseudoscience, why so little pseudotechnology? *Axiomathes, 30*(6), 685–696.

Harlan, E. & Schnuck, O. (2021). Objective or biased? BR24 (Bayerische Rundfunk), February 21, 2021. Retrieved March 7, 2025, from https://interaktiv.br.de/ki-bewerbung/en/

Hashemi, M., & Hall, M. (2020). RETRACTED ARTICLE: Criminal tendency detection from facial images and the gender bias effect. *Journal of Big Data, 7*(2020), article number 2.

Hashemi, M., & Hall, M. (2020). Retraction Note: Criminal tendency detection from facial images and the gender bias effect. *Journal of Big Data, 7*(2020), article number 40.

Heaven, D. (2020). Why faces don't always tell the truth about feelings. *Nature, 578*(27 February 2020), 502–504.

Hempel, C. G. (1965). Science and Human Values. In *Aspects of Scientific Explanation and Other Essay in the Philosophy of Science*. (pp. 81–96). The Free Press.

Hirschman, C. (2004). The origins and demise of the concept of race. *Population and Development Review, 30*(3), 385–415.

Howick, J., & Hoffmann, T. (2018). How placebo characteristics can influence estimates of intervention effects in trials. *CMAJ : Canadian Medical Association Journal, 30*(190), E908–E911.

Jacobs, A. Z., & Wallach, H. (2021). Measurement and Fairness. *FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 375–385. https://doi.org/10.1145/3442188.3445901

Jones, O. P., Alfaro-Almagro, F., & Jbabdi, S. (2018). An empirical, 21st century evaluation of phrenology. *Cortex, 106*(1), 26–35. https://doi.org/10.1016/j.cortex.2018.04.011

Kachur, A., Osin, E., Davydov, D., Shutilov, K., & Novokshonov, A. (2020). Assessing the Big Five personality traits using real-life static facial images. *Scientific Reports, 10*(1), 8487. https://doi.org/10.1038/s41598-020-65358-6

Kaplan, H. B., Johnson, R. J., & Bailey, C. A. (1991). Negative social sanctions and juvenile delinquency: Effects of labeling in a model of deviant behavior. *Social Science Quarterly, 71*(1), 98–122.

Keith, V. M., & Herring, C. (1991). Skin tone and stratification in the Black community. *American Journal of Sociology, 97*(3), 760–778. https://doi.org/10.1086/229819

Khanam, R., Johri, P., & Diván, M. J. (2021). Human Skin Color Detection Technique Using Different Color Models. In P. Johri, M. J. Diván, R. Khanam, M. Marciszack & A. Will (eds.) *Trends and Advancements of Image Processing and Its Applications*. (pp. 261–279). Springer. https://doi.org/10.1007/978-3-030-75945-2_14.

Kärkkainen, K., & Joo, J. (2021). Fairface: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision* (pp. 1548–1558).
Lee, W. (2024). Revisiting labeling theory: empirical test of informal labeling process. *Journal of Crime and Justice*, *47*(2), 188–200. https://doi.org/10.1080/0735648X.2023.2264263
Le Mau, T., Hoemann, K., Lyons, S. H., Fugate, J. M. B., Brown, E. N., Gendron, M., & Barrett, L. F. (2021). Professional actors demonstrate variability, not stereotypical expressions, when portraying emotional states in photographs. *Nature Communications, 12*(1), 5037. https://doi.org/10.1038/s41467-021-25352-6
Lemert, E. M. (1967). *Human deviance, social problems, and social control*. Prentice Hall.
Lichtenberg G. C. (1778). On Physiognomy; Against the Physiognomists: For the Promotion of Human Knowledge and the Love of Humanity (1778). German History Intersections. Retrieved December 05, 2024 from https://germanhistory-intersections.org/en/knowledge-and-education/ghis:document-19.
Macrae, C. N., & Shepherd, J. W. (1989). Do criminal stereotypes mediate juridic judgements? *British Journal of Social Psychology, 28*(2), 189–191.
Mahner, M. (2007). Demarcating Science from Non-Science. In T. Kuipers (Ed.), *Handbook of the Philosophy of Science: General Philosophy of Science: Focal Issues* (pp. 515–575). Elsevier.
Marchionni, C., Zahle, J., & Godman, M. (2024). Reactivity in the human sciences. *European Journal for Philosophy of Science, 14*(1), Article 8. https://doi.org/10.1007/s13194-024-00571
Markets and Market (2023). Emotion Recognition and Detection (ERD) market. Retrieved March 10, 2025, from https://www.marketsandmarkets.com/Market-Reports/emotion-detection-recognition-market-23376176.html.
Matsueda, R. L. (1992). Reflected appraisals, parental labeling, and delinquency: Specifying a symbolic interactionist theory. *American Journal of Sociology, 97*(6), 1577–1611.
McStay A. (2018). *Emotional AI: The Rise of an Empathetic Media*. Sage.
McStay, A. (2020). Emotional AI, soft biometrics and the surveillance of emotional life: An unusual consensus on privacy. *Big Data & Society, 7*(1), 922–961.
McStay, A., & Bakir, V. (2025). Soft law for unintentional empathy: Addressing the governance gap in emotion-recognition AI technologies. *Journal of Responsible Technology, 23*, 100126. https://doi.org/10.1016/j.jrt.2025.100126
Merler, M., Ratha, N., Feris, R. S., & Smith, J. R. (2019). Diversity in faces. *arXiv*. https://doi.org/10.48550/arxiv.1901.10436
Miller, M. B., Horn, J. D. V., Wolford, G. L., Handy, T. C., Valsangkar-Smyth, M., Inati, S., Grafton, S., & Gazzaniga, M. S. (2002). Extensive individual differences in brain activations associated with episodic retrieval are reliable over time. *Journal of Cognitive Neuroscience, 14*(8), 1200–1214. https://doi.org/10.1162/089892902760807203
Monk, E. P. (2016). The consequences of "Race and Color" in Brazil. *Social Problems, 63*(3), 413–430.
Monk, E. P. (2022). Inequality without groups: Contemporary theories of categories, intersectional typicality, and the disaggregation of difference. *Sociological Theory, 40*(1), 3–27.
Monteiro, S. (2023). Gaming faces: Diagnostic scanning in social media and the legacy of racist face analysis. *Information, Communication & Society, 26*(8), 1601–1617.
Morrison, M. & Morgan, M. S. (1999). Models as mediating instruments. In M. S. Morgan & M. Morrison (eds.) *Models as Mediators: Perspectives on Natural and Social Science*. Ideas in Context. (pp. 10–37). Cambridge University Press.
Nieves Delgado, A. (2023). Race and statistics in facial recognition: Producing types, physical attributes, and genealogies. *Social Studies of Science, 53*(6), 916–937. https://doi.org/10.1177/03063127221127666
Oh, Y.-H., See, J., Ngo, A. C. L., Phan, R.C.-W., & Baskaran, V. M. (2018). A survey of automatic facial micro-expression analysis: Databases, methods, and challenges. *Frontiers in Psychology, 9*, 1128. https://doi.org/10.3389/fpsyg.2018.01128
Omi, M., & Winant, H. (2015). *Racial Formation in the United States*. Third edition. Routledge.
O'Shea J., Crockett K., Khan W., Kindynis P, Antoniades A., and Boultadakis G. (2018). Intelligent Deception Detection through Machine Based Interviewing. *International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, 2018* (pp. 1–8) https://doi.org/10.1109/IJCNN.2018.8489392
Pasquale, F. (2018). When machine learning is facially invalid. *Communications of the ACM, 61*(9), 25–27.
Pfaffenberger, B. (1988). Fetishised objects and humanised nature: Towards an anthropology of technology. *Man, 23*(2), 236–252. https://doi.org/10.2307/2802804

Raghavan, M., Barocas, S., Kleinberg, J. & Levy, K (2020). Mitigating bias in algorithmic hiring. *ACM Conference on Fairness, Accountability, and Transparency (FAT*)*, 2020, (pp. 469–481). https://doi.org/10.1145/3351095.3372828

Reid-Merritt, P (2017). *Race in America: How a Pseudoscientific Concept Shaped Human Interaction*. Praeger.

Rhea, A. K., Markey, K., D'Arinzo, L., et al. (2022). An external stability audit framework to test the validity of personality prediction in AI hiring. *Data Mining and Knowledge Discovery, 36*(1), 2153–2193.

Roemmich, K., Rosenberg, T., Fan, S., & Andalibi, N. (2023). Values in emotion Artificial Intelligence hiring services: Technosolutions to organizational problems. *Proceedings of the ACM on Human-Computer Interaction, 7*(Issue CSCW1), 109. https://doi.org/10.1145/3579543

Rudner, R. (1953). The scientist *qua* scientist makes value judgments. *Philosophy of Science, 20*, 1–6.

Scheuerman, M. K., Pape, M., & Hanna, A. (2021). Auto-essentialization: Gender in automated facial analysis as extended colonial project. *Big Data & Society, 8*(2), 1–15. https://doi.org/10.1177/20539517211053712

Schon, D. A. (1967). *Technology and change*. Delacorte Press.

Seppälä, P., & Małecka, M. (2024). AI and discriminative decisions in recruitment: Challenging the core assumptions. *Big Data & Society, 11*(1), 1–15. https://doi.org/10.1177/20539517241235872

Sloane, M., Moss, E., & Chowdhury, R. (2022). A Silicon Valley love triangle: Hiring algorithms, pseudo-science, and the quest for auditability. *Patterns, 3*(2), 1–9.

Small, A. W. (1898). Seminar notes: The methodology of the social problem. Division I. The sources and uses of material. *American Journal of Sociology, 4*(1), 113–144.

Sánchez-Monedero, J., Dencik, L. & Edwards, L. (2020). What does it mean to "solve" the problem of discrimination in hiring? *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency*, 458–468. https://doi.org/10.1145/3351095.3372849

Social- og Boligstyrelsen (2025). Retningslinjer for udarbejdelse og anvendelse af forældrekompetence undersøgelse, 3. udgave. https://www.sbst.dk/Media/638821193571490312/Retningslinjer%20for%20udarbejdelse%20og%20anvendelse%20af%20forldrekompetenceundersgelse%20(3.%20udgave).pdf

Stark, L. (2019). Facial recognition is the plutonium of AI. *XRDS: Crossroads, the ACM Magazine for Students, 25*(3), 50–55. https://doi.org/10.1145/33131

Stark, L. & Hoey, J. (2021). The ethics of emotion in artificial intelligence systems. *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 782–793. https://doi.org/10.1145/3442188.3445939

Stark, L., & Hutson, J. (2021). Physiognomic Artificial Intelligence. *Fordham Intellectual Property, Media & Entertainment Law Journal, 32*(4), 922–961.

Stinson, C. (2020). Algorithms associating appearance and criminality have a dark past. *Aeon*. Retrieved March 10, 2025, from https://aeon.co/ideas/algorithms-associating-appearance-and-criminality-have-a-dark-past

Sublime, J. (2024). The return of pseudosciences in artificial intelligence: Have machine learning and deep learning forgotten lessons from statistics and history? *arXiv*. https://doi.org/10.48550/arxiv.2411.18656

Tai, C., Wang, Y. & Hong, S. How China is using artificial intelligence in classrooms. *The Wall Street Journal* (YouTube news video 1. October 2019), https://www.youtube.com/watch?v=JMLsHI8aV0g

Templeton, A. R. (2013). Biological races in humans. *Studies in History and Philosophy of Science Part c: Studies in History and Philosophy of Biological and Biomedical Sciences, 44*(3), 262–271. https://doi.org/10.1016/j.shpsc.2013.04.010

Tippins, N., Oswald, F. & McPhail, S. M. (2021). Scientific, legal, and ethical concerns about ai-based personnel selection tools: A call to action. *Personnel Assessment and Decisions*, *7*(2). https://doi.org/10.25035/pad.2021.02.001

Todorov, A., Olivola, C. Y., Dotsch, R., & Mende-Siedlecki, P. (2015). Social attributions from faces: Determinants, consequences, accuracy, and functional significance. *Annual Review of Psychology, 66*(1), 519–545.

Treem, J., van Zoonen, W., & Sivunen, A. (2024). Examining communication visibility and social technology platform use in organizations. *New Media & Society, 26*(5), 2633–2658.

Tuomela, R. (1987). Science, Protoscience, and Pseudoscience. In J. C. Pitt & M. Pera (Eds.), *Rational Changes in Science* (pp. 83–101). D. Reidel Publishing Company.

Ullstein, C., Engelmann, S., Papakyriakopoulos, O., Hohendanner, M., & Grossklags, J. (2022). AI-Competent Individuals and Laypeople Tend to Oppose Facial Analysis AI. *EAAMO '22: Proceedings of the 2nd ACM Conference on Equity and Access in Algorithms, Mechanisms, and Optimization,* pp. 1–12. https://doi.org/10.1145/3551624.3555294

van Atteveldt, N., Janssen, T. W. P., & Davidesco, I. (2020). Measuring brain waves in the classroom. *Frontiers for Young Minds, 8*(1), 96. https://doi.org/10.3389/frym.2020.00096

Van Noorden, R. V. (2020). The ethical questions that haunt facial-recognition research. *Nature, 587*(19 November 2020), 354–358.

van Rooij, I., Guest, O., Adolfi, F., Haan, R., Kolokolova, A., & Rich, P. (2024). Reclaiming AI as a theoretical tool for cognitive science. *Computational Brain & Behavior, 7*(4), 616–636. https://doi.org/10.1007/s42113-024-00217-5

Walton, D. (2003). *Relevance in Argumentation*. Lawrence Erlbaum Associates.

Wang, Y., & Kosinski, M. (2018). Deep Neural Networks are more accurate than humans at detecting sexual orientation from facial images. *Journal of Personality and Social Psychology, 114*(2), 246–257. https://doi.org/10.1037/pspa0000098

Whittaker, M, Crawford, K., Dobbe, R., et al. (2018). *AI Now Report 2018*. AI Now. Retrieved March 22, 2022, from ainowinstitute.org/AI_Now_2018_Report.pdf

Wilkes, M., Wright, C. Y., du Plessis, J. L., & Reeder, A. (2015). Fitzpatrick skin type, individual typology angle, and melanin index in an African population: Steps toward universally applicable skin photosensitivity assessments. *JAMA Dermatology, 151*(8), 902–903. https://doi.org/10.1001/jamadermatol.2015.0351

Winston, A. S. (2020). Scientific racism and North American psychology. in I. Johnsruhe (ed.) *Oxford Research Encyclopedia of Psychology*. Oxford University Press. https://doi.org/10.1093/acrefore/9780190236557.013.516

Wu, X., & Zhang, X. (2016a). Automated inference on criminality using face images. *arXiv*. https://doi.org/10.48550/arXiv.1611.04135

Wu, X., & Zhang, X (2016b). Responses to critiques on machine learning of criminality perceptions (Addendum of arXiv:1611.04135). arXiv (2016). https://doi.org/10.48550/arxiv.1611.04135

Yoonesi, S., Abedi Azar, R., Arab Bafrani, M., et al. (2025). Facial expression deep learning algorithms in the detection of neurological disorders: A systematic review and meta-analysis. *BioMedical Engineering OnLine, 24*, 64. https://doi.org/10.1186/s12938-025-01396-3

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15*(6), 603–623.

Zell, E., & Lesick, T. L. (2022). Big five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality, 90*(4), 559–573.

Zunhammer, M., Goltz, G., Schweifel, M., Stuck, B. A., & Bingel, U. (2022). Savor the flavor: A randomized double- blind study assessing taste- enhanced placebo analgesia in healthy volunteers. *Clinical and Translational Science, 15*(11), 2709–2719.